Mathematical
Surveys
and
Monographs

Volume 86

Number Theoretic Density and Logical Limit Laws

Stanley N. Burris

American Mathematical Society

http://dx.doi.org/10.1090/surv/086

Selected Titles in This Series

86 **Stanley N. Burris,** Number theoretic density and logical limit laws, 2001

85 **V. A. Kozlov, V. G. Maz′ya, and J. Rossmann,** Spectral problems associated with corner singularities of solutions to elliptic equations, 2001

84 **László Fuchs and Luigi Salce,** Modules over non-Noetherian domains, 2001

83 **Sigurdur Helgason,** Groups and geometric analysis: Integral geometry, invariant differential operators, and spherical functions, 2000

82 **Goro Shimura,** Arithmeticity in the theory of automorphic forms, 2000

81 **Michael E. Taylor,** Tools for PDE: Pseudodifferential operators, paradifferential operators, and layer potentials, 2000

80 **Lindsay N. Childs,** Taming wild extensions: Hopf algebras and local Galois module theory, 2000

79 **Joseph A. Cima and William T. Ross,** The backward shift on the Hardy space, 2000

78 **Boris A. Kupershmidt,** KP or mKP: Noncommutative mathematics of Lagrangian, Hamiltonian, and integrable systems, 2000

77 **Fumio Hiai and Dénes Petz,** The semicircle law, free random variables and entropy, 2000

76 **Frederick P. Gardiner and Nikola Lakic,** Quasiconformal Teichmüller theory, 2000

75 **Greg Hjorth,** Classification and orbit equivalence relations, 2000

74 **Daniel W. Stroock,** An introduction to the analysis of paths on a Riemannian manifold, 2000

73 **John Locker,** Spectral theory of non-self-adjoint two-point differential operators, 2000

72 **Gerald Teschl,** Jacobi operators and completely integrable nonlinear lattices, 1999

71 **Lajos Pukánszky,** Characters of connected Lie groups, 1999

70 **Carmen Chicone and Yuri Latushkin,** Evolution semigroups in dynamical systems and differential equations, 1999

69 **C. T. C. Wall (A. A. Ranicki, Editor),** Surgery on compact manifolds, second edition, 1999

68 **David A. Cox and Sheldon Katz,** Mirror symmetry and algebraic geometry, 1999

67 **A. Borel and N. Wallach,** Continuous cohomology, discrete subgroups, and representations of reductive groups, second edition, 2000

66 **Yu. Ilyashenko and Weigu Li,** Nonlocal bifurcations, 1999

65 **Carl Faith,** Rings and things and a fine array of twentieth century associative algebra, 1999

64 **Rene A. Carmona and Boris Rozovskii, Editors,** Stochastic partial differential equations: Six perspectives, 1999

63 **Mark Hovey,** Model categories, 1999

62 **Vladimir I. Bogachev,** Gaussian measures, 1998

61 **W. Norrie Everitt and Lawrence Markus,** Boundary value problems and symplectic algebra for ordinary differential and quasi-differential operators, 1999

60 **Iain Raeburn and Dana P. Williams,** Morita equivalence and continuous-trace C^*-algebras, 1998

59 **Paul Howard and Jean E. Rubin,** Consequences of the axiom of choice, 1998

58 **Pavel I. Etingof, Igor B. Frenkel, and Alexander A. Kirillov, Jr.,** Lectures on representation theory and Knizhnik-Zamolodchikov equations, 1998

57 **Marc Levine,** Mixed motives, 1998

56 **Leonid I. Korogodski and Yan S. Soibelman,** Algebras of functions on quantum groups: Part I, 1998

55 **J. Scott Carter and Masahico Saito,** Knotted surfaces and their diagrams, 1998

(*Continued in the back of this publication*)

Number Theoretic Density and Logical Limit Laws

Mathematical
Surveys
and
Monographs

Volume 86

Number Theoretic Density and Logical Limit Laws

Stanley N. Burris

American Mathematical Society

2000 *Mathematics Subject Classification.* Primary 03C13, 05A15, 05A16, 05C30, 11M41, 11M45, 11N45, 11N80, 11P82, 11U09.

This project has been supported by the Natural Sciences and Engineering Research Council of Canada.

ABSTRACT. This book shows how a study of generating series (power series in the additive case and Dirichlet series in the multiplicative case), combined with structure theorems for the finite models of a sentence, lead to general and powerful results on limit laws, including 0–1 laws. Furthermore, a wealth of evidence is collected to support the thesis that local results in the additive case lift to global results in the multiplicative case.

Library of Congress Cataloging-in-Publication Data

Burris, Stanley.
Number theoretic density and logical limit laws / Stanley N. Burris.
p. cm. — (Mathematical surveys and monographs ; v. 86)
Includes bibliographical references and index.
ISBN 0-8218-2666-2 (alk. paper)
1. Generating functions. 2. Limit theorems (Probability theory) I. Title. II. Mathematical surveys and monographs ; no. 86.
QA164.8 .B87 2000
511′.62—dc21 00-048512

---

10 9 8 7 6 5 4 3 2 1 06 05 04 03 02 01

Contents

Preface

In the late 1980s Kevin Compton published papers showing how to apply an analysis of the popular partition identities to obtain monadic second-order limit laws, including 0–1 laws, for numerous classes of graphs, posets, etc. I was fascinated by this work. When he gave a colloquium talk in the early 1990s in Waterloo on his recent work on a limit law for Abelian groups, the temptation to learn more about this subject was irresistible. This engaged a favorite area of my own research, algebraic structures.

Compton's papers can be somewhat opaque for specialists in combinatorics and number theory, as well as for specialists in logic, because of the intimate way he has woven these subjects together. After seeing the books [**29**], [**30**] of John Knopfmacher on abstract analytic number theory, it seemed worthwhile to separate Compton's treatment into two parts, one on density in number systems, and the other on the application of number theoretic density results to obtain logical limit laws. Each of these parts is of interest in its own right.

The reader will find a leisurely and detailed exposition of Compton's investigations, and closely related work of others, including recent contributions of Jason Bell, Edward Bender, Peter Cameron, Paweł Idziak, Arnold Knopfmacher, John Knopfmacher, Andrew Odlyzko, Bruce Richmond, András Sárközy, Cameron Stewart, Richard Warlimont, Alan Woods, and the author. The presentation is from the perspective of abstract number systems, in the spirit of John Knopfmacher's work in abstract analytic number theory.

This book has been used as an undergraduate special topics reading text at the University of Waterloo. Part 1, on Additive Number Systems, is completely accessible to an advanced undergraduate student. All chapters preceding Chapter 6 are devoted to number theoretic density, requiring only the usual undergraduate background in analysis, especially in power series, and an exposure to abstract mathematics. The well known ratio test plays a central role. The section on asymptotics, at the end of Chapter 5, uses basic complex analysis, including the Cauchy integral formula. Chapter 6 covers the logical aspects for Part 1. This chapter is self-contained so that one can work through it without prior exposure to logic. It features one of the most delightful tools of logic, the Ehrenfeucht-Fraissé games. (Having had a first course in logic, so that one is comfortable with first-order languages and structures, will no doubt make the chapter more rapid reading.)

Part 2, on Multiplicative Number Systems, offers the challenge and reward of becoming reasonably comfortable with Dirichlet series. The parallels with Part 1 show Dirichlet series as a natural companion of power series. Having worked through Part 1, one will be able to predict many of the results to be proved—the local density results of Part 1 seem, as if by magic, to reappear as global density results in Part 2. There is surely some deep connection between power series and Dirichlet series that we have not yet understood. The last chapter introduces the reader to the Feferman-Vaught Theorem, a favorite tool to analyze direct products, and Skolem's analysis of first-order sentences about Boolean algebras.

The reader will find all the material needed to thoroughly understand the method of Compton for proving logical limit laws. Above all, I think one will be delighted to see so many interesting tools from elementary mathematics pull together to help answer the question "What is the probability that a randomly chosen structure has a given property?"

Thanks go to Paweł Idziak for his contributions to the study of limit laws when I was first starting to work in the area, to Andrew Odlyzko for helping me understand what was going on with the Dirichlet series, to András Sárközy for helping develop the general multiplicative theory of limit laws during a visit to Waterloo, to Cameron Stewart for discussions of the ratio test, to Dejan Delic for help with proofreading an early draft of the book, to Bruce Richmond for helping me to locate and understand several relevant results from asymptotics, to John Knopfmacher for challenging me to take a harder look at what was going on with logical limit laws, to Richard Warlimont for a remarkable amount of beautifully handwritten correspondence regarding ways to improve the presentation, to Jason Bell for keeping me informed of his recent research in this area, and to Karen Yeats, a second year mathematics undergraduate at U. Waterloo who eagerly read the entire manuscript during the summer of 2000, giving me detailed feedback on how the text comes across to an undergraduate. Kevin Compton is, of course, the ultimate inspiration for this work, through his publications and his elegant lectures over the years.

Edward Dunne, Christine Thivierge and Elaine Becker of the AMS Book Program did everything possible to make the transition from manuscript to book a pleasant and trouble-free experience; and Barbara Beeton of the AMS Technical Support group was (as always!) able to solve all of my Tex related problems with the document.

Finally, I want to thank the Natural Sciences and Engineering Research Council of Canada for their long standing support of my investigations in universal algebra, logic, and computation, the support that has made it possible to write this book.

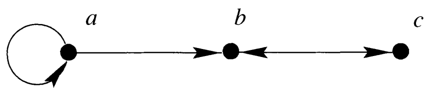
For errata and updates see www.thoralf.uwaterloo.ca.

Stanley Burris
Waterloo, 2000

Overview

When studying a subject with a bewildering variety of specimens one hopes to discover simple structural patterns. This happened in the study of finite relational structures with the discovery of 0–1 laws in the mid 1970s.

A finite relational structure $\mathbf{S} = (S, R_1, \ldots, R_n)$ is a set S with a list of relations $R_1, \ldots, R_n$ on S. The most popular relations R are binary, that is, $R \subseteq S \times S$. But $R \subseteq S^k$, for any $k \geq 1$, is also permitted. For the purpose of this overview it suffices to consider relational structures $\mathbf{S} = (S, R)$ with a single binary relation. One can think of R as a directed edge relation on the set of vertices S and draw a picture, for example:



In this picture one sees that aRa, aRb, bRc, cRb. Such structures are *directed graphs.*

As specializations of directed graphs one has several well known classes. If the relation R is irreflexive (not aRa, for any $a \in S$) and symmetric ($aRb \Rightarrow bRa$) one has a *graph.* If it is reflexive, symmetric and transitive, then one has an *equivalence relation.* And if it is reflexive, antisymmetric and transitive, then one has a *poset.* Finite (directed) graphs exhibit incredible diversity and have provided a rich source of examples and problems, from Euler's analysis of the Seven Bridges of Königsberg problem to the modern study of complexity in computer science. Indeed, the interest in finite structures has blossomed in tandem with the growth of theoretical computer science.

One of the fundamental questions is 'What can be said about a randomly chosen structure?' Consider a property $\mathcal{P}$ and a class $\mathcal{K}$ of finite structures. One can ask 'What is the probability that a finite structure chosen randomly from $\mathcal{K}$ satisfies $\mathcal{P}$?' The simplest and most natural definition of this probability is to let p_n be the proportion of structures in $\mathcal{K}$ of size n that have the property $\mathcal{P}$, and then to let the probability that $\mathcal{P}$ holds for a randomly chosen structure from $\mathcal{K}$ be the limit of p_n as n goes to infinity (whenever this limit exists). This probability is called *the asymptotic density* of the collection of structures in $\mathcal{K}$ that satisfy φ.

There is a problem with this definition if there are infinitely many n such that $\mathcal{K}$ has no structures of size n, for then p_n is infinitely often undefined.

This is handled by considering only those p_n that are defined. With this understanding it turns out that the general theory is a minor variation of the theory where one assumes that the p_n are eventually well-defined. So, for the purpose of this overview, it will be assumed that the classes $\mathcal{K}$ have structures of size n for all n greater than some N.

Glebskij, Logan, Liogonkij and Talanov (1969), and independently Fagin (1976), considered properties defined by first-order sentences. For example, the sentence $\forall x \exists y (xRy)$ says, in the case of finite graphs, that there are no isolated vertices; and the sentence $\forall x \forall y \big[(xRy) \rightarrow \exists z (xRz \wedge zRy)\big]$ says one can interpolate a vertex between adjacent vertices. For $\mathcal{K}$ the class of finite directed graphs, or the class of finite graphs, they showed, for any first-order sentence φ, that the probability of φ holding is either 0 or 1. Hence [directed] graphs have a *first-order 0–1 law.* If the probability of a property holding in a class $\mathcal{K}$ is 1 then the property is *almost certainly true* in $\mathcal{K}$.

The method used to prove these results does not generalize readily. First the result is proved for *labeled structures*, using the set of vertices $\{1, \dots, n\}$ for n-element structures. The following two directed graphs are identified when counting up to isomorphism, but are considered distinct when counting labeled structures:

When counting up to isomorphism one is said to be counting *unlabeled* structures.

To prove a 0–1 law for labeled structures, the original method is to understand the structures in $\mathcal{K}$ well enough to propose a basis Φ for the almost certainly true sentences. Then one proves that each member of Φ is indeed almost certainly true. Finally, to take a labeled 0–1 law back to the unlabeled case one needs to know that the property of being *rigid*[1] is almost certainly true in $\mathcal{K}$. Finding a basis Φ, and proving that rigidity is almost certainly true, are both serious obstacles to extending the applications of this method. Compton was able to carry this out for *posets*, but there are precious few other examples.

In the 1980s an alternate approach to proving 0–1 laws was developed by Compton. By considering *adequate* classes $\mathcal{K}$ of finite structures, that is, classes closed under disjoint union and components, he was able to show that the single condition $a(n-1)/a(n) \rightarrow 1$ is sufficient for a first-order 0–1 law. Here $a(n)$ counts the total number of structures of size n in $\mathcal{K}$ (counting up to isomorphism). This method does not apply to the original examples of graphs and directed graphs as these classes grow so rapidly that one has $a(n-1)/a(n) \rightarrow 0$. However, it is a beautiful technique that does apply to a truly wide range of slowly growing classes such as equivalence relations, permutations, linear forests, etc.

[1] A structure is rigid if the only automorphism is the identity map.

Furthermore, Compton's technique yields a 0–1 law for *monadic second-order logic.* This logic extends the well known first-order logic by allowing quantification over subsets—it is able to express far more properties than first-order logic. For example, in first-order logic one cannot say that a graph is *connected*, but in monadic second-order logic this is expressible by saying that 'if the domain is partitioned into two sets then it is always possible to find two vertices, one from each partition set, with an edge between them':

$$\forall U \forall V \Big[\Big(\exists x(Ux) \wedge \exists y(Vy) \wedge \forall x \big(Ux \leftrightarrow \neg Vx\big) \Big) \\ \rightarrow \exists x \exists y \big(Ux \wedge Vy \wedge xRy\big) \Big].$$

A considerable portion of the work in Compton's treatment is devoted to studying *Dirichlet density.* (This density is the limiting value of a quotient of two generating series.) He also uses the *fundamental identity*

$$\sum_{n \geq 0} a(n) x^n = \prod_{n \geq 1} \left(1 - x^n\right)^{-p(n)}$$

that relates $a(n)$ to $p(n)$, where $p(n)$ is the number of *components* of size n (counting up to isomorphism). On the surface, Dirichlet density appears to have little to do with the probability that a randomly chosen structure has a property, but under a natural hypothesis one can show that if the probability exists then so does the Dirichlet density, and they are equal. Thus Dirichlet density extends probability (defined as asymptotic density), and the main goal is to find theorems for a converse result that guarantees that if the Dirichlet density exists then it is the asymptotic density. Such converses are called Tauberian Theorems. For the development of the very large part of Compton's theory just described in this paragraph, one does not need the details of the relations involved in the individual structures, but rather just the additive number system associated with $\mathcal{K}$.

Consider the example of the class $\mathcal{K}$ of finite graphs. If one defines the sum of two finite graphs to be their disjoint union then, by identifying isomorphic graphs, one ends up with an additive number system $\mathcal{A}$. The graphs are the 'numbers' in this abstract system. The graph on the empty set gives the zero element of $\mathcal{A}$. The only information that is needed for the asymptotic work with additive number systems is the addition table for the elements and the 'size' (or *norm*) of a graph (defined as the number of vertices). The indecomposable elements, the nonzero elements that cannot be written as the sum of two nonzero elements, are precisely the connected graphs. Clearly every nonzero element of $\mathcal{A}$ can be expressed uniquely as a sum of indecomposable elements since every finite graph with a nonempty set of vertices is uniquely expressible as a disjoint union of connected graphs. The essential features of an additive number system are just the addition operation and the size function. The adjective 'additive' is derived from the fact that the size function $\|\mathsf{a}\|$ is additive, that is, $\|\mathsf{a} + \mathsf{b}\| = \|\mathsf{a}\| + \|\mathsf{b}\|$.

To simplify the presentation, additive number systems are interpolated between adequate classes of structures and the fundamental identities:

Adequate Classes of Structures	—	Additive Number Systems	—	Fundamental Identities

This interpolation allows one to separate the logical aspects from the number-theoretic aspects. Every adequate class gives rise to a (unique) additive number system, and every additive number system can be derived from (many) adequate classes; and each additive number system gives rise to a (unique) fundamental identity that essentially defines the number system.

Given an additive number system $\mathcal{A}$, subsets called *partition sets* play a crucial role in this work. A partition set has the form $\gamma_1 \mathsf{P}_1 + \cdots + \gamma_k \mathsf{P}_k$, where $\mathsf{P}_1, \ldots, \mathsf{P}_k$ is a partition of the set P of indecomposable elements of $\mathcal{A}$, and where each of the γ_i is in one of the three forms $(\geq m_i)$, m_i, $(\leq m_i)$. The elements of such a partition set are the members of $\mathcal{A}$ which can be written as a sum of γ_1 members of P_1 plus ... plus γ_k members of P_k. (Repeats are allowed when counting elements.) Thus $1\mathsf{P}$ is just P, and $(\geq 0)\mathsf{P}$ is the entire set of 'numbers' in $\mathcal{A}$. And if $\mathsf{P}_1, \mathsf{P}_2, \mathsf{P}_3$ is a partition of the set P of indecomposable elements of $\mathcal{A}$ then the partition set $(\leq 5)\mathsf{P}_1 + (\geq 4)\mathsf{P}_2 + 2\mathsf{P}_3$ is the set of elements which can be expressed as the sum of at most 5 elements from P_1 plus the sum of at least 4 elements from P_2 plus the sum of exactly 2 elements from P_3.

Now, if $\mathcal{K}$ is an adequate class of structures and $\mathcal{A}$ the corresponding additive number system, then the members of $\mathcal{K}$ that satisfy a given monadic second-order sentence can be described, when viewed in $\mathcal{A}$, as a disjoint union of finitely many partition sets. Thus, if one can show that every partition set of $\mathcal{A}$ has density 0 or 1 then $\mathcal{K}$ has a monadic second-order 0–1 law.

The method of Compton, to establish a 0–1 law for $\mathcal{K}$, is indeed to show that all partition sets of $\mathcal{A}$ have asymptotic density 0 or 1. The necessary and sufficient condition for this to hold is that $a(n-1)/a(n) \to 1$. To obtain interesting conditions that guarantee $a(n-1)/a(n) \to 1$, one turns to the fundamental identity. This ties in with the established work in asymptotic additive number theory, a subject often described by referring to the famous results of Hardy and Ramanujan on the number of partitions of an integer. The applications presented in Chapter 4 are based on an analysis of Bateman and Erdös of additive number systems which have at most one indecomposable of each size.

In a subsequent paper Compton turned to more general limit laws for logic, where one only requires that each sentence φ have a probability of holding (the probability need not be 0 or 1). Again, additive number systems are used to shift to the study of conditions that guarantee that every partition set has an asymptotic density. A rather delicate analysis is needed to establish the main result of Compton. One of the striking corollaries is that if $a(n)$ is asymptotic to $C\beta^n$ then $\mathcal{K}$ has a limit law. Further applications come from the asymptotics of Knopfmacher, Knopfmacher, and Warlimont. By applying the Cauchy integral formula

$$a(n) \;=\; \frac{1}{2\pi i}\int_C \frac{\mathbf{A}(z)}{z^{n+1}}dz$$

to $\mathbf{A}(z) = \sum a(k)z^k$, they are able to find asymptotics for $a(n)$ when $p(n) = C\beta^n + \mathrm{o}(\eta^n)$, where $0 < \eta < \beta$. This gives numerous applications of Compton's theorem, for example, to two-colored linear forests.

In the late 1980s Compton was considering the application of his ideas to the study of randomly chosen members of classes $\mathcal{K}$ of finite algebraic structures—the class of finite Abelian groups is an excellent example. Here the key operation is *direct product*, not disjoint union. The analog of a component (of a relational structure) is, in this setting, a (*directly*) *indecomposable* structure—that is, a structure which has at least two elements and is not isomorphic to a direct product of smaller structures. The indecomposable finite Abelian groups are precisely the $\mathbf{Z}_{p^n}$, where $\mathbf{Z}_{p^n}$ is the group of integers modulo p^n, for p a prime number. As is well known, every nontrivial finite Abelian group can be uniquely expressed as a product of indecomposable Abelian groups. This is the *unique factorization property* for finite Abelian groups. However, unique factorization does not hold for many important classes of algebraic structures, for example, semigroups.

