Partitions into Prime Powers and Related Divisor Functions

by

Roger Mullen Woodford

B.Sc., The University of Manitoba, 2003
M.Sc., The University of British Columbia, 2005

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

The Faculty of Graduate Studies

(Mathematics)

The University Of British Columbia

(Vancouver)

August, 2008

Abstract

In this thesis, we will study a class of divisor functions: the prime symmetric functions. These are polynomials over $\mathbb{Q}$ in the so-called *elementary* prime symmetric functions, whose values lie in $\mathbb{Z}$. The latter are defined on the nonnegative integers and take the values of the elementary symmetric functions applied to the multi-set of prime factors (with repetition) of an integer n.

Initially we look at basic properties of prime symmetric functions, and consider analogues of questions posed for the usual sum of proper divisors function, such as those concerning perfect numbers or Aliquot sequences. We consider the inverse question of when, and in how many ways a number n can be expressed as $f(m)$ for certain prime symmetric functions f. Then we look at asymptotic formulae for the average orders of certain fundamental prime symmetric functions, such as the arithmetic function whose value at n is the sum of k-th powers of the prime divisors (with repetition) of n. For these last functions in particular, we also look at statistical results by comparing their distribution of values with the distribution of the largest prime factor dividing n.

In addition to average orders, we look at the modular distribution of prime symmetric functions, and show that for a fundamental class, they are uniformly distributed over any fixed modulus. Then our focus shifts to the related area of partitions into prime powers. We compute the appropriate asymptotic formulae, and demonstrate important monotonicity properties.

We conclude by looking at iteration problems for some of the simpler prime symmetric functions. In doing so, we consider the empirical basis for certain conjectures, and are left with many open problems.

Table of Contents

Abstract ii

Table of Contents iii

List of Tables v

Glossary of Fundamental Notation vi

Acknowledgements vii

Dedication viii

Preface ix

1 Divisor Functions 1
1.1 Perfect Numbers, Variations on Perfect Numbers, and Aliquot Sequences 2

2 Prime Symmetric Divisor Functions 4
2.1 Elementary Prime Symmetric Divisor Functions 5
2.2 Basic Properties of the Elementary Prime Symmetric Functions 8

3 The Second Elementary Prime Symmetric Function 13
3.1 $\Omega(n) = 3$ 13
3.2 $\Omega(n) = 4$ 14
3.3 $\Omega(n) = 5$ 14
3.4 Inverse Problems 16

4 Higher Elementary Prime Symmetric Functions 22
4.1 The Hunt for s_3-cycles 25

5 The Power Prime Symmetric Divisor Functions 30
5.1 The Average Order of e_k 32
5.2 A Statistical Look at e_k 36

6 The Average Order of $s_{k,\ell}$ 47

7 Modular Distribution of Prime Symmetric Functions . . . 58
7.1 Perron's Formula 58
7.2 Bounding $L(s, \chi)$ and $\zeta(s)$ 61
7.3 Additive Prime Symmetric Functions 65
7.4 The Distribution of s_k Modulo q 70
7.4.1 The Dirichlet Series 71
7.5 Examples 77
7.5.1 s_2 modulo 3 77
7.5.2 s_3 modulo 2 79
7.5.3 s_k modulo a prime q 81

8 Partitions into Prime Powers 82
8.1 Introduction 82
8.2 Monotonicity 82
8.3 Asymptotic Formulae 98
8.3.1 Asymptotic Formula for the Generating Function . . 99
8.3.2 Bounding from Above 107
8.3.3 Bounding from Below 109

9 Sequences of Iterates 117
9.1 Density of Sets with Given Terminal Value 117
9.2 Distribution of Products in Partitions into Primes 125
9.3 The Number of Iterations of s until termination 126
9.3.1 The Average order of $t(n)$ 129
9.4 Iterating $s + b$ 130

Bibliography 135

Appendices

A Maple Algorithms 139

List of Tables

Table 1. Initial values of $r(k)$ . . . 11
Table 2. s_k^*-perfect numbers . . . 25
Table 3. $s_3(n)$ modulo 2 . . . 81
Table 4. Selected values of $\mathfrak{B}_p(x)$ . . . 119
Table 5. Selected values of $\mathfrak{B}_p(x)/x$ . . . 121
Table 6. Iterations of s_2 terminating in 39 . . . 124
Table 7. Average order of $t(n)$ . . . 130
Table 8. $(s+b)$-cycles . . . 134

Glossary of Fundamental Notation

$\mathbb{N}_0$, $\mathbb{N}_\infty$, $\mathbb{N}_{0,\infty}$, $\mathbb{P}$	4
s_k	4
$s_{k,\ell}$	4
e_k	4
$r(k)$	9
$r_f, r_{k,\ell}, r_k$	16
$E_k(n,r)$	22
$p_A(n)$	37
$p_A^{(k)}(n)$	37, 82
$P(n), b_k(x,y), \Psi(x,y)$	38
$B_m, \mathfrak{B}_m$	117
t_f	126

Acknowledgements

Two faculty members at UBC have given me the invaluable insight, direction, encouragement and support that has made this possible. For these things it is my privilege to thank and acknowledge my supervisor Dr. Izabella Laba, and Dr. Greg Martin, who is also on my supervisory committee.

I also wish to express my gratitude to the Natural Sciences and Engineering Research Council, and the University of British Columbia for generous funding and employment opportunities.

Dedication

To my wife Trish, a friend and source of inspiration and moderation through this and many other adventures.