In the context of finite algebraic structures let us say that $\mathcal{K}$ is *adequate* if it is closed under direct product, direct factors, and it has the unique factorization property. Let $a(n)$ count the number of structures in $\mathcal{K}$ of size n, and let $p(n)$ count the number of indecomposables in $\mathcal{K}$ of size n. (All counting is done up to isomorphism.) From the fact that $\mathcal{K}$ is adequate one has the fundamental identity

$$\sum_{n\geq 1} a(n)n^{-s} \;=\; \prod_{n\geq 2}\left(1-n^{-s}\right)^{-p(n)}.$$

What is the probability that a randomly chosen structure from $\mathcal{K}$ satisfies a given property $\mathcal{P}$? If one defines p_n as before, namely the proportion of structures in $\mathcal{K}$ of size n that have the property $\mathcal{P}$, then very little can be said. However, if one uses the *global* count, that is, let P_n be the proportion of structures in $\mathcal{K}$ of size *at most* n that have the property $\mathcal{P}$, and if one lets the probability be the limit of P_n as n goes to infinity (provided this exists), then the results for relational structures have remarkable parallels in the algebraic setting.

In Part 2 multiplicative number systems are interpolated between adequate classes and fundamental identities, giving an exact analog of the use of additive number systems in Part 1. This is used, as in Part 1, to give a separation of the number theoretic and logical aspects.

Adequate Classes of Structures	——	Multiplicative Number Systems	——	Fundamental Identities

In the case of Abelian groups one obtains the associated number system by letting the 'numbers' be the finite groups $\mathbf{G}$, and letting multiplication be direct product. These number systems are said to be multiplicative because the size function is multiplicative, that is, $\|\mathsf{a} \cdot \mathsf{b}\| = \|\mathsf{a}\| \cdot \|\mathsf{b}\|$.

The multiplicative number system analog of Compton's theorem (on general limit laws for relational structures) has, as a corollary, the fact that if $A(x)$ is asymptotic to cx^α, where $A(x) = \sum_{n \leq x} a(n)$, then $\mathcal{K}$ has a first-order limit law. This applies to the example of Abelian groups, and one can show, for example, that the probability of a finite Abelian group having an element of order 2 is $1 - \prod_{n \geq 1} \left(1 - 2^{-n}\right) \approx 0.71$.

Many of the interesting examples, like Abelian groups, have limit laws, but not 0–1 laws. Further examples are again found by turning to complex analysis, using Perron's integral formula, to generalize Oppenheim's asymptotics for the number of ways to factor the numbers less than or equal to n. From this analysis one concludes that if $p(n) = Cn^\alpha + O(n^\beta)$ with $C > 0$, $\alpha \geq 0$, and $\beta < \alpha$, then $\mathcal{K}$ has a first-order limit law. Thus the class of finite lattices that decompose into a product of chains has a first-order limit law.

Chapter 6, the last chapter of Part 1, covers the logic results for adequate classes (with respect to disjoint union) of purely relational structures; and Chapter 12, the last chapter of Part 2, covers the logic results for adequate classes (with respect to direct product) of structures.

These two chapters can be briefly summarized as follows. Given an adequate class $\mathcal{K}$ and a sentence φ, let $\mathcal{K}_\varphi$ be the class of members of $\mathcal{K}$ that satisfy φ. The goal is to prove that $\mathcal{K}_\varphi$ is a disjoint union of finitely many partition classes. Then the number theoretic results on asymptotic density of partition sets can be used.

Three tools are needed to prove these logic results: the Ehrenfeucht-Fraissé games in the additive case; and the Feferman-Vaught Theorem, with Skolem's analysis of sentences about finite Boolean algebras, in the multiplicative case. These tools are fully developed in Chapters 6 and 12.

Notation Guide

The most basic convention used in this text is that **power series** and **Dirichlet series** are designated by upper case boldface letters, with the corresponding lower case italic letters used for the coefficients, and the corresponding upper case italic leters used for the partial sums of the coefficients.

Power series (Part 1)	Dirichlet series (Part 2)	Sums of coefficients
$\mathbf{A}(x) = \sum_n a(n)x^n$	$\mathbf{A}(x) = \sum_n a(n)n^{-x}$	$A(x) = \sum_{n\le x} a(n)$
$\vdots$	$\vdots$	$\vdots$
$\mathbf{S}(x) = \sum_n s(n)x^n$	$\mathbf{S}(x) = \sum_n s(n)n^{-x}$	$S(x) = \sum_{n\le x} s(n)$
$\vdots$	$\vdots$	$\vdots$

For work with abstract number systems the following notational conventions are used:

$\mathcal{A}, \mathcal{B}, \ldots$	number systems
$\mathcal{A} = (\mathsf{A}, \mathsf{P}, +, 0, \Vert\ \Vert)$	an additive number system
$\mathcal{A} = (\mathsf{A}, \mathsf{P}, \cdot, 1, \Vert\ \Vert)$	a multiplicative number system
$\mathsf{a}, \mathsf{b}, \ldots, \mathsf{p}, \ldots$	elements of the number system
A	set of elements of the number system
P	set of indecomposable elements
$\mathsf{B}, \mathsf{C}, \ldots$	subsets of A
$b(n)$	local count function for B
$B(x) = \sum_{n\le x} b(n)$	global count function for B
$\mathbf{B}(x) = \sum_n b(n)x^n$	generating series for B (additive case)
$\mathbf{B}(x) = \sum_n b(n)n^{-x}$	generating series for B (multiplicative case)

(continued)

ρ	radius of convergence of $\mathbf{A}(x)$ (Part 1)
α^-	abscissa of convergence of $\mathbf{A}(x)$ (Part 2)
α	abscissa of absolute convergence of $\mathbf{A}(x)$ (Part 2)
$\delta(\mathsf{B}) = \lim\limits_{\substack{n\to 0\\ a(n)\neq 0}} b(n)/a(n)$	local asymptotic density of B
$\Delta(\mathsf{B}) = \lim\limits_{n\to\infty} B(n)/A(n)$	global asymptotic density of B
$\partial(\mathsf{B}) = \lim\limits_{x\to\rho} \mathbf{B}(x)/\mathbf{A}(x)$	Dirichlet density of B (additive case)
$\partial(\mathsf{B}) = \lim\limits_{x\to\alpha} \mathbf{B}(x)/\mathbf{A}(x)$	Dirichlet density of B (multiplicative case)
$f(x) \in \mathcal{PS}_c$	$f(x)$ has a power series expansion that converges absolutely at c
$f(x) \in \mathcal{DS}_c$	$f(x)$ has a Dirichlet series expansion that converges absolutely at c
$\mathsf{supp}\, f(n) = \{n : f(n) > 0\}$	support of $f(n)$
$\mathbf{S}\vert_m(x) = \sum\limits_{k\leq m} s(k)x^k$	mth truncation of the power series $\mathbf{S}(x)$
$\mathbf{S}\vert_m(x) = \sum\limits_{k\leq m} s(k)k^{-x}$	mth truncation of the Dirichlet series $\mathbf{S}(x)$
$0\mathsf{B} = \{0\}$	(for additive systems: Part 1)
$\mathsf{B}^0 = \{1\}$	(for multiplicative systems: Part 2)
$m\mathsf{B} = \{\mathsf{a}_1 + \cdots + \mathsf{a}_m : \mathsf{a}_i \in \mathsf{B}\}$	(for additive systems: Part 1)
$\mathsf{B}^m = \{\mathsf{a}_1 \cdots \mathsf{a}_m : \mathsf{a}_i \in \mathsf{B}\}$	(for multiplicative systems: Part 2)
$(\leq m)\mathsf{B} = 0\mathsf{B} \cup \cdots \cup m\mathsf{B}$	(for additive systems: Part 1)
$\mathsf{B}^{\leq m} = \mathsf{B}^0 \cup \cdots \cup \mathsf{B}^m$	(for multiplicative systems: Part 2)
$(\geq m)\mathsf{B} = \bigcup_{n\geq m} n\mathsf{B}$	(for additive systems: Part 1)
$\mathsf{B}^{\geq m} = \bigcup_{n\geq m} \mathsf{B}^n$	(for multiplicative systems: Part 2)

$\boxed{\text{for } \mathsf{Q} \subseteq \mathsf{P}}$

$\mathsf{A}_\mathsf{Q} = \omega\mathsf{Q} = \bigcup_{n\geq 0} n\mathsf{Q}$	(for additive systems: Part 1)
$\mathsf{A}_\mathsf{Q} = \mathsf{Q}^\omega = \bigcup_{n\geq 0} \mathsf{Q}^n$	(for multiplicative systems: Part 2)
$\partial_\mathsf{Q}(\gamma\mathsf{Q}) = \lim\limits_{x\to\rho} \dfrac{[\gamma\mathbf{Q}](x)}{\mathbf{A}_\mathsf{Q}(x)}$	relativized Dirichlet density, additive case
$\partial_\mathsf{Q}(\mathsf{Q}^\gamma) = \lim\limits_{x\to\alpha} \dfrac{[\mathbf{Q}^\gamma](x)}{\mathbf{A}_\mathsf{Q}(x)}$	relativized Dirichlet density, multiplicative case

APPENDIX A

Formal Power Series

A detailed and elementary treatment of the foundations of formal power series is given. To motivate the development, first the basic theory of functions is reviewed.

Real functions

A function is defined to be what is commonly called the *graph of a function*, that is, a function is to be considered as a collection of ordered pairs of real numbers.[1] This set theoretic approach may seem unnecessarily cumbersome at first, but it leads to the possibility of presenting simply, precisely formulated results. A number of key concepts are collected together in the following.

Definition A.1.

- A *real function* f is a subset of $\mathbb{R} \times \mathbb{R}$, the set of ordered pairs of real numbers, with the *single valued* property:
$$\left.\begin{array}{l}(a,b) \in f\\(a,c) \in f\end{array}\right\} \Rightarrow b = c.$$
- $\mathcal{F}$ is the set of all real functions.
- For $f \in \mathcal{F}$ the *domain* of f, $\mathrm{Dom}(f)$, and the *range* of f, $\mathrm{Rg}(f)$, are given by
$$\begin{aligned}\mathrm{Dom}(f) &= \big\{a \in \mathbb{R} : (\exists b \in \mathbb{R})\big((a,b) \in f\big)\big\}\\ \mathrm{Rg}(f) &= \big\{b \in \mathbb{R} : (\exists a \in \mathbb{R})\big((a,b) \in f\big)\big\}.\end{aligned}$$
$\mathrm{Dom}(f)$ is also called *the domain of definition* of f.
- For $a \in \mathrm{Dom}(f)$, let $f(a)$ denote the unique $b \in \mathbb{R}$ such that $(a,b) \in f$. $f(a)$ is the *value of f at a*.
- Given $c \in \mathbb{R}$, $\mathbf{c}$ is the function on $\mathbb{R}$ whose value is always c.
- x denotes the identity function on $\mathbb{R}$.
- $\emptyset$ denotes the *empty* function, the function with no ordered pairs in it.

For $f \in \mathcal{F}$, one has the trivial observation
$$f = \{(a, f(a)) : a \in \mathrm{Dom}(f)\}.$$

[1] The focus here is on real valued functions with their domain in the reals. This discussion can be applied to other situations, for example, complex valued functions defined on a subset of the complex numbers.

In calculus courses a function is described by an expression, like $\sqrt{x}$. The convention is to give the function the largest domain for which the expression is defined, so $\mathrm{Dom}(\sqrt{x}) = [0, \infty)$. In the more precise notation

$$\surd \;=\; \big\{(a,b) : a, b \geq 0,\ a = b^2\big\}.$$

Since functions are sets of order pairs, the usual notation $f \subseteq g$ is used to signify that every pair in f is also in g.

Definition A.2. If $f \subseteq g$ then f is a *restriction* of g and g is an *extension* of f.

For example the function defined by the power series $\sum_n x^n$ is a restriction of the function (defined by) $(1-x)^{-1}$.

Definition A.3. Suppose $f \in \mathfrak{F}$ and $S \subseteq \mathbb{R}$. Then $f|_S$, the *restriction of f to S*, is given by

$$f|_S \;=\; f \cap (S \times \mathbb{R}),$$

that is, $f|_S \;=\; \{(a,b) \in f : a \in S\}$.

Clearly $f|_S \subseteq f$, so the restriction of f to S is indeed a restriction of f. Furthermore

$$\mathrm{Dom}(f|_S) \;=\; \mathrm{Dom}(f) \cap S,$$

that is, a is in the domain of $f|_S$ iff it is in the domain of f as well as in S.

Thus the function defined by the power series $\sum_n x^n$ is $(1-x)^{-1}\big|_{(-1,1)}$, and not the function $(1-x)^{-1}$ with the larger domain of definition $\mathbb{R} \setminus \{1\}$.

Operations on real functions. Next the common operations used with real functions are considered.

negative: For $f \in \mathfrak{F}$

$$-f \;=\; \big\{\big(a, -f(a)\big) : a \in \mathrm{Dom}(f)\big\}$$

so $(-f)(a) \;=\; -f(a)$, for $a \in \mathrm{Dom}(f)$.

scalar multiplication: For $c \in \mathbb{R}$ and $f \in \mathfrak{F}$

$$cf \;=\; \big\{\big(a, c \cdot f(a)\big) : a \in \mathrm{Dom}(f)\big\}$$

so $(cf)(a) \;=\; c \cdot f(a)$, for $a \in \mathrm{Dom}(f)$.

addition: For $f, g \in \mathfrak{F}$

$$f + g \;=\; \big\{\big(a, f(a) + g(a)\big) : a \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g)\big\}$$

so $(f+g)(a) \;=\; f(a) + g(a)$, for $a \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g)$.

The empty function comes in handy here—if f and g have disjoint domains, then $f + g$ is just the empty function. A similar remark applies to the next two operations.

subtraction: Reduce this to addition by defining

$$f - g \;=\; f + (-g)$$

Here is another example of how this notation can be more precise than the usual, namely

$$\surd - \surd \;=\; \mathbf{0}\big|_{[0,\infty)}$$

and not $\surd - \surd = \mathbf{0}$.

multiplication: For $f, g \in \mathcal{F}$

$$f \cdot g \;=\; \big\{\big(a, f(a) \cdot g(a)\big) : a \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g)\big\}$$

so $(f \cdot g)(a) \;=\; f(a) \cdot g(a)$, for $a \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g)$.

composition: For $f, g \in \mathcal{F}$

$$f \circ g \;=\; \big\{\big(a, f(g(a))\big) : a \in \mathrm{Dom}(g),\ g(a) \in \mathrm{Dom}(f)\big\}$$

so $(f \circ g)(a) \;=\; f\big(g(a)\big)$, for $a \in \mathrm{Dom}(g) \cap g^{-1}\big(\mathrm{Dom}(f)\big)$.

Composition is often thought of as *substitution*, for example, if $f(x) = g(x)$ then $f(\sin x) = g(\sin x)$ by substitution. But the latter can also be written as $f \circ \sin = g \circ \sin$.

differentiation: Given $f \in \mathcal{F}$, define the derivative f' to be the function whose domain is the set of all $a \in \mathrm{Dom}(f)$ for which one can find a neighborhood[2] N_a of a with $N_a \subseteq \mathrm{Dom}(f)$, that is, f is defined on N_a, and the following limit exists:

$$\lim_{\substack{u \to 0 \\ u \in N_a}} \frac{f(a+u) - f(a)}{u}.$$

If a is in the domain of f', then $f'(a)$ is this limit.

Remark A.4. The operations $+, \cdot, \circ, '$ and scalar multiplication are *always defined*, no matter which functions they are applied to.

Identities of real functions. Now the basic properties of these operations on functions are listed. (It took many years during the 19th and early 20th centuries for mathematicians to explicitly list exactly which rules they were using.)

Proposition A.5. *For $a, b \in \mathbb{R}$ and $f, g, h \in \mathcal{F}$:*

scalar multiplication:

$$\begin{aligned} a\mathbf{0} &= \mathbf{0} & a\mathbf{1} &= \mathbf{a} \\ (-1)f &= -f & a(-f) &= (-a)f \;=\; -(af) \\ a(bf) &= (ab)f & (a+b)f &= (af) + (bf) \\ a(f+g) &= (af) + (ag) & a(f \cdot g) &= (af) \cdot g \;=\; f \cdot (ag) \end{aligned}$$

[2] A *neighborhood* of a is just an interval (c, d) with $c < a < d$.

addition, multiplication, and composition:

$$\begin{aligned} f + \mathbf{0} &= f & f \cdot \mathbf{1} &= f \\ f + g &= g + f & f \cdot g &= g \cdot f \\ (f+g)+h &= f+(g+h) & (f \cdot g) \cdot h &= f \cdot (g \cdot h) \\ f \cdot (g+h) &= (f \cdot g) + (f \cdot h) & (f+g) \circ h &= (f \circ h) + (g \circ h) \\ (f \cdot g) \circ h &= (f \circ h) \cdot (g \circ h) & (f \circ g) \circ h &= f \circ (g \circ h) \end{aligned}$$

differentiation:

$$\begin{aligned} (af)' &\supseteq af' & (f+g)' &\supseteq f' + g' \\ (f \cdot g)' &\supseteq (f' \cdot g) + (f \cdot g') & (f \circ g)' &\supseteq g' \cdot (f' \circ g) \\ \left(e^f\right)' &= f' \cdot e^f \end{aligned}$$

In this list one sees the famous product rule, and the chain rule of differential calculus. The last item is just a special case of the chain rule, but one that we want to emphasize.

Polynomials

If one regards $s(0) + \cdots + s(n)x^n$ merely as a formal expression, and not as a function, one is dealing with *formal polynomials*. There is an obvious correspondence between such polynomials and their sequences of coefficients $\big(s(0), \ldots, s(n)\big)$. So the sequence of coefficients will be taken as the definition of a formal polynomial. To make the description of addition and multiplication of polynomials simple to express, and to be able to consider polynomials as special cases of power series, a polynomial is viewed as having an infinite sequence of coefficients (that are eventually 0).

Throughout this appendix all polynomials are to be thought of as formal polynomials. A new notation will be introduced to discuss the functions they define. Consequently, from this point on the adjective 'formal' is dropped when talking about polynomials, but feel free to insert it everywhere you see the word polynomial.

Definition A.6. A *polynomial* is an infinite sequence $\big(s(n)\big)$ of real numbers such that the $s(n)$ are eventually 0. The $s(n)$ are the *coefficients* of the polynomial.

By tradition the polynomial $\mathbf{S} = \big(s(n)\big)$ is written in longer form as $\mathbf{S}(x) = s(0) + s(1)x + \cdots + s(k)x^k$, when $s(m) = 0$ for $m > k$, and also as $\mathbf{S}(x) = \sum_n s(n)x^n$. At this point x is merely a symbol, as are $+$ and $\sum$. It is only when real numbers are substituted for x that one is dealing with functions.

Functions defined by polynomials.

Definition A.7. Given a polynomial $\mathbf{S}$ let $f_{\mathbf{S}}$ be the real valued function defined on the reals by $f_{\mathbf{S}}(a) = \sum_n s(n)a^n$. $f_{\mathbf{S}}$ is called a *polynomial function*.

Knowing how functions defined by polynomials behave under the standard operations is crucial to the development.

$f^{(n)}$ means the *nth derivative* of f.

Proposition A.8. *Let* $\mathbf{S}, \mathbf{T}$ *be polynomials and let* $a, c \in \mathbb{R}$. *Then one has*

$$\begin{aligned}
(cf_{\mathbf{S}})(a) &= \sum_n cs(n)a^n \\
(f_{\mathbf{S}} + f_{\mathbf{T}})(a) &= \sum_n \big(s(n) + t(n)\big)a^n \\
(f_{\mathbf{S}} \cdot f_{\mathbf{T}})(a) &= \sum_n \Big(\sum_k s(k) \cdot t(n-k)\Big)a^n \\
(f_{\mathbf{S}} \circ f_{\mathbf{T}})(a) &= \sum_n \Big(\sum_k s(k) \sum_{\ell_1 + \cdots + \ell_k = n} t(\ell_1) \cdots t(\ell_k)\Big)a^n \\
f_{\mathbf{S}}{}'(a) &= \sum_{n \geq 1} ns(n)a^{n-1} \\
s(n) &= f_{\mathbf{S}}^{(n)}(0)/n! \qquad \textit{for } n \geq 0 \\
f_{\mathbf{S}} = f_{\mathbf{T}} &\Leftrightarrow \mathbf{S} = \mathbf{T}.
\end{aligned}$$

PROOF. Only basic facts about finite sums and products are needed. For $cf_{\mathbf{S}}$:

$$(cf_{\mathbf{S}})(a) = cf_{\mathbf{S}}(a) = c\sum_n s(n)a^n = \sum_n cs(n)a^n.$$

For $f_{\mathbf{S}} + f_{\mathbf{T}}$:

$$\begin{aligned}
(f_{\mathbf{S}} + f_{\mathbf{T}})(a) = f_{\mathbf{S}}(a) + f_{\mathbf{T}}(a) &= \sum_n s(n)a^n + \sum_n t(n)a^n \\
&= \sum_n \big(s(n) + t(n)\big)a^n.
\end{aligned}$$

For $f_{\mathbf{S}} \cdot f_{\mathbf{T}}$:

$$\begin{aligned}
(f_{\mathbf{S}} \cdot f_{\mathbf{T}})(a) &= f_{\mathbf{S}}(a) \cdot f_{\mathbf{T}}(a) = \sum_n s(n)a^n \cdot \sum_n t(n)a^n \\
&= \sum_{i,j} \big(s(i) \cdot t(j)\big)a^{i+j} = \sum_n \Big(\sum_{i+j=n} s(i) \cdot t(j)a^{i+j}\Big) \\
&= \sum_n \Big(\sum_{i+j=n} s(i) \cdot t(j)\Big)a^n.
\end{aligned}$$

For $f_{\mathbf{S}} \circ f_{\mathbf{T}}$:

$$
\begin{aligned}
(f_{\mathbf{S}} \circ f_{\mathbf{T}})(a) \;=\; f_{\mathbf{S}}\big(f_{\mathbf{T}}(a)\big) &= \sum_k s(k)\Big(\sum_m t(m)a^m\Big)^k \\
&= \sum_k s(k)\Big(\sum_{\ell_1,\dots,\ell_k \geq 0} t(\ell_1)\cdots t(\ell_k)a^{\ell_1+\cdots+\ell_k}\Big) \\
&= \sum_k s(k)\sum_n\Big(\sum_{\ell_1+\cdots+\ell_k=n} t(\ell_1)\cdots t(\ell_k)a^n\Big) \\
&= \sum_k\sum_n s(k)\Big(\sum_{\ell_1+\cdots+\ell_k=n} t(\ell_1)\cdots t(\ell_k)\Big)a^n \\
&= \sum_n\Big(\sum_k s(k)\sum_{\ell_1+\cdots+\ell_k=n} t(\ell_1)\cdots t(\ell_k)\Big)a^n.
\end{aligned}
$$

For $f_{\mathbf{S}}{}'$, apply the rules for differentiation of sums and products, along with $x' = 1$, to obtain the usual termwise differentiation formula. This leads to the formula for $s(n)$, and to the proof that $f_{\mathbf{S}} = f_{\mathbf{T}}$ iff $\mathbf{S} = \mathbf{T}$. □

Operations on polynomials. Now use Proposition A.8 to suggest appropriate definitions for the operations on polynomials.

Definition A.9. Let $\mathbf{S}, \mathbf{T}$ be polynomials and let $c \in \mathbb{R}$.

coefficients: $\big[x^n\big]\mathbf{S}$ is $s(n)$.

equality: $\mathbf{S} = \mathbf{T}$ holds precisely when the corresponding coefficients are equal, that is, for all $n \geq 0$, $\big[x^n\big]\mathbf{S} = \big[x^n\big]\mathbf{T}$.

constant polynomials:

$$\big[x^n\big]\mathbf{c} \;=\; \begin{cases} c & \text{if } n = 0 \\ 0 & \text{otherwise.} \end{cases}$$

scalar multiplication: $\big[x^n\big]\big(c\mathbf{S}\big) \;=\; cs(n)$.

addition: $\big[x^n\big]\big(\mathbf{S} + \mathbf{T}\big) \;=\; s(n) + t(n)$.

multiplication: Use the Cauchy product:

$$\big[x^n\big]\big(\mathbf{S} * \mathbf{T}\big) \;=\; \sum_{k \leq n} s(k) \cdot t(n-k).$$

composition:

$$\big[x^n\big]\big(\mathbf{S} \circ \mathbf{T}\big) \;=\; \sum_k s(k) \sum_{\ell_1+\cdots+\ell_k=n} t(\ell_1)\cdots t(\ell_k).$$

differentiation: $\big[x^n\big]\mathbf{S}' \;=\; (n+1)s(n+1)$.

Since the operations on polynomials have been defined to harmonize with Proposition A.8, immediately one has the following.

Identities of polynomial functions.

Proposition A.10. *Let* $\mathbf{S}, \mathbf{T}$ *be polynomials and let* $c \in \mathbb{R}$. *Then one has*

$$\begin{aligned} f_{c\mathbf{S}} &= c f_{\mathbf{S}} \\ f_{\mathbf{S}+\mathbf{T}} &= f_{\mathbf{S}} + f_{\mathbf{T}} \\ f_{\mathbf{S}*\mathbf{T}} &= f_{\mathbf{S}} \cdot f_{\mathbf{T}} \\ f_{\mathbf{S}\circ\mathbf{T}} &= f_{\mathbf{S}} \circ f_{\mathbf{T}} \\ f_{\mathbf{S}}' &= f_{\mathbf{S}'}. \end{aligned}$$

With this, and the fact that $f_{\mathbf{S}} = f_{\mathbf{T}} \Leftrightarrow \mathbf{S} = \mathbf{T}$, the entire Proposition A.5 can be lifted to the setting of polynomials.[3] The results on differentiation are improved to equality as the functions defined by polynomials have domain $\mathbb{R}$, and they are everywhere differentiable.

Proposition A.11. *For* $a, b \in \mathbb{R}$ *and polynomials* $\mathbf{R}, \mathbf{S}, \mathbf{T}$ *the following hold:*

scalar multiplication:

$$\begin{aligned} a\mathbf{0} &= \mathbf{0} & a\mathbf{1} &= \mathbf{a} \\ (-1)\mathbf{R} &= -\mathbf{R} & a(-\mathbf{R}) &= (-a)\mathbf{R} = -(a\mathbf{R}) \\ a(b\mathbf{R}) &= (ab)\mathbf{R} & (a+b)\mathbf{R} &= (a\mathbf{R}) + (b\mathbf{R}) \\ a(\mathbf{R}+\mathbf{S}) &= (a\mathbf{R}) + (a\mathbf{S}) & a(\mathbf{R}\cdot\mathbf{S}) &= (a\mathbf{R})\cdot\mathbf{S} = \mathbf{R}\cdot(a\mathbf{S}) \end{aligned}$$

addition, multiplication, and composition:

$$\begin{aligned} \mathbf{R}+\mathbf{0} &= \mathbf{R} & \mathbf{R}\cdot\mathbf{1} &= \mathbf{R} \\ \mathbf{R}+\mathbf{S} &= \mathbf{S}+\mathbf{R} & \mathbf{R}\cdot\mathbf{S} &= \mathbf{S}\cdot\mathbf{R} \\ (\mathbf{R}+\mathbf{S})+\mathbf{T} &= \mathbf{R}+(\mathbf{S}+\mathbf{T}) & (\mathbf{R}\cdot\mathbf{S})\cdot\mathbf{T} &= \mathbf{R}\cdot(\mathbf{S}\cdot\mathbf{T}) \\ \mathbf{R}\cdot(\mathbf{S}+\mathbf{T}) &= (\mathbf{R}\cdot\mathbf{S}) + (\mathbf{R}\cdot\mathbf{T}) & (\mathbf{R}+\mathbf{S})\circ\mathbf{T} &= (\mathbf{R}\circ\mathbf{T})+(\mathbf{S}\circ\mathbf{T}) \\ (\mathbf{R}\cdot\mathbf{S})\circ\mathbf{T} &= (\mathbf{R}\circ\mathbf{T})\cdot(\mathbf{S}\circ\mathbf{T}) & (\mathbf{R}\circ\mathbf{S})\circ\mathbf{T} &= \mathbf{R}\circ(\mathbf{S}\circ\mathbf{T}) \end{aligned}$$

differentiation:

$$\begin{aligned} (a\mathbf{R})' &= a\mathbf{R}' & (\mathbf{R}+\mathbf{S})' &= \mathbf{R}'+\mathbf{S}' \\ (\mathbf{R}\cdot\mathbf{S})' &= (\mathbf{R}'\cdot\mathbf{S}) + (\mathbf{R}\cdot\mathbf{S}') & (\mathbf{R}\circ\mathbf{S})' &= \mathbf{S}'\cdot(\mathbf{R}'\circ\mathbf{S}). \end{aligned}$$

[3] An algebraist would say that the algebraic structure $\{\text{Poly}, +, *, \circ, ', (c)_{c\in\mathbb{R}}, 0, 1\}$, where Poly is the set of polynomials over $\mathbb{R}$, is *isomorphic* to the corresponding algebraic structure of polynomial functions under the mapping $\mathbf{S} \mapsto f_{\mathbf{S}}$.

Power series

As with polynomials, if one regards $\sum a_n x^n$ merely as a formal expression, and not as a function, then one is dealing with *formal power series.* Many power series have radius of convergence 0, and one cannot learn much from a function that is only defined at 0. However, as formal expressions they contain information, namely one can read off the sequence $s(0), s(1), \ldots$, and this sequence may very well be of interest. Of course, if all that is wanted is the sequence $\big(s(n)\big)$ then the formal power series $\sum s(n)x^n$ would seem a rather clumsy way of capturing this information.

However, it turns out that identities between formal power series can capture valuable information about sequences, for example in Appendix C the formal identity

$$\mathbf{A}(x) \;=\; e^{\mathbf{H}(x)}$$

plays a pivotal role. By appropriately defining operations on formal power series, this identity holds even when the radius of convergence of $\mathbf{A}(x)$ is 0, and then one can differentiate and obtain another formal identity

$$\mathbf{A}'(x) \;=\; \mathbf{H}'(x) * \mathbf{A}(x).$$

Thus new information is obtained about the coefficients. This is the significance of formal power series for us.

A precise definition of a formal power series is given by focusing on the information contained in the sequence of coefficients. One might say that one writes $\big(s(n)\big)$ and thinks $\sum s(n)x^n$. Throughout this appendix all power series are to be thought of as formal power series, just as all polynomials are formal polynomials. Notation will be introduced to discuss the functions that power series define. Consequently, from this point on the adjective 'formal' is dropped, but feel free to insert it anywhere you see the words power series.

Definition A.12. A *power series* is a sequence $\big(s(n)\big)$ of real numbers. The $s(n)$ are the *coefficients* of the power series.

By tradition the power series $\mathbf{S} = \big(s(n)\big)$ is written as $\mathbf{S}(x) = \sum_n s(n)x^n$. Of course it will also be written as $\mathbf{S}(x) = s(0) + s(1)x + s(2)x^2 + \cdots$. At this point x is merely a symbol, as are $\sum$ and $+$. It is only when a real number is substituted for x that one is talking about infinite series.

Functions defined by power series.

Definition A.13. The *domain of convergence* $\mathrm{Dom}(\mathbf{S})$ of a power series $\mathbf{S}$ is the set of reals a such that the series $\sum_n s(n)a^n$ converges.

Given a power series $\mathbf{S}$ let $f_{\mathbf{S}}$ be the real valued function defined on $\mathrm{Dom}(\mathbf{S})$ by $f_{\mathbf{S}}(a) = \sum_n s(n)a^n$. $f_{\mathbf{S}}$ is called a *power series function.*

Definition A.14. Given a power series $\mathbf{S}$ the *radius of convergence* $\rho_{\mathbf{S}}$ is given by

$$\rho_{\mathbf{S}} \;=\; 1/\limsup_{n\to\infty} |s(n)|^{1/n}.$$

($\rho_{\mathbf{S}}$ is allowed to be ∞.)

The well known result of Cauchy-Hadamard is:

$$(-\rho_{\mathbf{S}}, \rho_{\mathbf{S}}) \;\subseteq\; \mathrm{Dom}(\mathbf{S}) \;\subseteq\; [-\rho_{\mathbf{S}}, \rho_{\mathbf{S}}].$$

The following is a modification of Proposition A.8 to give the corresponding facts for power series.

Proposition A.15. *Let* $\mathbf{S}, \mathbf{T}$ *be power series and let* $a, c \in \mathbb{R}$. *Then one has*

$$\begin{aligned}
(cf_{\mathbf{S}})(a) &= \sum_n cs(n)a^n \qquad \textit{for } a \in \mathrm{Dom}(\mathbf{S})\\
(f_{\mathbf{S}} + f_{\mathbf{T}})(a) &= \sum_n \big(s(n) + t(n)\big)a^n \qquad \textit{for } a \in \mathrm{Dom}(\mathbf{S}) \cap \mathrm{Dom}(\mathbf{T})\\
(f_{\mathbf{S}} \cdot f_{\mathbf{T}})(a) &= \sum_n \Big(\sum_k s(k)\cdot t(n-k)\Big)a^n \qquad \textit{for } |a| < \min(\rho_{\mathbf{S}}, \rho_{\mathbf{T}})\\
(f_{\mathbf{S}} \circ f_{\mathbf{T}})(a) &= \sum_n \Big(\sum_k s(k) \sum_{\ell_1,\ldots,\ell_k \ge 0} t(\ell_1)\cdots t(\ell_k)\Big)a^n\\
&\qquad\qquad \textit{for } |a| < \rho_{\mathbf{T}} \textit{ and } |f_{\mathbf{T}}(a)| < \rho_{\mathbf{S}}\\
f_{\mathbf{S}}{}'(a) &= \sum_{n\ge 1} ns(n)a^{n-1} \qquad \textit{if } |a| < \rho_{\mathbf{S}}\\
s(n) &= f_{\mathbf{S}}^{(n)}(0)/n! \qquad \textit{if } \rho_{\mathbf{S}} > 0\\
\mathbf{S} = \mathbf{T} &\Leftrightarrow f_{\mathbf{S}} = f_{\mathbf{T}} \textit{ on a neighborhood of } 0 \quad (\textit{provided } \rho_{\mathbf{S}}, \rho_{\mathbf{T}} > 0).
\end{aligned}$$

PROOF. These are standard results about infinite series. In the item on addition, note that one can choose any $a \in \mathrm{Dom}(f_{\mathbf{S}}) \cap \mathrm{Dom}(f_{\mathbf{T}})$, even if a is an endpoint of the radius of convergence for one or both functions. However, for the multiplication item, the proof involves a consideration of the series of absolute values of the terms, and to be sure that this converges one is forced inside the interval of convergence. The same comment applies to the composition and differentiation of series. □

Operations on power series. Now Proposition A.15 is used to suggest appropriate definitions for operations on power series.

Definition A.16. Let $\mathbf{S}, \mathbf{T}$ be power series and let $c \in \mathbb{R}$.

coefficients: $\big[x^n\big]\mathbf{S}(x)$ is $s(n)$.

equality: $\mathbf{S} = \mathbf{T}$ holds iff $\big[x^n\big]\mathbf{S} = \big[x^n\big]\mathbf{T}$, for $n \ge 0$.

constant power series:

$$[x^n]\mathbf{c} \;=\; \begin{cases} c & \text{if } n = 0 \\ 0 & \text{otherwise.} \end{cases}$$

scalar multiplication: $[x^n](c\mathbf{S}) \;=\; cs(n)$.
addition: $[x^n](\mathbf{S}+\mathbf{T}) \;=\; s(n)+t(n)$.
multiplication: Use the Cauchy product:

$$[x^n](\mathbf{S} * \mathbf{T}) \;=\; \sum_{k\le n} s(k)\cdot t(n-k).$$

composition: $\mathbf{S}\circ\mathbf{T}$ will only be defined when $t(0) = 0$:[4]

$$[x^n](\mathbf{S}\circ\mathbf{T}) \;=\; \sum_{k\le n} s(k) \sum_{\ell_1+\ldots+\ell_k=n} t(\ell_1)\cdots t(\ell_k).$$

exponentiation: This is an instance of composition, and is defined only for $\mathbf{T}$ with $t(0) = 0$:

$$[x^n]e^{\mathbf{T}} \;=\; \begin{cases} 1 & \text{for } n = 0 \\ \displaystyle\sum_{k\le n} [x^n]\frac{1}{k!}\mathbf{T}^k & \text{for } n \ge 1. \end{cases}$$

differentiation: $[x^n]\mathbf{S}' \;=\; (n+1)s(n+1)$.

Connection between power series and functions. Operations on power series were defined to mimic what happens with power series functions, so one has the following translation of Proposition A.15.

Proposition A.17. *Let $\mathbf{S},\mathbf{T}$ be power series and let $a,c\in\mathbb{R}$. Then one has*

$$\begin{array}{lcll}
f_{c\mathbf{S}}(a) & = & cf_{\mathbf{S}}(a) & \textit{for } |a| < \rho_{\mathbf{S}} \\
f_{\mathbf{S}+\mathbf{T}}(a) & = & f_{\mathbf{S}}(a) + f_{\mathbf{T}}(a) & \textit{for } |a| < \min(\rho_{\mathbf{S}},\rho_{\mathbf{T}}) \\
f_{\mathbf{S}*\mathbf{T}}(a) & = & f_{\mathbf{S}}(a)\cdot f_{\mathbf{T}}(a) & \textit{for } |a| < \min(\rho_{\mathbf{S}},\rho_{\mathbf{T}}) \\
f_{\mathbf{S}\circ\mathbf{T}}(a) & = & (f_{\mathbf{S}}\circ f_{\mathbf{T}})(a) & \textit{for } |a| < \rho_{\mathbf{T}},\ |f_{\mathbf{T}}(a)| < \rho_{\mathbf{S}} \\
{f_{\mathbf{S}}}'(a) & = & f_{\mathbf{S}'}(a) & \textit{for } |a| < \rho_{\mathbf{S}}.
\end{array}$$

Truncation of power series. Since the function defined by a power series may easily have radius of convergence 0, one cannot proceed as with polynomials and simply say that all the identities of functions lift immediately. The key to lifting the results to power series is the observation that the fundamental operations are such that the definition of the nth coefficient in an application of an operation depends only on the first n coefficients of the given series, with the exception of differentiation which requires the $n+1$st coefficient.

[4]The restriction $t(0) = 0$ guarantees that the definition of the nth coefficient of the composition $\mathbf{S}\circ\mathbf{T}$ involves only finitely many of the coefficients of $\mathbf{S}$ and $\mathbf{T}$. [Some condition is needed to guarantee that $\sum_k s(k) \sum_{\ell_1,\ldots,\ell_k\ge 0} t(\ell_1)\cdots t(\ell_k)$ converges!]

Definition A.18. The mth *truncation* $\mathbf{S}|_m$ of a power series $\mathbf{S}$ is the polynomial defined by

$$[x^n]\mathbf{S}|_m = \begin{cases} s(n) & \text{if } n \le m \\ 0 & \text{if } n > m. \end{cases}$$

Thus $\mathbf{S}|_m(x) = s(0) + \cdots + s(m)x^m$.

Lemma A.19. *Given power series $\mathbf{S}, \mathbf{T}$, a constant $c \in \mathbb{R}$, and $n \ge 0$:*

$$\begin{aligned}
[x^n](c\mathbf{S}) &= [x^n](c\mathbf{S}|_m) && \textit{for } m \ge n \\
[x^n](\mathbf{S}+\mathbf{T}) &= [x^n](\mathbf{S}|_m + \mathbf{T}|_m) && \textit{for } m \ge n \\
[x^n](\mathbf{S}\cdot\mathbf{T}) &= [x^n](\mathbf{S}|_m \cdot \mathbf{T}|_m) && \textit{for } m \ge n \\
[x^n](\mathbf{S}\circ\mathbf{T}) &= [x^n](\mathbf{S}|_m \circ \mathbf{T}|_m) && \textit{for } m \ge n \\
[x^n]\mathbf{S}' &= [x^n](\mathbf{S}|_m{}') && \textit{for } m \ge n+1.
\end{aligned}$$

Identities of power series.

Proposition A.20. *For $a, b \in \mathbb{R}$ and power series $\mathbf{R}, \mathbf{S}, \mathbf{T}$ the following hold:*

scalar multiplication:

$$\begin{aligned}
a\mathbf{0} &= \mathbf{0} & a\mathbf{1} &= \mathbf{a} \\
(-1)\mathbf{R} &= -\mathbf{R} & a(-\mathbf{R}) &= (-a)\mathbf{R} = -(a\mathbf{R}) \\
a(b\mathbf{R}) &= (ab)\mathbf{R} & (a+b)\mathbf{R} &= (a\mathbf{R}) + (b\mathbf{R}) \\
a(\mathbf{R}+\mathbf{S}) &= (a\mathbf{R}) + (a\mathbf{S}) & a(\mathbf{R}\cdot\mathbf{S}) &= (a\mathbf{R})\cdot\mathbf{S} = \mathbf{R}\cdot(a\mathbf{S})
\end{aligned}$$

addition, multiplication, and composition:

$$\begin{aligned}
\mathbf{R}+\mathbf{0} &= \mathbf{R} & \mathbf{R}\cdot\mathbf{1} &= \mathbf{R} \\
\mathbf{R}+\mathbf{S} &= \mathbf{S}+\mathbf{R} & \mathbf{R}\cdot\mathbf{S} &= \mathbf{S}\cdot\mathbf{R} \\
(\mathbf{R}+\mathbf{S})+\mathbf{T} &= \mathbf{R}+(\mathbf{S}+\mathbf{T}) & (\mathbf{R}\cdot\mathbf{S})\cdot\mathbf{T} &= \mathbf{R}\cdot(\mathbf{S}\cdot\mathbf{T}) \\
\mathbf{R}\cdot(\mathbf{S}+\mathbf{T}) &= (\mathbf{R}\cdot\mathbf{S}) + (\mathbf{R}\cdot\mathbf{T}) & (\mathbf{R}+\mathbf{S})\circ\mathbf{T} &= (\mathbf{R}\circ\mathbf{T}) + (\mathbf{S}\circ\mathbf{T}) \\
(\mathbf{R}\cdot\mathbf{S})\circ\mathbf{T} &= (\mathbf{R}\circ\mathbf{T})\cdot(\mathbf{S}\circ\mathbf{T}) & (\mathbf{R}\circ\mathbf{S})\circ\mathbf{T} &= \mathbf{R}\circ(\mathbf{S}\circ\mathbf{T})
\end{aligned}$$

differentiation:

$$\begin{aligned}
(a\mathbf{R})' &= a\mathbf{R}' & (\mathbf{R}+\mathbf{S})' &= \mathbf{R}'+\mathbf{S}' \\
(\mathbf{R}\cdot\mathbf{S})' &= (\mathbf{R}'\cdot\mathbf{S}) + (\mathbf{R}\cdot\mathbf{S}') & (\mathbf{R}\circ\mathbf{S})' &= \mathbf{S}'\cdot(\mathbf{R}'\circ\mathbf{S}) \\
(e^{\mathbf{R}})' &= \mathbf{R}'\cdot e^{\mathbf{R}}.
\end{aligned}$$

Proof. To verify each identity one only has to show that the corresponding coefficients of the two sides are equal. This is straightforward with scalar multiplication and addition.

The commutativity of multiplication is easy to see from the definition. For the associativity let us employ truncations,[5] and use the fact that multiplication of polynomials is associative:

$$\begin{aligned}\big[x^n\big]\big((\mathbf{R}*\mathbf{S})*\mathbf{T}\big) &= \big[x^n\big]\big((\mathbf{R}|_n*\mathbf{S}|_n)*\mathbf{T}|_n\big) = \big[x^n\big]\big(\mathbf{R}|_n*(\mathbf{S}|_n*\mathbf{T}|_n)\big)\\ &= \big[x^n\big]\big(\mathbf{R}*(\mathbf{S}*\mathbf{T})\big).\end{aligned}$$

This technique of reduction to polynomials can be used for all the identities under the heading of 'addition, multiplication, and composition'. This can also be used for differentiation. The differentiation of scalar multiplication and of addition are straightforward. For the product rule first note that

$$\begin{aligned}\mathbf{S}|_{n+1}{}' &= \mathbf{S}'|_n\\ \big[x^n\big]\mathbf{S}' &= \big[x^n\big]\mathbf{S}|_{n+1}{}'.\end{aligned}$$

Then

$$\begin{aligned}\big[x^n\big](\mathbf{R}*\mathbf{S})' &= \big[x^n\big](\mathbf{R}*\mathbf{S})|_{n+1}{}'\\ &= \big[x^n\big]\big(\mathbf{R}|_{n+1}*\mathbf{S}|_{n+1}\big)'\\ &= \big[x^n\big]\big(\mathbf{R}|_{n+1}{}'*\mathbf{S}|_{n+1}+\mathbf{R}|_{n+1}*\mathbf{S}|_{n+1}{}'\big)\\ &= \big[x^n\big]\big(\mathbf{R}'|_n*\mathbf{S}|_{n+1}+\mathbf{R}|_{n+1}*\mathbf{S}'|_n\big)\\ &= \big[x^n\big]\big(\mathbf{R}'|_n*\mathbf{S}|_n+\mathbf{R}|_n*\mathbf{S}'|_n\big)\\ &= \big[x^n\big]\big(\mathbf{R}'*\mathbf{S}+\mathbf{R}*\mathbf{S}'\big).\end{aligned}$$

For the chain rule:

$$\begin{aligned}\big[x^n\big](\mathbf{R}\circ\mathbf{S})' &= \big[x^n\big](\mathbf{R}\circ\mathbf{S})|_{n+1}{}'\\ &= \big[x^n\big]\big(\mathbf{R}|_{n+1}\circ\mathbf{S}|_{n+1}\big)'\\ &= \big[x^n\big]\big(\mathbf{S}|_{n+1}{}'*(\mathbf{R}|_{n+1}{}'\circ\mathbf{S}|_{n+1})\big)\\ &= \big[x^n\big]\big(\mathbf{S}'|_n*(\mathbf{R}'|_n\circ\mathbf{S}|_{n+1})\big)\\ &= \big[x^n\big]\big(\mathbf{S}'|_n*(\mathbf{R}'|_n\circ\mathbf{S}|_n)\big)\\ &= \big[x^n\big]\big(\mathbf{S}'*(\mathbf{R}'\circ\mathbf{S})\big).\end{aligned}$$

Thus $\big(e^{\mathbf{R}}\big)' = \mathbf{R}'*e^{\mathbf{R}}$. □

An identity for number systems

For any additive number system $\mathcal{A}$, even with $\rho = 0$, the alternate form of the partition identity, $\mathbf{A}(x) = e^{\mathbf{H}(x)}$, still holds as an identity between power series. This will be used later (see Theorems C.3 and C.6).

[5]In these arguments using truncations the reader may find that the individual steps do not follow immediately from what has been written before, but they are not difficult to verify.

Proposition A.21. *For $\mathcal{A}$ an additive number system define* $\mathbf{H}(x) = \sum h(n)x^n$, *where* $h(0) = 0$ *and, for* $n \geq 1$, $h(n) = \sum_{ij=n} p(i)/j$. *Then*

$$\mathbf{A}(x) \;=\; e^{\mathbf{H}(x)}.$$

PROOF. Clearly both sides have the same constant term 1. So let $m \geq 1$ be given, and let $\mathcal{B} = (\mathsf{B}, \mathsf{Q}, +, 0, \|\ \|)$ be the subsystem of $\mathcal{A}$ generated by the set Q of indecomposables of size at most m. Let $\mathbf{H}_{\mathsf{B}}(x) = \sum h_{\mathsf{B}}(n)x^n$, where $h_{\mathsf{B}}(0) = 0$ and, for $n \geq 1$, $h_{\mathsf{B}}(n) = \sum_{ij=n} q(i)/j$, $q(n)$ being the counting function for Q.

Let $\mathbf{S}(x)$ be the power series $e^{\mathbf{H}_{\mathsf{B}}(x)}$. Since $\mathcal{B}$ is finitely generated the radius of convergence of $\mathbf{B}(x)$ is 1, so from §2.5 one has, for $x \in [0, 1)$,

$$f_{\mathbf{B}}(x) \;=\; f_{\mathbf{S}}(x).$$

This gives the equality of the power series $\mathbf{B}(x)$ and $\mathbf{S}(x)$, that is,

$$\mathbf{B}(x) \;=\; e^{\mathbf{H}_{\mathsf{B}}(x)}.$$

Now $a(1), \ldots, a(m)$ are completely determined by Q, and hence must agree with the corresponding $b(1), \ldots, b(m)$, that is, $b(j) = a(j)$, for $j \leq m$. Since $p(j) = q(j)$, for $1 \leq j \leq m$, it follows that $h_{\mathsf{B}}(j) = h(j)$, for $1 \leq j \leq m$. From this, for the given m, one has

$$\begin{aligned} a(m) \;=\; b(m) \;=\; \big[x^m\big]e^{\mathbf{H}_{\mathsf{B}}(x)} \;&=\; \big[x^m\big]e^{\mathbf{H}_{\mathsf{B}}|_m(x)} \\ &=\; \big[x^m\big]e^{\mathbf{H}|_m(x)} \;=\; \big[x^m\big]e^{\mathbf{H}(x)}. \end{aligned}$$

This proves the proposition. □

Division of power series

Proposition A.22. *If* $\mathbf{R}(x), \mathbf{S}(x)$ *are power series with* $s(0) \neq 0$ *then there is a unique power series* $\mathbf{T}(x)$ *such that* $\mathbf{R}(x) = \mathbf{S}(x) * \mathbf{T}(x)$. *One writes* $\mathbf{R}(x)/\mathbf{S}(x)$ *for* $\mathbf{T}(x)$, *and calls it the* quotient *of* $\mathbf{R}(x)$ *by* $\mathbf{S}(x)$.

PROOF. Given $\mathbf{R}(x) = \sum_n r(n)x^n$ and $\mathbf{S}(x) = \sum_n s(n)x^n$ with $s(0) \neq 0$, the goal is to show that there is a unique $\mathbf{T}(x) = \sum_n t(n)x^n$ such that $\mathbf{R}(x) = \mathbf{S}(x) * \mathbf{T}(x)$. This is equivalent to proving that there is a unique solution for $t(0), t(1), \ldots$ to the system of equations

$$r(n) = \sum_{k \leq n} s(k) \cdot t(n-k),$$

or, in more detail,

$$\begin{aligned} r(0) \;&=\; s(0)t(0) \\ r(1) \;&=\; s(0)t(1) + s(1) \cdot t(0) \\ r(2) \;&=\; s(0)t(2) + s(1) \cdot t(1) + s(2) \cdot t(0) \\ &\textit{etc.} \end{aligned}$$

Since $s(0) \neq 0$ it is clear that indeed one can recursively solve this system for the $t(n)$, and there is only one solution. □

The composition inverse of a power series

Inversion (under composition) can be carried out over a wide range of power series.

Proposition A.23.

(a) *If* $\mathbf{S}(x)$ *is such that* $s(0) = 0$ *and* $s(1) \neq 0$ *then there is a unique* $\mathbf{R}(x)$ *such that* $(\mathbf{R} \circ \mathbf{S})(x) = x$*. For this* $\mathbf{R}(x)$ *one has* $r(0) = 0$ *and* $r(1) = 1/s(1)$*.* $\mathbf{R}(x)$ *is the* left inverse *of* $\mathbf{S}(x)$ *under composition.*

(b) *Given* $\mathbf{S}(x)$ *with* $s(0) = 0$ *and* $s(1) \neq 0$ *there is a unique power series* $\mathbf{T}(x) = \sum_n t(n)x^n$ *with* $t(0) = 0$ *such that* $(\mathbf{S} \circ \mathbf{T})(x) = x$*.* $\mathbf{T}(x)$ *is the* right inverse *of* $\mathbf{S}(x)$*.*

PROOF. (a) This is equivalent to showing that there is a unique solution for the coefficients $r(0), r(1), \ldots$ in the infinite system of equations

$$\big[x^n\big](\mathbf{R} \circ \mathbf{S}) \;=\; \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise,} \end{cases}$$

or, in more detail,

$$\begin{array}{lcl} r(0) & = & 0 \\ r(1) \cdot s(1) & = & 1 \\ r(2) \cdot s(1)^2 + r(1) \cdot s(2) & = & 0 \\ & \vdots & \\ r(n) \cdot s(1)^n + q_n\big(r(1), \ldots, r(n-1), s(1), \ldots, s(n)\big) & = & 0 \\ & \vdots & \end{array}$$

for suitable polynomials q_n. Since $s(1) \neq 0$, clearly this system is uniquely solvable for the $r(n)$, starting with $r(0) = 0$, $r(1) = 1/s(1)$.

(b) Since $t(0) = 0$ this is equivalent to showing that there is a unique solution $t(1), t(2), \ldots$ to the infinite system of equations

$$\big[x^n\big](\mathbf{S} \circ \mathbf{T}) \;=\; \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise,} \end{cases}$$

or, in more detail,

$$\begin{array}{lcl} s(1) \cdot t(1) & = & 1 \\ s(1) \cdot t(2) + s(2) \cdot t(1)^2 & = & 0 \\ s(1) \cdot t(3) + 2s(2) \cdot t(1) \cdot t(2) + s(3) \cdot t(1)^3 & = & 0 \\ & \vdots & \\ s(1) \cdot t(n) + q_n\big(s(1), \ldots, s(n), t(1), \ldots, t(n-1)\big) & = & 0 \\ & \vdots & \end{array}$$

for suitable polynomials q_n. Since $s(1) \neq 0$, clearly this system is uniquely solvable for the $t(n)$, starting with $t(1) = 1/s(1)$. $\square$

A group under composition

Having established a left inverse $\mathbf{R}(x)$ and a right inverse $\mathbf{T}(x)$ of any power series $\mathbf{S}(x)$ satisfying $s(0) = 0$ and $s(1) \neq 0$, now one would like to show that these two inverses are the same. The simplest proof is to show that the collection of power series $\mathbf{S}(x)$ that satisfy $s(0) = 0$ and $s(1) \neq 0$ forms a group under composition.

Definition A.24. Let

$$\mathcal{G} \;=\; \{\mathbf{S}(x) : s(0) = 0,\ s(1) \neq 0\}.$$

Recall that a group is a collection of elements with a binary operation that is associative, there is a left and right identity element with respect to this operation, and every element has a two-sided (both left and right) inverse.

Proposition A.25. *$\mathcal{G}$ is a group under composition.*

PROOF. It is easy to see that $\mathcal{G}$ is closed under the binary operation of composition. Composition is an associative operation on $\mathcal{G}$, by Proposition A.20, and x is a two-sided identity, that is, $x \circ \mathbf{S} = \mathbf{S}$ and $\mathbf{S} \circ x = \mathbf{S}$. For $\mathbf{S} \in \mathcal{G}$, the left inverse of $\mathbf{S}$ in $\mathcal{G}$ exists and is unique, by Proposition A.23 (a), so let $\mathbf{S}^{-1}$ be this left inverse. Now a simple algebraic argument shows that $\mathbf{S}^{-1}$ is also a right inverse of $\mathbf{S}$, namely:

$$\begin{aligned}
\mathbf{S} \circ \mathbf{S}^{-1} &= x \circ \left(\mathbf{S} \circ \mathbf{S}^{-1}\right) \\
&= \left(\left(\mathbf{S}^{-1}\right)^{-1} \circ \mathbf{S}^{-1}\right) \circ \left(\mathbf{S} \circ \mathbf{S}^{-1}\right) \\
&= \left(\mathbf{S}^{-1}\right)^{-1} \circ \mathbf{S}^{-1} \circ \mathbf{S} \circ \mathbf{S}^{-1} \\
&= \left(\mathbf{S}^{-1}\right)^{-1} \circ x \circ \mathbf{S}^{-1} \\
&= \left(\mathbf{S}^{-1}\right)^{-1} \circ \mathbf{S}^{-1} \\
&= x.
\end{aligned}$$

□

Since the left inverse of a power series in $\mathcal{G}$ is a right inverse, and since the right inverse is also unique, one simply refers to *the* inverse of $\mathbf{S} \in \mathcal{G}$.

Now the goal is to prove[6] that if $\mathbf{S} \in \mathcal{G}$ has a positive radius of convergence then so does $\mathbf{S}^{-1}$.

A special case. First the special case where $s(1) = 1$ is considered:

$$\mathbf{S}(x) \;=\; x + s(2)x^2 + s(3)x^3 + \cdots$$

Lemma A.26. *Suppose that $\mathbf{S}(x) = x + \sum_{n \geq 2} s(n)x^n$ and $\mathbf{R}(x) = \sum_n r(n)x^n$ is the inverse of $\mathbf{S}$ in $\mathcal{G}$. If $\rho_{\mathbf{S}} > 0$ then $\rho_{\mathbf{R}} > 0$.*

[6]This proof follows Knopp [**34**], pp. 184–188, which he says is due to Cauchy. Two other proofs can be found in Hille [**27**], an intermediate level proof on pp. 86–90, where one uses successive approximations; and a more advanced level proof on p. 265, where one gets an explicit form for the right inverse using an integral in the complex plane.

PROOF. $\mathbf{R}(x)$, as a left inverse, is specified by the system of equations

$$\begin{array}{lcl} r(0) & = & 0 \\ s(1)\cdot r(1) & = & 1 \\ s(1)^2\cdot r(2)+s(2)\cdot r(1) & = & 0 \\ s(1)^3\cdot r(3)+2s(1)\cdot s(2)\cdot r(2)+s(3)\cdot r(1) & = & 0 \\ & \vdots & \\ s(1)^n\cdot r(n)+q_n\big(s(1),\ldots,s(n),r(1),\ldots,r(n-1)\big) & = & 0 \\ & \vdots & \end{array}$$

Let $\lambda = 1 + 1/\rho_{\mathbf{S}}$. Then, for some K,

$$|s(n)| \;\le\; K\lambda^n, \qquad \text{for } n \ge 0.$$

Now define the power series $\mathbf{U}(x)$ by

$$\mathbf{U}(x) \;=\; x - \sum_{n\ge 2}(K\lambda^n)x^n.$$

$\mathbf{U}(x)$ is clearly in $\mathcal{G}$, so let $\mathbf{V}(x) = \sum_n v(n)x^n$ be its inverse. As a left inverse one has

$$\begin{array}{lcl} v(0) & = & 0 \\ v(1) & = & 1 \\ v(2)-(K\lambda^2)v(1) & = & 0 \\ v(3)-2(K\lambda^2)v(2)-(K\lambda^3)v(1)^3 & = & 0 \\ & \vdots & \end{array}$$

so the $v(n)$ are positive, and indeed, $v(n) \ge |r(n)|$. Thus

(A.26.1) $$\rho_{\mathbf{V}} > 0 \;\Rightarrow\; \rho_{\mathbf{R}} > 0.$$

There is a compact description of the function $f_{\mathbf{U}}(x)$ since there is a geometric series in the definition:

$$f_{\mathbf{U}}(x) \;=\; x - K\lambda^2x^2\Big(\frac{1}{1-\lambda x}\Big), \qquad \text{for } |x| < 1/\lambda.$$

A routine calculation shows that $f_{\mathbf{U}}{}'(0) = 1$, so $f_{\mathbf{U}}(x)$ is a one-one function in some neighborhood of 0. There is a simple formula for the inverse function, namely

$$\begin{aligned} y = f_{\mathbf{U}}(x) \;&\Rightarrow\; y - \lambda yx \;=\; x - \lambda(\lambda K+1)x^2 \\ &\Rightarrow\; \lambda(\lambda K+1)x^2 \;-\; (1+\lambda y)x \;+\; y \;=\; 0 \\ &\Rightarrow\; x \;=\; \frac{(1+\lambda y) \pm \sqrt{(1+\lambda y)^2 - 4\lambda(\lambda K+1)y}}{2\lambda(\lambda K+1)}. \end{aligned}$$

So if

$$g(x) \;=\; \frac{(1+\lambda x) - \sqrt{(1+\lambda x)^2 - 4\lambda(\lambda K+1)x}}{2\lambda(\lambda K+1)}$$

then $g(0) = 0$, g is defined on a neighborhood of 0 since the discriminant is a continuous function of x and takes on the value 1 at $x = 0$, and

$$f_{\mathbf{U}}\big(g(x)\big) \;=\; x$$

on some neighborhood of 0.

Let

$$g_0(x) = \sqrt{(1+\lambda x)^2 - 4\lambda(\lambda K + 1)x}.$$

Then $g_0(x)$ can be written in the form $\sqrt{1-q(x)}$, where $q(x)$ is a polynomial in x with $f_q(0) = 0$. Since $\sqrt{1-x}$ has a power series expansion on a neighborhood of 0, by the result on composition of functions defined by power series in Proposition A.17, $g_0(x) = \big(\sqrt{1-x}\big) \circ q(x)$ also has a power series expansion on a neighborhood of 0. Thus $g(x)$ has a power series expansion on some neighborhood of 0.

Let $\mathbf{W}(x) = \sum_n w(n)x^n$ be the power series expansion of $g(x)$. Then $\rho_{\mathbf{W}} > 0$, $w(0) = 0$, and $g(x) = f_{\mathbf{W}}(x)$ on some neighborhood of 0. From $f_{\mathbf{U}}\big(g(x)\big) = x$ on some neighborhood of 0 follows $f_{\mathbf{U}}\big(f_{\mathbf{W}}(x)\big) = x$ on a neighborhood of 0, and thus $\big(f_{\mathbf{U}} \circ f_{\mathbf{W}}\big)(x) = x$ on a neighborhood of 0. Now $\big(f_{\mathbf{U}} \circ f_{\mathbf{W}}\big)(x) = f_{\mathbf{U}\circ\mathbf{W}}(x)$ in that neighborhood, so $f_{\mathbf{U}\circ\mathbf{W}}(x) = x$ in a neighborhood of 0. This implies that $(\mathbf{U} \circ \mathbf{W})(x) = x$, so $\mathbf{W}(x)$ is the inverse of $\mathbf{U}(x)$.

But $\mathbf{V}(x)$ was defined to be that inverse. Consequently $\mathbf{V}(x) = \mathbf{W}(x)$, and thus $\rho_{\mathbf{V}} > 0$. From (A.26.1) it follows that $\rho_{\mathbf{R}} > 0$. □

Removing the condition $s(1) = 1$**.** Now the goal is to lift these results to any power series in $\mathcal{G}$.

Proposition A.27. *Suppose* $\mathbf{S} \in \mathcal{G}$*, that is,* $s(0) = 0$ *and* $s(1) \neq 0$*. If* $\rho_{\mathbf{S}} > 0$*, that is, if* $\mathbf{S}$ *has positive radius of convergence, then the composition inverse of* $\mathbf{S}$ *also has a positive radius of convergence.*

PROOF. Let $\mathbf{U}(x) = \sum_n u(n)x^n$ be the power series defined by

$$\mathbf{U}(x) \;=\; \mathbf{S}(x) \circ \Big(\frac{1}{s(1)}x\Big) \;=\; \sum_n \frac{s(n)}{s(1)^n}x^n.$$

Then $\rho_{\mathbf{U}} = s(1)\rho_{\mathbf{S}} > 0$ and $u(0) = 0$, $u(1) = 1$. Thus, by Lemma A.26, it follows that the inverse $\mathbf{V}(x) = \sum_n v(n)x^n$ of $\mathbf{U}(x)$ has a positive radius of convergence. Of course $v(0) = 0$ and $v(1) = 1$.

Define $\mathbf{T}(x)$ to be the power series $\big(1/s(1)\big)\mathbf{V}(x)$, which can be written as $\big(x/s(1)\big) \circ \mathbf{V}(x)$. Then $\rho_{\mathbf{T}} = \rho_{\mathbf{V}} > 0$, and

$$\begin{aligned}
\mathbf{S}(x) \circ \mathbf{T}(x) \;&=\; \mathbf{S}(x) \circ \Big(\frac{1}{s(1)}x\Big) \circ \mathbf{V}(x) \\
&=\; \mathbf{U}(x) \circ \mathbf{V}(x) \\
&=\; x.
\end{aligned}$$

This means that $\mathbf{T}$ is the inverse of $\mathbf{S}$, and it has a positive radius of convergence. □

This result will be used to prove Theorem C.3 of Bell, which implies that one is fully justified in avoiding additive number systems with radius of convergence zero when searching for systems with all partition sets having asymptotic density.

APPENDIX B

Refined Counting

The additive case

This topic concerns the number of elements in A of norm n that can be expressed as a sum of m indecomposables.

Definition B.1. The *refined (local) counting function* $a(n,m)$ of A is defined on $\mathbb{N}\times\mathbb{N}$ by

$$a(n,m) \;=\; \big|\{\mathsf{a}\in m\mathsf{P} : \|\mathsf{a}\| = n\}\big|.$$

Corresponding to the function $a(n,m)$ there is the *two variable generating series*

$$\mathbf{A}(x,y) \;=\; \sum_{m,n\geq 0} a(n,m)x^n y^m.$$

This can also be expressed by

$$\mathbf{A}(x,y) \;=\; \sum_{m\geq 0}\sum_{\mathsf{a}\in m\mathsf{P}} x^{\|\mathsf{a}\|}y^m.$$

The *two variable fundamental identity* of $\mathcal{A}$ is

$$\sum_{m,n\geq 0} a(n,m)x^n y^m \;=\; \prod_{n\geq 1}(1-x^n y)^{-p(n)}. \tag{B.1.1}$$

The next section shows that this identity describes a method to compute the values of $a(n,m)$ from $p(n)$, and the following section gives a proof that it holds as an identity between two functions when $0\leq x<\rho$ and $0\leq y<1$.

Meaning of the two variable fundamental identity. The meaning of the two variable fundamental identity (B.1.1), initially, is just that it is a shorthand device to remember how to compute $a(n,m)$ from $p(n)$. To this end think of each factor

$$(1-x^n y)^{-p(n)}$$

in the product part of the fundamental identity as a power series to the power $p(n)$, namely

$$(1-x^n y)^{-p(n)} \;=\; \big(1+x^n y + x^{2n}y^2+\cdots\big)^{p(n)}.$$

What follows explains why $a(n,m)$ is the coefficient of $x^n y^m$ in the expansion of

$$\big(1 + x + x^2 + \cdots\big)^{p(1)} \ \cdots \ \big(1 + x^k + x^{2k}y^2 + \cdots\big)^{p(k)},$$

for *any* choice of $k \geq n$. Actually, it is more transparent to say that $a(n,m)$ is the coefficient of x^n in the expansion of

$$\prod_{\|\mathsf{p}\| \leq k} \big(1 + x^{\|\mathsf{p}\|}y + x^{2\|\mathsf{p}\|}y^2 + \cdots\big),$$

for any choice of $k \geq n$.

To see that this is true proceed as follows. Let $\mathsf{p}_0, \mathsf{p}_1, \ldots$ be a listing of the indecomposable elements of $\mathcal{A}$. As before, one has a natural one-to-one correspondence between the elements of A and the sequences (m_i) of nonnegative integers with only finitely many of the m_i nonzero, namely (m_i) corresponds to the element a such that m_i is the coefficient of p_i in the decomposition of a into indecomposables. Thus the sequences $\big(x^{m_i\|\mathsf{p}_i\|}y^{m_i}\big)$ of monomials, where almost all monomials are 1, are also naturally in a one-to-one correspondence with A.

Now $a(n,m)$ is the number of sequences (m_i) of nonnegative integers such that $n = \sum m_i\|\mathsf{p}_i\|$ and $m = \sum m_i$. This is because each element of $m\mathsf{P}$ with norm n can be expressed as a sum $\sum m_i\mathsf{p}_i$ with $\sum m_i\|\mathsf{p}_i\| = n$ and $\sum m_i = m$, and distinct sequences (m_i) with $\sum m_i = m$ give rise to distinct elements $\sum m_i\mathsf{p}_i$ of $m\mathsf{P}$. Note that

$$n = \sum m_i\|\mathsf{p}_i\| \quad \text{and} \quad m = \sum m_i \quad \text{iff} \quad x^n y^m = \prod x^{m_i\|\mathsf{p}_i\|}y^{m_i}.$$

Thus $a(n,m)$ is the number of sequences (m_i) of nonnegative integers such that $x^n y^m = \prod x^{m_i\|\mathsf{p}_i\|}y^{m_i}$. Now consider the product

$$\text{(B.1.2)} \quad \Big(1 + x^{\|\mathsf{p}_0\|}y + x^{2\|\mathsf{p}_0\|}y^2 + \cdots\Big) \ \cdots \ \Big(1 + x^{\|\mathsf{p}_k\|}y + x^{2\|\mathsf{p}_k\|}y^2 + \cdots\Big),$$

where k is sufficiently large to ensure that all indecomposables p with norm at most n appear in the list $\mathsf{p}_0, \ldots, \mathsf{p}_k$. Then each sequence (m_i) with $x^n y^m = \prod x^{m_i\|\mathsf{p}_i\|}y^{m_i}$ corresponds to a choice of one monomial from each of the sums $1 + x^{\|\mathsf{p}_i\|}y + x^{2\|\mathsf{p}_i\|}y^2 + \cdots$ above, and thus $a(n,m)$ is the coefficient of $x^n y^m$ in the expansion of (B.1.2).

Validity of the two variable fundamental identity. Let $\mathcal{A}$ be an additive number system with radius of convergence ρ.

Proposition B.2. *The two variable fundamental identity*

$$\sum_{m,n \geq 0} a(n,m)x^n y^m \ = \ \prod_{n \geq 1}(1 - x^n y)^{-p(n)}$$

of $\mathcal{A}$ holds for $0 \leq x, y < 1$, with convergence if $0 \leq x < \rho$.

PROOF. This follows from the observation that, for $x, y \geq 0$, and, for any $\ell \geq 1$,

$$\sum_{0 \leq m,n \leq \ell} a(n,m)x^n y^m \;\leq\; \prod_{1 \leq j \leq \ell} \Big(\sum_{k \geq 0} x^{kj} y^k\Big)^{p(j)} \;\leq\; \sum_{m,n \geq 0} a(n,m)x^n y^m.$$

Thus, for $0 \leq x, y < 1$,

$$\sum_{0 \leq m,n \leq \ell} a(n,m)x^n y^m \;\leq\; \prod_{1 \leq j \leq \ell} (1 - x^j y)^{-p(j)} \;\leq\; \sum_{m,n \geq 0} a(n,m)x^n y^m,$$

so,

$$\sum_{m,n \geq 0} a(n,m)x^n y^m \;=\; \prod_{n \geq 1} (1 - x^n y)^{-p(n)}.$$

If $0 \leq x < \rho$ and $0 \leq y < 1$ then

$$\begin{aligned}
\sum_{m,n} a(n,m)x^n y^m \;&\leq\; \sum_{m,n} a(n,m)x^n \\
&=\; \sum_{n} \Big(\sum_{m} a(n,m)\Big)x^n \\
&=\; \sum_{n} a(n)x^n \\
&=\; \mathbf{A}(x) \;<\; \infty.
\end{aligned}$$

□

Alternate version of the two variable fundamental identity. Just as with the fundamental identity, there is a useful alternate form involving exponentiation.

Proposition B.3. *Let $\mathcal{A}$ be an additive number system with radius of convergence $\rho > 0$. Then the identity*

$$\mathbf{A}(x,y) \;=\; \exp\Big(\sum_{m \geq 1} \mathbf{P}(x^m)y^m/m\Big)$$

holds when $0 \leq x < \rho$ and $0 \leq y < 1$.

PROOF. For $0 \leq x < \rho$ and $0 \leq y < 1$,

$$\begin{aligned}
(1 - x^n y)^{-p(n)} \;&=\; \exp\Big(\log\big((1 - x^n y)^{-p(n)}\big)\Big) \\
&=\; \exp\big(-p(n)\log(1 - x^n y)\big) \\
&=\; \exp\Big(p(n)\sum_{m \geq 1} x^{mn} y^m/m\Big),
\end{aligned}$$

thus,

$$\prod_{1 \leq n \leq M} (1 - x^n y)^{-p(n)} \;=\; \prod_{1 \leq n \leq M} \exp\Big(p(n)\sum_{m \geq 1} x^{mn} y^m/m\Big).$$

So, from Proposition B.2,

$$\begin{aligned}
\mathbf{A}(x,y) &= \prod_{n\geq 1}\bigl(1-x^n y\bigr)^{-p(n)} \\
&= \lim_{M\to\infty}\prod_{1\leq n\leq M}\exp\Bigl(\sum_{m\geq 1}p(n)x^{mn}y^m/m\Bigr) \\
&= \lim_{M\to\infty}\exp\Bigl(\sum_{1\leq n\leq M}\sum_{m\geq 1}p(n)x^{mn}y^m/m\Bigr) \\
&= \exp\Bigl(\sum_{n\geq 1}\sum_{m\geq 1}p(n)x^{mn}y^m/m\Bigr) \\
&= \exp\Bigl(\sum_{m\geq 1}\sum_{n\geq 1}p(n)x^{mn}y^m/m\Bigr) \\
&= \exp\Bigl(\sum_{m\geq 1}\mathbf{P}(x^m)y^m/m\Bigr).
\end{aligned}$$

□

The generating series of $m\mathsf{P}$. The generating series of $m\mathsf{P}$ can be expressed as a polynomial in $\mathbf{P}(x),\ldots,\mathbf{P}(x^m)$ with nonnegative rational coefficients. The case $m=0$ is trivial as $[0\mathbf{P}](x)=1$, so a proof is only needed when $m>0$.

Proposition B.4. *Given a positive integer m, for any $x\geq 0$,*

$$[m\mathbf{P}](x) = \sum_{j=1}^{m}\frac{1}{j!}\sum_{\substack{m_1+\cdots+m_j=m\\ m_i\geq 1}}\frac{1}{m_1\cdots m_j}\mathbf{P}(x^{m_1})\cdots\mathbf{P}(x^{m_j}), \tag{B.4.1}$$

and the two sides converge iff $\mathbf{P}(x)$ converges.

PROOF. Since

$$[m\mathbf{P}](x) = \sum_{n\geq 0}a(n,m)x^n,$$

it follows that

$$\mathbf{A}(x,y) = \sum_{m\geq 0}[m\mathbf{P}](x)y^m, \tag{B.4.2}$$

so, from Proposition B.3,

$$\begin{aligned}
\sum_{m\geq 0}[m\mathbf{P}](x)y^m &= \exp\Bigl(\sum_{m\geq 1}\mathbf{P}(x^m)y^m/m\Bigr) \\
&= \sum_{j\geq 0}\frac{1}{j!}\Bigl(\sum_{m\geq 1}\mathbf{P}(x^m)y^m/m\Bigr)^j. \qquad \text{(B.4.3)}
\end{aligned}$$

This means that one obtains $[m\mathbf{P}](x)$ by collecting the terms from (B.4.3) that contribute to the coefficient of y^m, and this gives equation (B.4.1), at least for $0 \le x < \rho$.

To see that (B.4.1) holds when $x \ge 0$, with convergence precisely when $\mathbf{P}(x)$ converges, note that $\mathbf{P}(x), \dots, \mathbf{P}(x^m)$ are power series with nonnegative integer coefficients, so, when $x \ge 0$, if they converge at x then they converge absolutely. Furthermore, if $\mathbf{P}(x)$ converges at $x \ge 0$ then so do each of $\mathbf{P}(x^2), \dots, \mathbf{P}(x^m)$. After noting that one of the summands in the right side of (B.4.1) is

$$\frac{1}{m!}\mathbf{P}(x)^m$$

(this comes from $j = m$), it follows that, for $x \ge 0$, the right side of (B.4.1) converges precisely when $\mathbf{P}(x)$ converges. By Lemma 1.19, there is a power series $\mathbf{R}(x)$ such that the right side of (B.4.1) equals $\mathbf{R}(x)$ when $x \ge 0$, so

$$[m\mathbf{P}](x) \;=\; \mathbf{R}(x)$$

on $[0, \rho)$. Since $[m\mathbf{P}](x)$ and $\mathbf{R}(x)$ are both power series, they must be the same power series. Thus (B.4.1) holds when $x \ge 0$, with the two sides converging precisely when $\mathbf{P}(x)$ converges. $\square$

Here are two examples to illustrate (B.4.1):

$$\begin{aligned}
[2\mathbf{P}](x) &= \tfrac{1}{2}\mathbf{P}(x)^2 + \tfrac{1}{2}\mathbf{P}(x^2)\\
[3\mathbf{P}](x) &= \tfrac{1}{6}\mathbf{P}(x)^3 + \tfrac{1}{2}\mathbf{P}(x)\cdot\mathbf{P}(x^2) + \tfrac{1}{3}\mathbf{P}(x^3).
\end{aligned}$$

The reader may have noticed that the sums of the coefficients in each of the two examples is 1. This is no accident. For from (B.4.1) there is a polynomial $\mathbf{r}(y_1, \dots, y_m)$ with rational coefficients such that

$$[m\mathbf{P}](x) \;=\; \mathbf{r}\big(\mathbf{P}(x), \dots, \mathbf{P}(x^m)\big).$$

The sum of the coefficients is $\mathbf{r}(1, \dots, 1)$. From (B.4.3)

$$[m\mathbf{P}](x) \;=\; \big[y^m\big]\sum_{j\ge 0}\frac{1}{j!}\Big(\sum_{k\ge 1}\mathbf{P}(x^k)y^k/k\Big)^j.$$

Thus

$$\begin{aligned}
\mathbf{r}(1, \dots, 1) &= \big[y^m\big]\sum_{j\ge 0}\frac{1}{j!}\Big(\sum_{k\ge 1}y^k/k\Big)^j\\
&= \big[y^m\big]\exp\Big(\sum_{k\ge 1}y^k/k\Big)\\
&= \big[y^m\big]e^{-\log(1-y)}\\
&= \big[y^m\big]\frac{1}{1-y} \;=\; 1,
\end{aligned}$$

so $\mathbf{r}(1, \dots, 1) = 1$.

The multiplicative case

This section concerns the number of elements in A of norm n that can be expressed as a product of m indecomposables.

Definition B.5. The *refined (local) counting function* $a(n,m)$ of A, defined on $\mathbb{P} \times \mathbb{N}$, is given by

$$a(n,m) \;=\; \big|\{\mathsf{a} \in \mathsf{P}^m : \|\mathsf{a}\| = n\}\big|.$$

Corresponding to the function $a(n,m)$ there is the *two variable generating series*

$$\mathbf{A}(x,y) \;=\; \sum_{\substack{m\geq 0\\ n\geq 1}} a(n,m)n^{-x}y^m.$$

This can also be expressed by

$$\mathbf{A}(x,y) \;=\; \sum_{m\geq 0}\sum_{\mathsf{a}\in\mathsf{P}^m} \|\mathsf{a}\|^{-x}y^m.$$

The *two variable fundamental identity* of $\mathcal{A}$ is

$$\sum_{\substack{m\geq 0\\ n\geq 1}} a(n,m)n^{-x}y^m \;=\; \prod_{n\geq 2}(1-n^{-x}y)^{-p(n)}. \tag{B.5.1}$$

The next section shows that this identity describes a method to compute the values of $a(n,m)$ from $p(n)$, and the following section gives a proof that it holds as an identity between two functions when $x > \alpha$ and $0 \leq y < 1$.

Meaning of the two variable fundamental identity. The meaning of the two variable fundamental identity (B.5.1), initially, is just that it is a shorthand device to remember how to compute $a(n,m)$ from $p(n)$. (This is a 'formal' aspect of the identity—in the next section convergence is examined.) To this end think of each factor

$$(1-n^{-x}y)^{-p(n)}$$

in the product part of the fundamental identity as a geometric series to the power $p(n)$, namely,

$$(1-n^{-x}y)^{-p(n)} \;=\; \big(1+n^{-x}y+n^{-2x}y^2+\cdots\big)^{p(n)}.$$

What follows explains why $a(n,m)$ is the coefficient of $n^{-x}y^m$ in the expansion of

$$\big(1+2^{-x}y+2^{-2x}y^2\cdots\big)^{p(2)}\;\cdots\;\big(1+k^{-x}y+k^{-2x}y^2+\cdots\big)^{p(k)},$$

for *any* choice of $k \geq n$. Actually it is more transparent to say that $a(n,m)$ is the coefficient of $n^{-x}y^m$ in the expansion of

$$\prod_{\|\mathsf{p}\| \leq k} \left(1 + \|\mathsf{p}\|^{-x}y + \|\mathsf{p}\|^{-2x}y^2 + \cdots\right),$$

for any choice of $k \geq n$.

To see that this is true proceed as follows. Let $\mathsf{p}_0, \mathsf{p}_1, \ldots$ be a listing of the indecomposable elements of $\mathcal{A}$. As before, one has a natural one-to-one correspondence between the elements of A and the sequences (m_i) of nonnegative integers with only finitely many of the m_i nonzero, namely, (m_i) corresponds to the element a such that m_i is the exponent of p_i in the factorization of a into indecomposables. Thus the sequences $\left(\|\mathsf{p}_i\|^{-m_i x} y^{m_i}\right)$ of monomials, where almost all monomials are 1, are also naturally in a one-to-one correspondence with A.

Now $a(n,m)$ is the number of sequences (m_i) of nonnegative integers such that $n = \prod \|\mathsf{p}_i\|^{m_i}$ and $m = \sum m_i$. This is because each element of P^m with norm n can be expressed as a product $\prod \mathsf{p}_i^{m_i}$ with $\prod \|\mathsf{p}_i\|^{m_i} = n$ and $\sum m_i = m$, and distinct sequences (m_i) with $\sum m_i = m$ give rise to distinct elements $\prod \mathsf{p}_i^{m_i}$ of P^m. Note that

$$n = \prod \|\mathsf{p}_i\|^{m_i} \quad \text{and} \quad m = \sum m_i \quad \text{iff} \quad n^{-x}y^m = \prod \|\mathsf{p}_i\|^{-m_i x} y^{m_i}.$$

Thus $a(n,m)$ is the number of sequences (m_i) of nonnegative integers such that $n^{-x}y^m = \prod \|\mathsf{p}_i\|^{-m_i x} y^{m_i}$. Now consider the product

(B.5.2)
$$\left(1 + \|\mathsf{p}_0\|^{-x}y + \|\mathsf{p}_0\|^{-2x}y^2 + \cdots\right) \cdots \left(1 + \|\mathsf{p}_k\|^{-x}y + \|\mathsf{p}_k\|^{-2x}y^2 + \cdots\right),$$

where k is sufficiently large to ensure that all indecomposables p with norm at most n appear in the list $\mathsf{p}_0, \ldots, \mathsf{p}_k$. Then each sequence (m_i) with $n^{-x}y^m = \prod \|\mathsf{p}_i\|^{-m_i x} y^{m_i}$ corresponds to a choice of one monomial from each of the sums $1 + \|\mathsf{p}_i\|^{-x}y + \|\mathsf{p}_i\|^{-2x}y^2 + \cdots$ above, and thus $a(n,m)$ is the coefficient of $n^{-x}y^m$ in the expansion of (B.5.2).

Validity of the two variable fundamental identity. Let $\mathcal{A}$ be a multiplicative number system with abscissa of convergence α.

Proposition B.6. *The two variable fundamental identity*

$$\sum_{\substack{m \geq 0 \\ n \geq 1}} a(n,m) n^{-x} y^m \;=\; \prod_{n \geq 2} (1 - n^{-x}y)^{-p(n)}$$

of $\mathcal{A}$ holds for $x \geq 0$ and $0 \leq y < 1$, with convergence if $x > \alpha$.

PROOF. This follows from the observation that, for $x \geq 0$ and $0 \leq y < 1$, and for any $\ell \geq 1$,

$$\sum_{\substack{0 \leq m \leq \ell \\ 1 \leq n \leq \ell}} a(n,m) n^{-x} y^m \;\leq\; \prod_{2 \leq k \leq \ell} \Big(\sum_{j \geq 0} k^{-jx} y^j \Big)^{p(k)} \;\leq\; \sum_{\substack{m \geq 0 \\ n \geq 1}} a(n,m) n^{-x} y^m,$$

and thus,

$$\sum_{\substack{0 \leq m \leq \ell \\ 1 \leq n \leq \ell}} a(n,m) n^{-x} y^m \;\leq\; \prod_{2 \leq k \leq \ell} (1 - k^{-x} y)^{-p(k)} \;\leq\; \sum_{\substack{m \geq 0 \\ n \geq 1}} a(n,m) n^{-x} y^m,$$

so,

$$\sum_{\substack{m \geq 0 \\ n \geq 1}} a(n,m) n^{-x} y^m \;=\; \prod_{n \geq 2} (1 - n^{-x} y)^{-p(n)}.$$

If $x > \alpha$ and $0 \leq y < 1$ then

$$\begin{aligned} \sum_{m,n} a(n,m) n^{-x} y^m \;&\leq\; \sum_{m,n} a(n,m) n^{-x} \\ &=\; \sum_n \Big(\sum_m a(n,m) \Big) n^{-x} \\ &=\; \sum_n a(n) n^{-x} \\ &=\; \mathbf{A}(x) \;<\; \infty. \end{aligned}$$

$\square$

Alternate version of the two variable fundamental identity. Just as with the fundamental identity, there is a useful alternate form involving exponentiation.

Proposition B.7. *Let $\mathcal{A}$ be a multiplicative number system with abscissa of convergence $\alpha < \infty$. Then the identity*

$$\mathbf{A}(x,y) \;=\; \exp\Big(\sum_{m \geq 1} \mathbf{P}(mx) y^m / m \Big)$$

holds when $x > \alpha$ and $0 \leq y < 1$.

PROOF. For $x > \alpha$ and $0 \leq y < 1$,

$$\begin{aligned} (1 - n^{-x} y)^{-p(n)} \;&=\; \exp\Big(\log\big((1 - n^{-x} y)^{-p(n)} \big) \Big) \\ &=\; \exp\big(-p(n) \log(1 - n^{-x} y) \big) \\ &=\; \exp\Big(p(n) \sum_{m \geq 1} n^{-mx} y^m / m \Big), \end{aligned}$$

thus,

$$\prod_{1\le n\le M}\left(1-n^{-x}y\right)^{-p(n)} = \prod_{1\le n\le M}\exp\Big(p(n)\sum_{m\ge1}n^{-mx}y^m/m\Big).$$

So, from Proposition B.6,

$$\begin{aligned}
\mathbf{A}(x,y) &= \prod_{n\ge2}\left(1-n^{-x}y\right)^{-p(n)}\\
&= \lim_{M\to\infty}\prod_{2\le n\le M}\exp\Big(\sum_{m\ge1}p(n)n^{-mx}y^m/m\Big)\\
&= \lim_{M\to\infty}\exp\Big(\sum_{2\le n\le M}\sum_{m\ge1}p(n)n^{-mx}y^m/m\Big)\\
&= \exp\Big(\sum_{n\ge2}\sum_{m\ge1}p(n)n^{-mx}y^m/m\Big)\\
&= \exp\Big(\sum_{m\ge1}\sum_{n\ge2}p(n)n^{-mx}y^m/m\Big)\\
&= \exp\Big(\sum_{m\ge1}\mathbf{P}(mx)y^m/m\Big).
\end{aligned}$$

□

The generating series of P^m. Now it will be proved that the generating series of P^m can be expressed as a polynomial in $\mathbf{P}(x),\dots,\mathbf{P}(mx)$ with nonnegative rational coefficients. The case $m=0$ is trivial as one has $\mathbf{P}^0(x)=1$, so a proof is only needed when $m>0$.

Proposition B.8. *Given a positive integer m, for any $x\ge0$,*

$$\mathbf{P}^m(x) = \sum_{j=1}^{m}\frac{1}{j!}\sum_{\substack{m_1+\cdots+m_j=m\\ m_i\ge1}}\frac{1}{m_1\cdots m_j}\mathbf{P}(m_1x)\cdots\mathbf{P}(m_jx), \tag{B.8.1}$$

and the two sides converge iff $\mathbf{P}(x)$ converges.

Proof. Since $\mathbf{P}^m(x) = \sum_{n\ge2}a(n,m)n^{-x}$, it follows that

$$\mathbf{A}(x,y) = \sum_{m\ge0}\mathbf{P}^m(x)y^m, \tag{B.8.2}$$

so, from Proposition B.7,

$$\begin{aligned}
\sum_{m\ge0}\mathbf{P}^m(x)y^m &= \exp\Big(\sum_{m\ge1}\mathbf{P}(mx)y^m/m\Big)\\
&= \sum_{j\ge0}\frac{1}{j!}\Big(\sum_{m\ge1}\mathbf{P}(mx)y^m/m\Big)^j. \qquad \text{(B.8.3)}
\end{aligned}$$

This means that one has $\mathbf{P}^m(x)$ by collecting the terms from (B.8.3) that contribute to the coefficient of y^m, and this gives equation (B.8.1), at least for $x > \alpha$.

To see that (B.8.1) holds when $x \geq 0$, with convergence precisely when $\mathbf{P}(x)$ converges, note that $\mathbf{P}(x), \ldots, \mathbf{P}(mx)$ are Dirichlet series with non-negative integer coefficients, so when $x \geq 0$, if they converge at x then they converge absolutely. Furthermore, if $\mathbf{P}(x)$ converges at $x \geq 0$ then so do each of $\mathbf{P}(2x), \ldots, \mathbf{P}(mx)$. After noting that one of the summands in the right side of (B.8.1) is

$$\frac{1}{m!}\mathbf{P}(x)^m$$

(this comes from $j = m$), it follows that, for $x \geq 0$, the right side of (B.8.1) converges precisely when $\mathbf{P}(x)$ converges. By Lemma 7.15 there is a Dirichlet series $\mathbf{R}(x)$ such that the right side of (B.8.1) equals $\mathbf{R}(x)$ when $x \geq 0$, so

$$\mathbf{P}^m(x) \;=\; \mathbf{R}(x)$$

on (α, ∞). Since $\mathbf{P}^m(x)$ and $\mathbf{R}(x)$ are both Dirichlet series, they must be the same Dirichlet series. Thus (B.8.1) holds when $x \geq 0$, with the two sides converging precisely when $\mathbf{P}(x)$ converges. □

Here are two examples to illustrate (B.8.1):

$$\begin{aligned} \mathbf{P}^2(x) &= \tfrac{1}{2}\mathbf{P}(x)^2 + \tfrac{1}{2}\mathbf{P}(2x) \\ \mathbf{P}^3(x) &= \tfrac{1}{6}\mathbf{P}(x)^3 + \tfrac{1}{2}\mathbf{P}(x)\cdot\mathbf{P}(2x) + \tfrac{1}{3}\mathbf{P}(3x). \end{aligned}$$

In the additive case one has a polynomial $\mathbf{r}\big(\mathbf{P}(x), \ldots, \mathbf{P}(x^m)\big)$ to express $[m\mathbf{P}](x)$, and now one has, with the same polynomial,

$$\mathbf{P}^m(x) \;=\; \mathbf{r}\big(\mathbf{P}(x), \ldots, \mathbf{P}(mx)\big),$$

since the expression (B.8.1) looks just like the corresponding expression (B.4.1) for the additive case.

APPENDIX C

Consequences of $\delta(\mathsf{P}) = 0$

The major theme in this study of density is to find necessary as well as sufficient conditions that all partition sets have asymptotic density. From Proposition 3.28, the condition $\delta(\mathsf{P}) = 0$ is necessary. Bell [**5**] and Bell, Bender, Cameron and Richmond [**7**] show that this has strong consequences, namely $\rho > 0$ and $\mathbf{A}(\rho) = \infty$.

Note: The Standard Assumption about $\mathcal{A}$ is not in effect in this appendix.

Bell's Positive Radius Theorem: $\delta(\mathsf{P}) = 0$ implies $\rho > 0$

Knowledge of compositions of power series can be useful in showing that a radius of convergence is positive.

Lemma C.1. *Suppose*

$$\mathbf{R}(x) \;=\; \sum_{n=1}^{\infty} r(n)x^n \qquad \textit{and} \qquad \mathbf{T}(x) \;=\; \sum_{n=1}^{\infty} t(n)x^n$$

are two power series such that $t(1) \neq 0$, $\rho_{\mathbf{T}} > 0$ *and* $\rho_{\mathbf{T}\circ\mathbf{R}} > 0$. *Then* $\rho_{\mathbf{R}} > 0$.

PROOF. Since $t(1) \neq 0$, the inverse $\mathbf{T}^{-1}(x)$ of the power series $\mathbf{T}(x)$ exists. Since $\rho_{\mathbf{T}} > 0$, by Proposition A.27 one has $\rho_{\mathbf{T}^{-1}} > 0$. Since $\mathbf{R} = \mathbf{T}^{-1} \circ \mathbf{T} \circ \mathbf{R}$ it follows that $\rho_{\mathbf{R}} > 0$. $\square$

Here is the key tool.

Proposition C.2. *Suppose*

$$\mathbf{S}(x) \;=\; \sum_{n=1}^{\infty} s(n)x^n \qquad \textit{and} \qquad \mathbf{T}(x) \;=\; \sum_{n=1}^{\infty} t(n)x^n$$

are two power series, with $t(1) \neq 0$, $s(n) \geq 0$, *for* $n \geq 1$, *and* $\rho_{\mathbf{T}} > 0$, *such that*

$$2s(n) \;\leq\; [x^n]e^{\mathbf{S}(x)} \;+\; t(n).$$

Then $\rho_{\mathbf{S}} > 0$.

PROOF. Let $\mathbf{R}(x) = \sum_{n=1}^{\infty} r(n)x^n$ be the power series that satisfies the equation

$$2\mathbf{R}(x) \;=\; e^{\mathbf{R}(x)} + \mathbf{T}(x) \;-\; 1, \qquad \text{(C.2.1)}$$

that is, $\mathbf{R}(x)$ is defined by the system of equations

$$\begin{aligned} r(0) &= 0 \\ 2r(n) &= [x^n]e^{\mathbf{R}(x)} + t(n) - 1 \qquad (n \geq 1). \end{aligned}$$

(This system indeed has exactly one solution.) Since

$$r(1) = t(1) \neq 0,$$

it follows that $\mathbf{R}(x)$ has an inverse power series $\mathbf{R}^{-1}(x)$. Now, substituting $x = \mathbf{R}^{-1}(u)$ into the equation (C.2.1) gives

$$\mathbf{T}\big(\mathbf{R}^{-1}(u)\big) = -e^u + 2u + 1.$$

Thus $\rho_{\mathbf{R}^{-1}} > 0$ by Lemma C.1. Then, by Lemma A.27, $\rho_{\mathbf{R}} > 0$.

Now it is proved, by induction on n, that $0 \leq s(n) \leq r(n)$, for $n \geq 1$. For $n = 1$,

$$2s(1) \leq [x]e^{\mathbf{S}(x)} + t(1) = s(1) + t(1),$$

so $s(1) \leq t(1) = r(1)$.

Hence the claim is true when $n = 1$. Suppose it is true for values less than n. Then

$$\begin{aligned} 2s(n) &\leq [x^n]e^{\mathbf{S}(x)} + t(n) \\ &= [x^n]e^{s(1)x+s(2)x^2+\cdots+s(n)x^n} + t(n) \\ &\leq [x^n]e^{r(1)x+r(2)x^2+\cdots+r(n)x^n+(s(n)-r(n))x^n} + t(n), \end{aligned}$$

since $s(k) \leq r(k)$, for $k < n$. Thus

$$\begin{aligned} 2s(n) &\leq [x^n]\Big(e^{r(1)x+\cdots+r(n)x^n} \cdot e^{(s(n)-r(n))x^n}\Big) + t(n) \\ &= [x^n]\Big(e^{\mathbf{R}(x)} \cdot \Big(1 + \big(s(n) - r(n)\big)x^n\Big)\Big) + t(n) \\ &= [x^n]e^{\mathbf{R}(x)} + s(n) - r(n) + t(n) \\ &= \big(2r(n) - t(n)\big) + s(n) - r(n) + t(n) \\ &= r(n) + s(n). \end{aligned}$$

Hence $0 \leq s(n) \leq r(n)$, for $n \geq 1$. Since $\rho_{\mathbf{R}} > 0$, it follows that $\rho_{\mathbf{S}} > 0$. This completes the proof. □

Theorem C.3 (Bell [**5**]). *Let $\mathcal{A}$ be an additive number system such that $\delta(\mathsf{P}) = 0$. Then $\rho > 0$.*

PROOF. Let $\widetilde{\mathcal{A}} = (\widetilde{\mathsf{A}}, \widetilde{\mathsf{P}}, +, 0, \|\ \|)$ be an additive number system with counting functions $\widetilde{a}(n), \widetilde{p}(n)$, where $\widetilde{p}(n) = p(n) + 1$. Then, since $\mathbf{P}(x)$ and $\widetilde{\mathbf{P}}(x)$ have the same radius of convergence, it follows from Corollary 2.25 (b) that $\rho = \widetilde{\rho}$. Also, $\delta(\widetilde{\mathsf{P}}) = 0$ since

$$\frac{\widetilde{p}(n)}{\widetilde{a}(n)} = \frac{p(n)+1}{\widetilde{a}(n)} \leq \frac{p(n)+1}{a(n)} \to 0.$$

So assume, without loss of generality, that $\mathcal{A}$ is such that $p(n) > 0$, for $n \geq 1$ (by replacing $\mathcal{A}$ by $\widetilde{\mathcal{A}}$).

Suppose $\rho = 0$. Since $\delta(\mathsf{P}) = 0$ there exists an integer $N \geq 1$ such that

$$\frac{a(n)}{p(n)} > 3, \qquad \text{for } n > N. \tag{C.3.1}$$

Define the two sets

$$S_1 = \left\{ n > N : \frac{a(n)}{h(n)} \geq 2 \right\} \tag{C.3.2}$$

and

$$S_2 = \left\{ n > N : \frac{a(n)}{h(n)} < 2 \right\}. \tag{C.3.3}$$

From $h(n) = \sum_{ij=n} p(i)/j$, by Möbius inversion one has

$$p(n) = \sum_{j|n} \frac{\mu(j)h(n/j)}{j}.$$

Suppose $n \in S_2$. By (C.3.1) one has $\dfrac{p(n)}{h(n)} < \dfrac{2}{3}$. Thus

$$\sum_{j|n} \frac{\mu(j)h(n/j)}{j} = p(n) < \frac{2h(n)}{3}. \tag{C.3.4}$$

But

$$\begin{aligned} \sum_{j|n} \frac{\mu(j)h(n/j)}{j} &= h(n) + \sum_{\substack{j|n \\ j\neq 1}} \frac{\mu(j)h(n/j)}{j} \\ &\geq h(n) - \sum_{\substack{j|n \\ j\neq 1}} \frac{h(n/j)}{j}. \end{aligned}$$

Combining this result with (C.3.4) gives

$$\begin{aligned} 2h(n)/3 &> \sum_{j|n} \frac{\mu(j)h(n/j)}{j} \\ &\geq h(n) - \sum_{\substack{j|n \\ j\neq 1}} \frac{h(n/j)}{j}, \end{aligned}$$

so,

$$\sum_{\substack{j|n \\ j\neq 1}} \frac{h(n/j)}{j} > \frac{h(n)}{3}.$$

Choose a divisor $j \neq 1$ of n such that $h(n/j)/j$ is maximum among the values $h(n/2)/2, \ldots, h(n/n)/n$. Then

$$d(n)\frac{h(n/j)}{j} \;>\; \frac{h(n)}{3},$$

where $d(n)$ counts the number of divisors of n, so

$$\frac{h(n/j)}{j} \;>\; \frac{h(n)/3}{d(n)} \;\geq\; \frac{h(n)}{3n}.$$

For this choice of j, and since $n \in S_2$,

$$\begin{aligned}
2h(n) \;>\; a(n) \;&=\; [x^n]e^{\mathbf{H}(x)} \\
&\geq\; h(n) \;+\; \frac{h(n/j)^j}{j!} \\
&\geq\; h(n) \;+\; \frac{\big(jh(n)\big)^j}{(3n)^j j!} \\
&\geq\; h(n) \;+\; \frac{h(n)^j}{(3n)^j},
\end{aligned}$$

so

$$h(n) \;<\; (3n)^{j/(j-1)} \;\leq\; 9n^2.$$

Thus, for $n \in S_2$, one has $h(n) < 9n^2$. So, there is a $c > 0$ such that $h(n) < cn^2$, for $n \in S_2 \cup \{1, 2, \ldots, N\}$. Thus $h(n) < cn^2$ if $n \notin S_1$. Let

$$\mathbf{T}(x) \;=\; 2c\sum_{j \notin S_1} j^2 x^j.$$

By the Cauchy-Hadamard formula for the radius of convergence, $\mathbf{T}(x)$ has a radius of convergence equal to 1 if S_2 is infinite, and ∞ if S_2 is finite.

For $n \notin S_1$,

$$\begin{aligned}
[x^n]\big(2\mathbf{H}(x) - \mathbf{T}(x)\big) \;&=\; 2\big(h(n) - cn^2\big) \\
&<\; 0 \\
&\leq\; a(n) \\
&=\; [x^n]e^{\mathbf{H}(x)}.
\end{aligned}$$

For $n \in S_1$,

$$[x^n]\big(2\mathbf{H}(x) - \mathbf{T}(x)\big) \;=\; 2h(n) \;\leq\; a(n) \;=\; [x^n]e^{\mathbf{H}(x)}.$$

Hence, for any $n \geq 1$,

$$2h(n) \;\leq\; [x^n]e^{\mathbf{H}(x)} \;+\; t(n).$$

Now $t(1) = 2c > 0$, so by Proposition C.2, $\rho_{\mathbf{H}} > 0$. Since $\rho = \rho_{\mathbf{H}}$ one has a contradiction. □

Remark C.4. Jason Bell actually proved the stronger result (for reduced $\mathcal{A}$):

$$\rho = 0 \qquad \Rightarrow \qquad \limsup_{n\to\infty} \frac{p(n)}{a(n)} \;=\; 1.$$

This was announced in October, 1999. Richard Warlimont, who had also been looking at this problem (motivated in part by an earlier draft of this book), found a completely different proof in November, and by December he could prove the analog for multiplicative number systems [**55**].

Showing $\mathbf{A}(\rho) = \infty$

Now it is proved that the condition that all partition sets of $\mathcal{A}$ have asymptotic density leads to $\mathbf{A}(\rho) = \infty$. The following technical result is needed.

Lemma C.5 (Stam). *Let $\mathbf{T}(x)$ be a power series with nonnegative coefficients, $t(0) = 0$, and $\rho = \rho_{\mathbf{T}} > 0$. Let*

$$\sum \frac{1}{n!}\mathfrak{q}_n(y)x^n \;=\; e^{y\mathbf{T}(x)}.$$

If $\mathbf{T}(\rho)$ converges, then

$$\limsup_{n\to\infty} \frac{\mathfrak{q}_n(y)}{\mathfrak{q}_n(1)} \;>\; 0, \qquad \textit{for } y > 0.$$

PROOF. See Theorem 3 of Stam [**53**]. □

The following result has stronger hypotheses than needed (see Remark C.9 below), but this version is quite sufficient for our purposes.

Theorem C.6 (Bell, Bender, Cameron and Richmond [**7**]). *Suppose that $\mathcal{A} \in \mathsf{RT}_\rho$ and $\delta(\mathsf{P}) = 0$. Then $\mathbf{A}(\rho) = \infty$.*

PROOF. By Theorem C.3, $\rho > 0$. For the case that $\rho = 1$ the result is trivial. So assume $0 < \rho < 1$. Since $\delta(\mathsf{P}) = 0$ one has $p(n)/a(n) \to 0$. Then from Proposition 3.46 follows

$$\frac{p(n)}{s(n)} \to 0,$$

where

$$\mathbf{S}(x) \;=\; \sum s(n)x^n \;=\; e^{\mathbf{P}(x)},$$

the power series defined by formal exponentiation.

Assume $\mathbf{A}(\rho)$ converges. Then $\mathbf{P}(\rho)$ converges, so apply Lemma C.5, with $\mathbf{T}(x) = 2\mathbf{P}(x)$ and $y = 1/2$, to obtain

$$\limsup_{n\to\infty} \frac{\mathfrak{q}_n(1/2)}{\mathfrak{q}_n(1)} \;>\; 0. \tag{C.6.1}$$

Now

$$\begin{aligned}\sum \frac{1}{n!}\mathbf{q}_n(1/2)x^n &= e^{\mathbf{P}(x)} = \mathbf{S}(x)\\ \sum \frac{1}{n!}\mathbf{q}_n(1)x^n &= e^{2\mathbf{P}(x)} = \mathbf{S}(x)^2\end{aligned}$$

so

$$\begin{aligned}\frac{1}{n!}\mathbf{q}_n(1/2) &= s(n)\\ \frac{1}{n!}\mathbf{q}_n(1) &= \sum_{k=0}^{n} s(k)s(n-k).\end{aligned}$$

Then, from $\mathbf{S}'(x) = \mathbf{P}'(x) \cdot \mathbf{S}(x)$ follows

$$\begin{aligned}\frac{1}{n!}\mathbf{q}_n(1/2) &= s(n)\\ &= \frac{1}{n}\sum_{k=0}^{n} kp(k)s(n-k)\\ &= \sum_{k=0}^{n} \frac{kp(k)}{ns(k)} s(k)s(n-k)\\ &< \sum_{k<n^{1/2}} n^{-1/2}s(k)s(n-k)\\ &\quad + \sum_{k\geq n^{1/2}} \frac{p(k)}{s(k)} s(k)s(n-k)\\ &\leq \mathrm{o}(1)\sum_{k=0}^{n} s(k)s(n-k)\\ &= \mathrm{o}(1)\,\frac{1}{n!}\mathbf{q}_n(1),\end{aligned}$$

contradicting (C.6.1). Thus $\mathbf{A}(\rho)$ diverges. □

Corollary C.7. *Suppose $\mathcal{A}$ is such that all partition sets have [global] asymptotic density. Then the [global] asymptotic density of a partition set B is the Dirichlet density $\partial(\mathsf{B})$.*

PROOF. It suffices to prove this when $\mathcal{A}$ is reduced. The hypothesis forces $\delta(\mathsf{P}) = 0$ by Proposition 3.28 (the proof of this proposition does not require that one assume $\rho > 0$). Then $\rho > 0$ follows from Theorem C.3, so $\mathcal{A} \in \mathsf{RT}_\rho$ by Corollary 3.30. Thus $\mathbf{A}(\rho) = \infty$ by Theorem C.6. This means Theorem 3.13 applies. □

Corollary C.8. *Suppose that $\mathcal{A} \in \mathsf{RT}_\rho$ and $\delta(\mathsf{P}) = 0$. Then either P is a finite set or $\mathbf{P}(\rho) = \infty$.*

PROOF. By Corollary 2.25,

$$\mathbf{A}(x) < \infty \quad \text{iff} \quad \mathbf{P}(x) < \infty,$$

for $0 \leq x < 1$. If $\rho = 1$ then $\mathbf{P}(\rho) < \infty$ holds iff P is a finite set. If $\rho < 1$ apply Theorem C.6. $\square$

Remark C.9. Actually Bell, et al [**7**] prove

$$\delta(\mathsf{P}) = 0 \qquad \Rightarrow \qquad \mathbf{A}(\rho) = \infty,$$

without assuming $\mathcal{A}$ has any special properties other than being an additive number system.

APPENDIX D

On the Monotonicity of $a(n)$ When $p(n) \leq 1$

When $\mathcal{A} \in \mathsf{RT}_\rho$ with $\rho < 1$ then it is easy to prove that $a(n)$ is eventually strictly increasing. One only has to apply Schur's Tauberian Theorem to $(1 - x) \cdot \mathbf{A}(x)$ to see that

$$a(n+1) - a(n) \ \sim\ (1-\rho)a(n).$$

When $\rho = 1$ it need not be the case that $a(n)$ is eventually strictly increasing. In 1956 Bateman and Erdös [**4**] investigated this question under the hypothesis $p(n) \leq 1$, that is, there is at most one indecomposable of each size. They found necessary and sufficient conditions on the support of $p(n)$ for $a(n)$ to be monotone. Recall the definition of $f^-(n)$ from Chapter 4.

Definition D.1. For $f : \mathbb{N} \to \mathbb{N}$ let

$$\begin{aligned} f^-(n) &= f(n) - f(n-1) \qquad \text{for } n > 0 \\ f^-(0) &= f(0). \end{aligned}$$

f^{--} means $\big(f^-\big)^-$.

Definition D.2. An additive number system $\mathcal{A}$ with counting functions $a(n), p(n)$ and with $r = \sum p(n)$ has the property BE[1] if:

(a) $r \geq 2$, and
(b) $\gcd\big(\{\|\mathsf{q}\| : \mathsf{q} \in \mathsf{P},\ \mathsf{q} \neq \mathsf{p}\}\big) \ =\ 1$, for any $\mathsf{p} \in \mathsf{P}$.

Thus $\mathcal{A}$ has BE means that if one deletes any indecomposable from P then the subsystem generated by the remaining indecomposables is (still) reduced.

First consider the case of finitely many indecomposables.

Theorem D.3. *Suppose $\mathcal{A}$ is an additive number system of finite rank r with property* BE. *Let* $\operatorname{supp} p(n) = \{d_1, \ldots, d_k\}$ *and let* $p(d_i) = m_i$. *Then*

$$a^-(n) \ \sim\ \frac{1}{(r-2)!\prod {d_i}^{m_i}} \cdot n^{r-2}.$$

Consequently $a^-(n)$ is eventually positive and

$$\lim_{n\to\infty} \frac{a^{--}(n)}{a^-(n)} \ =\ 0.$$

[1] Bateman and Erdös referred to this property as P_1—here the initials BE of Bateman and Erdös offer an appropriately mnemonic name.

PROOF. By Lemma 4.11

$$\begin{aligned}\sum_{n\geq 0} a^-(n)x^n &= (1-x)\cdot\sum_{n\geq 0} a(n)x^n \\ &= (1-x)\cdot\prod_{i=1}^{k}(1-x^{d_i})^{-m_i} \\ &= \frac{C}{(1-x)^{r-1}} + \sum_{\substack{j<r-1\\ \theta^\ell=1}} \frac{C_{j,\theta}}{(1-\theta x)^j},\end{aligned}$$

using the partial fraction decomposition technique from the proof of Theorem 2.48. The reason for the bound $r-1$ on the j's is that property BE guarantees that a given factor $1-\theta x$, $\theta\neq 1$, can be a factor in at most $r-2$ of the $1-x^{d_i}$ appearing in $\prod_{i=1}^{k}(1-x^{d_i})^{-m_i}$.

Now repeat the proof of Theorem 2.48 to derive the asymptotic expression for $a^-(n)$. The asymptotics show that $a^-(n)$ is eventually positive and that $\dfrac{a^-(n-1)}{a^-(n)}\to 1$ as n goes to infinity. Thus $\lim_{n\to\infty}\dfrac{a^{--}(n)}{a^-(n)} = 0$. □

Corollary D.4. *Suppose $\mathcal{A}$ is an additive number system of finite rank r with counting functions $a(n), p(n)$. Then $a(n)$ is eventually strictly increasing iff* BE *holds.*

PROOF. If BE holds then appeal to Theorem D.3.

For the converse, suppose that $a(n)$ is eventually strictly increasing. Let $\mathsf{supp}\,p(n) = \{d_1,\ldots,d_k\}$ and let $p(d_i) = m_i$. In the following δ_{ij} is the Kronecker delta function, namely $\delta_{ij}=1$ if $i=j$, otherwise $\delta_{ij}=0$. Then, for $j\in\{d_1,\ldots,d_k\}$,

$$\begin{aligned}(1-x^j)\cdot\sum_{n\geq 0} a^-(n)x^n &= (1-x^j)\cdot(1-x)\cdot\prod_{i=1}^{k}(1-x^{d_i})^{-m_i} \\ &= (1-x)\cdot\prod_{i=1}^{k}(1-x^{d_i})^{-m_i+\delta_{ij}}\end{aligned}$$

so

$$(1+x+x^2+\cdots+x^{j-1})\cdot\sum_{n\geq 0} a^-(n)x^n = \prod_{i=1}^{k}(1-x^{d_i})^{-m_i+\delta_{ij}}.$$

Since $a^-(n)$ is eventually positive it follows that in the expansion of the left side of the last equation the coefficients are eventually positive. Since $p(j)\geq 1$ this equation is the partition identity of a finitely generated additive number system, the subsystem of $\mathcal{A}$ generated by P with one indecomposable of norm j removed. It follows from Lemma 2.42 that $\gcd\big(\{d_i : m_i-\delta_{ij} > 0\}\big) = 1$. Thus $\mathcal{A}$ has the property BE. □

Theorem D.5. *Suppose $\mathcal{A}$ is an additive number system with counting functions $a(n), p(n)$ and $p(n) \leq 1$. Then $a(n)$ is eventually strictly increasing iff* BE *holds.*

PROOF. Suppose that $a(n)$ is eventually strictly increasing. The argument to show that BE holds is a repeat of the argument from the last proof, but the product may be infinite. If $p(j) = 1$ then

$$(1 - x^j) \cdot \sum_{n \geq 0} a^-(n) x^n \;=\; (1 - x) \cdot \prod_{n \geq 1} (1 - x^n)^{-p(n) + \delta_{nj}},$$

so

$$(1 + x + x^2 + \cdots + x^{j-1}) \cdot \sum_{n \geq 0} a^-(n) x^n \;=\; \prod_{n \geq 1} (1 - x^n)^{-p(n) + \delta_{nj}}.$$

Since $a^-(n)$ is eventually positive it follows that in the expansion of the left side the coefficients are eventually positive. Since this equation is the partition identity of an additive number system it follows by Lemma 2.42 that $\gcd\big(\{n : p(n) - \delta_{nj} > 0\}\big) = 1$. Thus $\mathcal{A}$ has the property BE.

Suppose that BE holds. The case that $r = \sum_n p(n)$ is finite is covered by Corollary D.4. So now suppose r is infinite.

Choose $p_1(n)$ with finite support such that $0 \leq p_1(n) \leq p(n)$ and $p_1(n)$ has the property BE. Let $p(n) = p_1(n) + p_2(n)$. Let $a_1(n), a_2(n)$ be defined by the additive partition identities:

$$\sum_{n \geq 0} a_i(n) x^n \;=\; \prod_{n \geq 1} (1 - x^n)^{-p_i(n)}, \qquad \text{for } \; i = 1, 2.$$

Then, from Theorem D.3,

$$\frac{a_1^-(n)}{1 + \big|a_1^{--}(n)\big|} \;\to\; \infty. \tag{D.5.1}$$

Let t be a positive integer. By (D.5.1) there exists a positive integer M such that

$$\frac{a_1^-(n)}{1 + |a_1^{--}(n)|} \;\geq\; \begin{cases} -M & \text{for all } n \\ t + 1 & \text{for } n \geq M. \end{cases} \tag{D.5.2}$$

Let

$$c \;=\; \max_{0 \leq n \leq M-1} \Big(1 + \big|a_1^{--}(n)\big|\Big). \tag{D.5.3}$$

From Lemma 4.9 $\dfrac{a_2(n)}{A_2(n)} \to 0$, as $\dfrac{a_2^\star(n-1)}{a_2^\star(n)} \to 1$ implies $\dfrac{A_2(n-1)}{A_2(n)} \to 1$. Thus there is a positive integer N such that

$$\frac{\displaystyle\sum_{m=n-M+1}^{n} a_2(m)}{A_2(n)} \;\leq\; \frac{1}{(M + t + 1)c}, \qquad \text{for } n \geq N. \tag{D.5.4}$$

For $n \geq \max(M, N)$,

$$
\begin{aligned}
a^-(n) &= \sum_{m=0}^{n} a_1^-(n-m) \cdot a_2(m) \qquad \text{by Lemma 4.12} \\
&\geq \sum_{m=0}^{n-M} a_1^-(n-m) \cdot a_2(m) \; - \sum_{m=n-M+1}^{n} \left|a_1^-(n-m)\right| \cdot a_2(m) \\
&\geq (t+1) \cdot \sum_{m=0}^{n-M} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \\
&\qquad - M \cdot \sum_{m=n-M+1}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \quad \text{by (D.5.2)} \\
&= (t+1) \cdot \sum_{m=0}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \\
&\qquad - (M+t+1) \cdot \sum_{m=n-M+1}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \\
&\geq (t+1) \cdot \sum_{m=0}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \\
&\qquad - (M+t+1) \cdot \sum_{m=n-M+1}^{n} c a_2(m) \qquad \text{by (D.5.3)} \\
&= (t+1) \cdot \sum_{m=0}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \\
&\qquad - (M+t+1) \cdot c \cdot \frac{\sum_{m=n-M+1}^{n} a_2(m)}{A_2(n)} \cdot A_2(n) \\
&\geq \left[(t+1) \cdot \sum_{m=0}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m)\right] - A_2(n) \quad \text{(D.5.4)} \\
&\geq t \cdot \sum_{m=0}^{n} \left(1 + \left|a_1^{--}(n-m)\right|\right) \cdot a_2(m) \\
&\geq t \cdot \sum_{m=0}^{n} a_2(m) \\
&= t \cdot A_2(n) \, .
\end{aligned}
$$

Thus, for $n \geq \max(M, N)$, $\dfrac{a^-(n)}{A_2(n)} \geq t$, and thus $a(n) - a(n-1)$ is eventually positive, so $a(n)$ is eventually strictly increasing. $\square$

APPENDIX E

Results of Woods

Woods was interested in modifying the work of Compton so that it would apply to additive number systems that are derived from (partial or total) unary functions with k additional unary predicates. The following Tauberian theorem of Woods [**57**] is a modification of a result of Meir and Moon [**37**].

Woods' Tauberian Theorem

Theorem E.1 (Woods, 1997). *Let* $\mathbf{S}(x), \mathbf{T}(x)$ *be two power series such that, for some* ρ *with* $0 < \rho \le 1$, $\mathbf{T}(x) \in \mathsf{RT}_\rho$. *Let* $\mathbf{R}(x) = \mathbf{S}(x) * \mathbf{T}(x)$. *If, for some* $C > 0$ *and* $\mu < 1$,

$$s(n) = \mathrm{O}\big(\rho^{-n}/n\big) \tag{E.1.1}$$

$$t(n) \sim C\rho^{-n}/n^{\mu} \tag{E.1.2}$$

and $\mathbf{S}(\rho)$ *converges absolutely then* $\displaystyle\lim_{n\to\infty}\frac{r(n)}{t(n)}$ *exists and*

$$\lim_{n\to\infty}\frac{r(n)}{t(n)} = \mathbf{S}(\rho) = \lim_{x\to\rho}\frac{\mathbf{R}(x)}{\mathbf{T}(x)}.$$

Proof. The case that $\mu \le 0$ is covered by Theorem 5.1, in view of Lemma 5.2, with $t_0(n) = Cn^{-\mu}$. So, for the rest of the proof, assume

$$0 < \mu < 1.$$

Assume (by the change of variable $x \leftarrow \rho x$) that $\rho = 1$. Then

$$\Big|\mathbf{S}(1) - \frac{r(n)}{t(n)}\Big| \le \underbrace{\Big|\mathbf{S}(1) - \sum_{k\le n} s(k)\Big|}_{\mathrm{I}} + \underbrace{\Big|\sum_{k\le n} s(k) - \frac{r(n)}{t(n)}\Big|}_{\mathrm{II}},$$

and term I goes to 0 as $n \to \infty$. So one only needs to show that term II vanishes as $n \to \infty$. For this,

$$\Big|\sum_{k\le n} s(k) - \frac{r(n)}{t(n)}\Big| \;\le\; \underbrace{\sum_{k\le\sqrt{n}} \Big|s(k)\cdot\Big(1-\frac{t(n-k)}{t(n)}\Big)\Big|}_{\mathrm{III}}$$
$$+ \underbrace{\sum_{\sqrt{n}<k\le n} |s(k)|}_{\mathrm{IV}} \;+\; \underbrace{\sum_{\sqrt{n}<k\le n} |s(k)|\frac{t(n-k)}{t(n)}}_{\mathrm{V}}.$$

Term IV goes to 0 as $n \to \infty$ since $\sum_k |s(k)|$ is convergent. Thus it only remains to show that terms III and V also go to 0.

For term III, let k_n be a value of k in $[0, \sqrt{n}]$ for which $\Big|1 - \dfrac{t(n-k)}{t(n)}\Big|$ achieves its maximum value. Then

$$\begin{aligned}\sum_{k\le\sqrt{n}} \Big|s(k)\cdot\Big(1-\frac{t(n-k)}{t(n)}\Big)\Big| \;&\le\; \sum_{k\le\sqrt{n}} \Big|s(k)\cdot\Big(1-\frac{t(n-k_n)}{t(n)}\Big)\Big|\\ &\le\; \Big(\sum_{k\ge 0}|s(k)|\Big)\cdot\Big|\Big(1-\frac{t(n-k_n)}{t(n)}\Big)\Big|.\end{aligned}$$

Since $k_n \le \sqrt{n}$ it follows that $n - k_n \to \infty$ as $n \to \infty$. Since $t(n) \sim Cn^{-\mu}$ it follows that

$$t(n-k_n) \;\sim\; C\cdot(n-k_n)^{-\mu}.$$

Now, from $\lim\limits_{n\to\infty} k_n/n = 0$,

$$(n-k_n)^{-\mu} \;\sim\; n^{-\mu},$$

and thus

$$t(n-k_n) \;\sim\; Cn^{-\mu} \;\sim\; t(n).$$

From this, for any $\varepsilon > 0$, for n sufficiently large,

$$\Big|1 - \frac{t(n-k_n)}{t(n)}\Big| \;<\; \varepsilon.$$

This shows that term III goes to 0 as $n \to \infty$.

So now one only needs to show that term V has limit 0 as $n \to \infty$. Let

$$C_n \;=\; \sum_{k>\sqrt{n}} |s(k)|,$$

the tail end of the convergent series $\sum_k |s(k)|$. Next define

$$M_n \;=\; n\sqrt{C_n} \;+\; \sqrt{n}.$$

Then the following hold:

(E.1.3) $$\lim_{n\to\infty} C_n \;=\; 0$$

$$
\begin{array}{lrcll}
\text{(E.1.4)} & \lim_{n\to\infty} M_n & = & \infty & \\
\text{(E.1.5)} & \lim_{n\to\infty} (n - M_n) & = & \infty & \\
 & \sqrt{n} & < & n - M_n & \text{for } n \text{ sufficiently large} \\
\text{(E.1.6)} & M_n & = & \mathrm{o}(n) & \\
\text{(E.1.7)} & nC_n & = & \mathrm{o}(M_n). &
\end{array}
$$

Break up term V into two sums as follows:

$$
\sum_{\sqrt{n}<k<n} |s(k)|\frac{t(n-k)}{t(n)} = \underbrace{\sum_{\sqrt{n}<k\le n-M_n} |s(k)|\frac{t(n-k)}{t(n)}}_{\mathrm{VI}} + \underbrace{\sum_{n-M_n<k\le n} |s(k)|\frac{t(n-k)}{t(n)}}_{\mathrm{VII}}.
$$

It will be proved that each of the terms VI and VII goes to 0 as $n \to \infty$.

Define k_n to be a k in $(\sqrt{n}, n - M_n]$ for which $\dfrac{t(n-k)}{t(n)}$ has maximum value. Then

$$
\begin{array}{lrcl}
 & \displaystyle\sum_{\sqrt{n}<k\le n-M_n} |s(k)|\frac{t(n-k)}{t(n)} & \le & \displaystyle\frac{t(n-k_n)}{t(n)} \cdot \sum_{\sqrt{n}<k\le n-M_n} |s(k)| \\
\text{(E.1.8)} & & \le & C_n \displaystyle\frac{t(n-k_n)}{t(n)}.
\end{array}
$$

Clearly $n - k_n \ge M_n$ and (E.1.4) imply

$$
\text{(E.1.9)} \qquad \lim_{n\to\infty} (n - k_n) = \infty.
$$

From (E.1.2), (E.1.4), and (E.1.9),

$$
\begin{array}{lrcl}
 & t(n-k_n) & \sim & C \cdot (n-k_n)^{-\mu} \\
\text{(E.1.10)} & t(M_n) & \sim & CM_n^{-\mu}
\end{array}
$$

and thus

$$
\text{(E.1.11)} \qquad \frac{t(n-k_n)}{t(M_n)} \sim \Big(\frac{M_n}{n-k_n}\Big)^{\mu}.
$$

From $k_n \le n - M_n$, one has $\dfrac{M_n}{n-k_n} \le 1$, so $\Big(\dfrac{M_n}{n-k_n}\Big)^{\mu} \le 1$. Combining this with (E.1.11) leads to

$$
\text{(E.1.12)} \qquad t(n-k_n) = \mathrm{O}\big(t(M_n)\big),
$$

and applying (E.1.12) to (E.1.8) gives

$$\begin{aligned}\sum_{\sqrt{n}<k\le n-M_n} |s(k)|\frac{t(n-k_n)}{t(n)} &= \mathrm{O}\Big(C_n\frac{t(M_n)}{t(n)}\Big)\\ &= \mathrm{O}\Big(C_n\cdot\big(n/M_n\big)^{\mu}\Big) && \text{(E.1.2), (E.1.10)}\\ &= \mathrm{O}\Big(\big(C_n\cdot n/M_n\big)^{\mu}\Big) && \text{(E.1.3), } 0<\mu<1,\end{aligned}$$

and the last line goes to 0 as $n\to\infty$ by (E.1.7). Thus term VI goes to 0 as $n\to\infty$.

For term VII,

$$\sum_{n-M_n<k\le n} |s(k)|\frac{t(n-k)}{t(n)} \;=\; \sum_{k<M_n} |s(n-k)|\frac{t(k)}{t(n)}.$$

Now choose k_n to be a $k<M_n$ such that $|s(n-k)|$ is maximal. Then

$$\sum_{k<M_n} |s(n-k)|\frac{t(k)}{t(n)} \;\le\; |s(n-k_n)|\sum_{k<M_n}\frac{t(k)}{t(n)}.$$

From (E.1.1) and (E.1.5),

$$s(n-k_n) \;=\; \mathrm{O}\Big(\frac{1}{n-k_n}\Big),$$

and, since $k_n=\mathrm{o}(n)$ by (E.1.6),

$$\frac{1}{n-k_n} \;\sim\; \frac{1}{n}.$$

Thus

$$s(n-k_n) \;=\; \mathrm{O}\big(1/n\big). \tag{E.1.13}$$

For $0<\mu<1$, one has an upper bound on $k^{-\mu}$ given by

$$\frac{1}{k^{\mu}} \;<\; \int_{k-1}^{k}\frac{dx}{x^{\mu}}, \qquad \text{for } x>0,$$

and this leads to

$$\begin{aligned}\sum_{1\le k<n}\frac{1}{k^{\mu}} &< 1 + \int_1^n\frac{dx}{x^{\mu}}\\ &= 1+\Big(\frac{n^{1-\mu}}{1-\mu}-\frac{1}{1-\mu}\Big)\\ &= \frac{n^{1-\mu}}{1-\mu} - \frac{\mu}{1-\mu},\end{aligned}$$

and thus

$$\sum_{1\le k<n} \frac{1}{k^\mu} \;=\; \mathrm{O}\big(n^{1-\mu}\big). \tag{E.1.14}$$

Now, from $t(k) \sim k^{-\mu}$,

$$t(k) \;=\; \mathrm{O}\big(k^{-\mu}\big) \tag{E.1.15}$$

$$\frac{1}{t(k)} \;=\; \mathrm{O}\big(k^{\mu}\big). \tag{E.1.16}$$

Then

$$\begin{aligned} |s(n-k_n)| \sum_{k<M_n} \frac{t(k)}{t(n)} \;&=\; \mathrm{O}\Big(\frac{1}{nt(n)} \sum_{k<M_n} t(k)\Big) && \text{by (E.1.13)}\\ &=\; \mathrm{O}\Big(\frac{1}{n^{1-\mu}} \sum_{1\le k<M_n} k^{-\mu}\Big) && \text{by (E.1.15), (E.1.16)}\\ &=\; \mathrm{O}\Big(\big(M_n/n\big)^{1-\mu}\Big) && \text{by (E.1.14).} \end{aligned}$$

From (E.1.6),

$$\lim_{n\to\infty} \big(M_n/n\big)^{1-\mu} \;=\; 0,$$

and thus term VII goes to 0 as $n \to \infty$. This finishes the proof. □

Woods' Density Theorem

The following theorem from [**57**] is used to prove monadic second-order limit laws for classes of unary functions with unary predicates.

Theorem E.2 (Woods, 1997). *Suppose $\mathcal{A}$ is an additive number system with*

$$p(n) \;=\; \mathrm{O}\big(\beta^n/n\big) \tag{E.2.1}$$

$$a(n) \;\sim\; C\beta^n/n^\mu, \tag{E.2.2}$$

for some $C > 0$, $\beta \ge 1$, and $\mu < 1$. Then all partition sets of $\mathcal{A}$ have asymptotic density which equals the Dirichlet density.

Proof. From $a(n) \sim C\beta^n/n^\mu$ one has $\rho = 1/\beta$, and $\mathcal{A} \in \mathsf{RT}_\rho$. If $\beta = 1$ then Theorem 4.2 applies. So assume $\beta > 1$. Let B be a partition set.

Case 1: $\partial(\mathsf{B}) \;=\; 0$.
From Corollary 5.7 one has $\delta(\mathsf{B}) = 0$.

Case 2: $\partial(\mathsf{B}) \;>\; 0$.
For this case it will be proved that Theorem E.1 applies with

$$\mathbf{S}(x) = [\overline{\mathbf{B}}/\mathbf{A}](x) \quad \text{and} \quad \mathbf{T}(x) = \mathbf{A}(x).$$

Note that $\mathbf{S}(x)$ is absolutely convergent on $[0,\rho]$ by Lemma 5.11.

Let $\mathsf{B} = \sum_{i=1}^{k} \gamma_i \mathsf{P}_i$, let J be the set of indices j in $\{1, \ldots, k\}$ such that $\mathbf{A}_{\mathsf{P}_j}(\rho) < \infty$, and let $\mathsf{Q} = \bigcup_{j \in J} \mathsf{P}_j$. Then, for $0 \leq x < \rho$, since $\partial(\overline{\mathsf{B}}) > 0$,

$$\begin{aligned}
[\overline{\mathbf{B}}/\mathbf{A}](x) &= \overline{\mathbf{B}}(x)/\mathbf{A}(x) \\
&= \prod_{j \in J} [\gamma_j \mathbf{P}_j](x)/\mathbf{A}_{\mathbf{P}_j}(x) \\
&= \Big(\prod_{j \in J} [\gamma_j \mathbf{P}_j](x)\Big)/\mathbf{A}_{\mathsf{Q}}(x) \\
&= \Big(\prod_{j \in J} [\gamma_j \mathbf{P}_j](x)\Big) \cdot \exp\Big(-\sum_{m \geq 1} \mathbf{Q}(x^m)/m\Big),
\end{aligned}$$

so it follows that

$$\begin{aligned}
|s(n)| &= \left| \big[x^n\big] \Big(\Big(\prod_{j \in J} [\gamma_j \mathbf{P}_j](x)\Big) \cdot \exp\Big(-\sum_{m \geq 1} \mathbf{Q}(x^m)/m\Big)\Big) \right| \\
&\leq \big[x^n\big] \Big(\Big(\prod_{j \in J} [\gamma_j \mathbf{P}_j](x)\Big) \cdot \exp\Big(\sum_{m \geq 1} \mathbf{Q}(x^m)/m\Big)\Big) \\
&= \big[x^n\big] \Big(\Big(\prod_{j \in J} [\gamma_j \mathbf{P}_j](x)\Big) \cdot \mathbf{A}_{\mathsf{Q}}(x)\Big) \\
&\leq \big[x^n\big] \big(\mathbf{A}_{\mathsf{Q}}(x) \cdot \mathbf{A}_{\mathsf{Q}}(x)\big) \\
&= \big[x^n\big] \exp\Big(2 \sum_{m \geq 1} \mathbf{Q}(x^m)/m\Big).
\end{aligned}$$

Let $\mathbf{U}(x) = \sum_{n \geq 0} u(n) x^n$ be the power series expansion of $2 \sum_{m \geq 1} \mathbf{Q}(x^m)/m$. Then $u(n) = \mathrm{O}\big(\rho^{-n}/n\big)$ follows from:

$$\begin{aligned}
u(n) &\leq 2 \sum_{k|n} p(n/k)/k \\
&= 2 \sum_{k|n} (k/n) p(k) \\
&= 2p(n) \ + \ 2 \sum_{\substack{k|n \\ 1 \leq k \leq n/2}} (k/n) p(k) \\
&= \mathrm{O}\Big(\rho^{-n}/n \ + \sum_{\substack{k|n \\ 1 \leq k \leq n/2}} (k/n) \rho^{-k}/k\Big) \\
&= \mathrm{O}\big(\rho^{-n}/n\big).
\end{aligned}$$

Now let $\mathbf{V}(x) = \sum_{n\geq 0} v(n)x^n$ be the power series expansion of $\exp\big(\mathbf{U}(x)\big)$. Since

$$\mathbf{A}_{\mathsf{Q}}(\rho) \;=\; \prod_{j\in J} \mathbf{A}_{\mathsf{P}_j}(\rho) \;<\; \infty,$$

$\mathbf{U}(\rho) < \infty$, so $\mathbf{V}(x) = \exp\big(\mathbf{U}(x)\big)$, convergent at ρ. The claim is that $v(n) \in \mathrm{O}(\rho^{-n}/n)$. Differentiating $\mathbf{V}(x) = \exp\big(\mathbf{U}(x)\big)$ gives

$$\begin{aligned} \mathbf{V}'(x) &= \mathbf{U}'(x)\cdot \exp\big(\mathbf{U}(x)\big) \\ &= \mathbf{U}'(x)\cdot \mathbf{V}(x). \end{aligned}$$

By equating the coefficients of x^{n-1},

$$\begin{aligned} nv(n) &= \sum_{k=1}^{n} ku(k)v(n-k) \\ &= \mathrm{O}\Big(\sum_{k=1}^{n} k\big(\rho^{-k}/k\big)v(n-k)\Big) \\ &= \mathrm{O}\Big(\sum_{k=1}^{n} \rho^{-k}v(n-k)\Big) \\ &= \mathrm{O}\Big(\rho^{-n}\cdot\sum_{k=1}^{n} \rho^{n-k}v(n-k)\Big) \\ &= \mathrm{O}\Big(\rho^{-n}\cdot\sum_{k=0}^{n-1} v(k)\rho^{k}\Big), \end{aligned}$$

so $v(n) = \mathrm{O}\big(\rho^{-n}/n\big)$ since

$$\sum_{k=0}^{n-1} v(k)\rho^k \;\leq\; \mathbf{V}(\rho) \;=\; \exp\big(\mathbf{U}(\rho)\big) \;<\; \infty.$$

Now, since $|s(n)| \leq v(n)$, it follows that $s(n) = \mathrm{O}\big(\rho^{-n}/n\big)$. Thus $\mathbf{S}(x)$ and $\mathbf{T}(x)$ fulfill the hypotheses of Theorem E.1. Since $\overline{\mathbf{B}}(x) = \mathbf{S}(x) * \mathbf{A}(x)$ it follows that $\overline{\mathsf{B}}$ has asymptotic density $\mathbf{S}(\rho)$, which is the Dirichlet density of $\overline{\mathsf{B}}$. Thus the same applies to B. $\square$

Bibliography

[1] Tom M. Apostol, *Introduction to Analytical Number Theory.* Springer Verlag, 1976.

[2] Tom M. Apostol, *Modular Functions and Dirichlet Series.* Springer Verlag, 2nd ed., 1990.

[3] Paul T. Bateman and Harold G. Diamond, *Asymptotic distribution of Beurling's generalized prime numbers.* In: W. LeVeque, ed., *Studies in Number Theory*, Math. Assoc. Amer., Studies in Math. **6**, Prentice-Hall, 1969, 152–210.

[4] P.T. Bateman and P. Erdös, *Monotonicity of partition functions.* Mathematika **3** (1956), 1–14.

[5] Jason P. Bell, *When structures are almost surely connected.* Electron. J. Combin. **7** (2000), R36.

[6] Jason P. Bell, *Sufficient conditions for zero-one laws.* Preprint, 2000.

[7] Jason P. Bell, Edward A. Bender, Peter J. Cameron, L. Bruce Richmond, *Asymptotics for the probability of connectedness and the distribution of number of components.* Electron. J. Combin. **7** (2000), R33.

[8] Jason P. Bell and Stanley N. Burris, *Asymptotics for logical limit laws. When the growth of the components is in an RT class.* Preprint, 2000.

[9] Edward A. Bender, *Asymptotic methods in enumeration.* SIAM Rev. **16** (1974), 485–515.

[10] Arne Beurling, *Analyse de la loi asymptotique de la distribution des nombres premiers généralisés, I.* Acta Math. **68** (1937), 255–291.

[11] David Bigelow and Stanley Burris, *Boolean algebras of factor congruences.* Acta Sci. Math. (Szeged) **54** (1990), 11–20.

[12] Stanley Burris, *Spectrally determined first-order limit laws.* Logic and Random Structures (New Brunswick, NJ, 1995), 33–52, ed. by Ravi Boppana and James Lynch. DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **33**, Amer. Math. Soc., Providence, RI, 1997.

[13] Stanley N. Burris, *Logic for Mathematics and Computer Science.* Prentice Hall, 1998.

[14] Stanley Burris, Kevin Compton, Andrew Odlyzko, Bruce Richmond, *Fine spectra and limit laws II. First-order 0–1 laws.* Canad. J. Math. **49** (1997), 641–652.

[15] Stanley Burris and Paweł Idziak, *A directly representable variety has a discrete first-order law.* Internat. J. Algebra Comput. **6** (1996), 269–276.

[16] Stanley Burris and H.P. Sankappanavar, *A Course in Universal Algebra.* Originally published by Springer, 1981. Now freely distributed as the *Millenium Edition* on the internet at `www.thoralf.uwaterloo.ca`.

[17] Stanley Burris and András Sárközy, *Fine spectra and limit laws I. First-order laws.* Canad. J. Math. **49** (1997), 468–498.

[18] Peter J. Cameron, *On the probability of connectedness.* 15th British Combinatorics Conference (Stirling 1995), Discrete Math. **167/168** (1997), 175–187.

[19] C.C. Chang and H.J. Keisler, *Model Theory.* Studies in Logic and Foundations of Mathematics **73**, North-Holland, 1973.

[20] Kevin J. Compton, *A logical approach to asymptotic combinatorics. I. First order properties.* Adv. in Math. **65** (1987), 65–96.

[21] Kevin J. Compton, *A logical approach to asymptotic combinatorics. II. Monadic second-order properties.* J. Combin. Theory, Ser. A **50** (1989), 110–131.

[22] Kevin J. Compton, *Laws in logic and combinatorics.* Algorithms and Order (Ottawa, ON, 1987), 353–383, ed. by I. Rival. Kluwer Acad. Publ., 1989.

[23] Heinz-Dieter Ebbinghaus and Jörg Flum, *Finite Model Theory.* Perspectives in Mathematical Logic, 2nd ed., Springer-Verlag, 1999.

[24] P. Erdös and G. Szekeres, *Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisches Problem.* Acta Litt. Sci. Reg. Univ. Hungar. Fr.–Jos. Sect. Sci. Math. **7** (1934), 94–103.

[25] J.L. Geluk and L. de Haan, *Regular variation, extensions and Tauberian theorems.* CWI Tract **40**, Centre for Mathematics and Computer Science, 1987.

[26] G.H. Hardy and E.M. Wright, *An Introduction to the Theory of Numbers.* Oxford Press, 5th ed., 1980.

[27] Einar Hille, *Analytic Function Theory, Vol. I.* Introduction to Higher Mathematics, Ginn and Company, 1959.

[28] Paweł M. Idziak and Jerzy Tyszkiewicz, *Monadic second order probabilities in algebra. Directly representable varieties and groups.* Logic and random structures (New Brunswick, NJ, 1995), 79–107, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., **33**. Amer. Math. Soc., Providence, RI, 1997.

[29] John Knopfmacher, *Abstract Analytic Number Theory.* North-Holland Mathematical Library **12**, North-Holland, 1975. (Available as a Dover Reprint.)

[30] John Knopfmacher, *Analytic Arithmetic of Algebraic Function Fields.* Lecture Notes in Pure and Applied Mathematics **50**, Marcel Dekker, Inc. 1979.

[31] Arnold Knopfmacher, John Knopfmacher and Richard Warlimont, *"Factorisatio numerorum" in arithmetical semigroups.* Acta Arith. **61** (1992), 327–336.

[32] Arnold Knopfmacher and John Knopfmacher, *Arithmetical semigroups related to trees and polyhedra.* J. Combin. Th., Ser. A. **86** (1999), 85–102.

[33] John Knopfmacher and Richard Warlimont, *Arithmetical semigroups related to trees and polyhedra, II — Maps on surfaces.* Preprint.

[34] Konrad Knopp, *Theory and Application of Infinite Series.* Dover edition, 1990.

[35] Ralph N. McKenzie, George F. McNulty and Walter F. Taylor, *Algebras, Lattices, and Varieties. I.* The Wadsworth and Brooks/Cole Mathematics Series, Wadsworth and Brooks/Cole, 1987.

[36] Günter Meinardus, *Asymptotische Aussagen über Partitionen.* Math. Zeitschr. Bd. **59** (1954), 388–398.

[37] A. Meir and J.W. Moon, *Some asymptotic results useful in enumeration problems.* Aequationes Math. **33** (1987), 260–268.

[38] Melvyn B. Nathanson, *Elementary Methods in Number Theory.* Graduate Texts in Mathematics **195**, Springer-Verlag, 2000.

[39] A.M. Odlyzko, *Asymptotic enumeration methods.* Handbook of Combinatorics, Vol. 1, 2, 1063–1229, Elsevier, Amsterdam, 1995.

[40] A. Oppenheim, *On an arithmetic function.* J. London Math. Soc. **1** (1926), 105–211.

[41] A. Oppenheim, *On an arithmetic function (II).* J. London Math. Soc. **2** (1927), 123–130.

[42] O. Perron, *Über das infinitäre Verhalten der Koeffizienten einer gewissen Potenzreihe.* Archiv d. Math. u. Phys., Ser. 3 **22** (1914), 329–340.

[43] O. Perron, *Über das Verhalten einer ausgearteten hypergeometrischen Reihe bei unbegrenztem Wachstum eines Parameters.* Journal f. d. reine u. angew. Mathematik, **151** (1921), 63–78.

[44] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis. I.* Springer-Verlag, 1970.

[45] L. Bruce Richmond, *Asymptotic relations for partitions.* J. Number Theory **7** (1975), 389–405.

[46] L. Bruce Richmond, *A general asymptotic result for partitions.* Canad. J. Math. **27** (1975), 1083–1091.

[47] L. Bruce Richmond, *Asymptotic results for partitions. I. The distribution of certain integers.* J. Number Theory **8** (1976), 372–389.

[48] L. Bruce Richmond, *Asymptotic results for partitions. II. A conjecture of Bateman and Erdös.* J. Number Theory **8** (1976), 390–396.

[49] L. Bruce Richmond, *The moments of partitions. II.* Acta Arith. **28** (1975), 229–243.

[50] L. Bruce Richmond, *Asymptotic relations for partitions.* Trans. Amer. Math. Soc. **219** (1976), 379–385.

[51] K.F. Roth and G. Szekeres, *Some asymptotic formulae in the theory of partitions.* Quart. J. Math. Oxford Ser. 2 **5** (1954), 241–259.

[52] Th. Skolem, *Untersuchung über die Axiome des Klassenkalküls und über Produktations- und Summationsprobleme, welche gewisse Klassen von Aussagen betreffen.* Videnskabsakademiet i Kristiania, Skrifter I, No. 3, 1919, pp. 37. Also in "Selected Works in Logic by Th. Skolem", ed. by Jens Erik Fenstak, Scand. Univ. Books, Universitetsforlaget, Oslo, 1970.

[53] A.J. Stam, *Polynomials of binomial type and compound Poisson processes.* J. Math. Anal. Appl. **130** (1988), 493–508.

[54] E.C. Titchmarsh, *The Theory of Functions.* Oxford University Press, (first ed. 1932) 1997.

[55] Richard Warlimont, *Handwritten notes on* $\Delta(\mathsf{P}) = 0$. January, 2000.

[56] Herbert S. Wilf, *Generatingfunctionology.* 2nd ed., Academic Press, Inc., 1994.

[57] Alan R. Woods, *Counting finite models.* J. Symbolic Logic **62** (1997), 925–949.

Symbol Index

Subject Index

Selected Titles in This Series

(Continued from the front of this publication)

54 **Casper Goffman, Togo Nishiura, and Daniel Waterman,** Homeomorphisms in analysis, 1997

53 **Andreas Kriegl and Peter W. Michor,** The convenient setting of global analysis, 1997

52 **V. A. Kozlov, V. G. Maz′ya, and J. Rossmann,** Elliptic boundary value problems in domains with point singularities, 1997

51 **Jan Malý and William P. Ziemer,** Fine regularity of solutions of elliptic partial differential equations, 1997

50 **Jon Aaronson,** An introduction to infinite ergodic theory, 1997

49 **R. E. Showalter,** Monotone operators in Banach space and nonlinear partial differential equations, 1997

48 **Paul-Jean Cahen and Jean-Luc Chabert,** Integer-valued polynomials, 1997

47 **A. D. Elmendorf, I. Kriz, M. A. Mandell, and J. P. May (with an appendix by M. Cole),** Rings, modules, and algebras in stable homotopy theory, 1997

46 **Stephen Lipscomb,** Symmetric inverse semigroups, 1996

45 **George M. Bergman and Adam O. Hausknecht,** Cogroups and co-rings in categories of associative rings, 1996

44 **J. Amorós, M. Burger, K. Corlette, D. Kotschick, and D. Toledo,** Fundamental groups of compact Kähler manifolds, 1996

43 **James E. Humphreys,** Conjugacy classes in semisimple algebraic groups, 1995

42 **Ralph Freese, Jaroslav Ježek, and J. B. Nation,** Free lattices, 1995

41 **Hal L. Smith,** Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems, 1995

40.4 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon,** The classification of the finite simple groups, number 4, 1999

40.3 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon,** The classification of the finite simple groups, number 3, 1998

40.2 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon,** The classification of the finite simple groups, number 2, 1995

40.1 **Daniel Gorenstein, Richard Lyons, and Ronald Solomon,** The classification of the finite simple groups, number 1, 1994

39 **Sigurdur Helgason,** Geometric analysis on symmetric spaces, 1994

38 **Guy David and Stephen Semmes,** Analysis of and on uniformly rectifiable sets, 1993

37 **Leonard Lewin, Editor,** Structural properties of polylogarithms, 1991

36 **John B. Conway,** The theory of subnormal operators, 1991

35 **Shreeram S. Abhyankar,** Algebraic geometry for scientists and engineers, 1990

34 **Victor Isakov,** Inverse source problems, 1990

33 **Vladimir G. Berkovich,** Spectral theory and analytic geometry over non-Archimedean fields, 1990

32 **Howard Jacobowitz,** An introduction to CR structures, 1990

31 **Paul J. Sally, Jr. and David A. Vogan, Jr., Editors,** Representation theory and harmonic analysis on semisimple Lie groups, 1989

30 **Thomas W. Cusick and Mary E. Flahive,** The Markoff and Lagrange spectra, 1989

29 **Alan L. T. Paterson,** Amenability, 1988

28 **Richard Beals, Percy Deift, and Carlos Tomei,** Direct and inverse scattering on the line, 1988

27 **Nathan J. Fine,** Basic hypergeometric series and applications, 1988

For a complete list of titles in this series, visit the AMS Bookstore at **www.ams.org/bookstore/**.

ISBN 0-8218-2666-2

9 780821 826669

SURV/86