Preface

Who so in pompe of proud estate (quoth she)
Does swim, and bathes himselfe in courtly blis,
Does waste his dayes in darke obscuritee,
And in oblivion ever buried is:
Where ease abounds, yt's eath to doe amis;
But who his limbs with labours, and his mind
Behaves with cares, cannot so easie mis.
Abroad in armes, at home in studious kind
Who seekes with painfull toile, shall honor soonest find.

Belphoebe to Braggadocchio
The Faerie Queene
Sir Edmund Spenser

The work found here began in 2001 when I began to think about iterating the sum of prime factors with repetition function. Later, I generalized this function by looking at other symmetric polynomials in the paper "A Variation on Perfect Numbers" [46] published in Integers: Electronic Journal of Combinatorial Number Theory, in 2004, which in turn became the starting point for my Masters thesis [47].

With tools from analytic number theory and the theory of partitions, I have been able to prove many new results and connect them to other areas, and research done by prominent mathematicians. One of the key examples of work that is important to us is that of Hardy and Ramanujan on partition functions [19] which in the instance of partitions into powers of primes we correct and generalize. To do this we cite the work of Bateman and Erdős ([3], [4]) on the monotonicity of partition functions, which we use to find bounds for when difference functions of partitions into primes cease to be negative.

Levan's work [27] on the average order of the sum of k-th powers of primes is important to us as well, as is that of Alladi and Erdős on the uniform distribution modulo 2 of the sum of prime factors function. We will compute the average orders of functions generalizing these, and extend the modular distribution results to other moduli. Work on the functions we generalize in this treatise is scattered across the literature of the twentieth century. It is our intent not only to bring this work together and add to it as much as possible, but to ask fresh questions, and provide some fresh answers as well.

The majority of the work found in this treatise is new, and much of that which was contained in my Masters thesis has been improved upon. Chapters 1 through 4, the initial pages of Chapter 5, and Section 9.4 originate in my Masters thesis, and are included for completeness with appropriate revisions and notational changes.

The average order and statistical results of Chapters 5 and 6 have now been published in a separate paper in Integers [50]. Chapter 7, dealing with modular distribution derives from a paper submitted to the Journal of Number Theory. The monotonicity results of Chapter 8 are based on my paper [49] published in the Journal of Integer Sequences, though they have been thoroughly revised with a fresh approach. As well, a paper in which the asymptotic formulae for partitions into prime powers and their difference functions are derived (also in Chapter 8) has been accepted for publication in the Canadian Journal of Mathematics [48].

Finally, a word on notation: much of the notation used in this thesis is chapter, or occasionally, section specific. For this reason, a comprehensive glossary of notation would be challenging, or even misleading in the few instances where the same, or similar symbols are used to mean different things in different sections. To remedy this problem I have done two things. First, I have endeavored to make it abundantly clear in the context of each chapter (or section) what meanings are to be attributed to which symbols. Second, in the "Glossary of Fundamental Notation" I have indicated the page on which some of the universal symbols are defined.

Chapter 1

Divisor Functions

The sum of divisors function $\sigma : \mathbb{N} \to \mathbb{N}$ defined by

$$\sigma(n) = \sum_{d|n} d$$

has been thoroughly studied for centuries. The first, most natural generalization of σ are the functions σ_α defined by

$$\sigma_\alpha(n) = \sum_{d|n} d^\alpha,$$

where $\alpha \geq 0$ is a real number. With this definition, σ_0 counts the number of positive divisors of n. Sometimes σ_0 is denoted by d, and σ_1 is simply σ.

Other relevant divisor-related functions are ω, which counts the number of distinct primes dividing a number n, and Ω, which counts the number of primes dividing n, with multiplicity.

The functions σ_α are multiplicative, that is if m and n are relatively prime, then $\sigma_\alpha(mn) = \sigma_\alpha(m)\sigma_\alpha(n)$.

Derivations for the asymptotic formulae of the average orders for σ_α can be found in Apostol [2]. In case $\alpha = 0$, we have that for all $x \geq 1$,

$$\sum_{n \leq x} d(n) = x \log x + (2C - 1)x + O(\sqrt{x}),$$

where C is Euler's constant. Also,

$$\sum_{n \leq x} \sigma(n) = \frac{\pi^2}{12} x^2 + O(x \log x).$$

Furthermore, for $\alpha > 0$, and $\alpha \neq 1$, we have

$$\sum_{n \leq x} \sigma_\alpha(n) = \frac{\zeta(\alpha + 1)}{\alpha + 1} x^{\alpha+1} + O(x^\beta),$$

where $\beta = \max\{1, \alpha\}$, and ζ is the Riemann-zeta function.

In this exposition, we will define and study a new class of divisor functions: the prime symmetric divisor functions. These functions are not multiplicative in general. The following notions, those of perfect numbers, and aliquot sequences, and some of their generalizations based on the usual multiplicative divisor functions, or related functions, will be extended to the prime symmetric divisor functions.

1.1 Perfect Numbers, Variations on Perfect Numbers, and Aliquot Sequences

Several good texts detailing the basic theory of perfect numbers exist (cf. [10], [15], [28] and [42]). In addition, many variations on perfect numbers have been defined and studied. Examples can be found in [12], [20], [23], [29], [30], [35], [40] and [41].

Let $\sigma^*(n) = \sigma(n) - n$, that is $\sigma^*(n)$ sums all the proper divisors of n. We say that n is perfect if $\sigma^*(n) = n$, excessive (or abundant) if $\sigma^*(n) > n$, and defective (or deficient) if $\sigma^*(n) < n$. The first few perfect numbers are 6, 28, 496, 8128, 33550336.

It has been known since Euclid's day that if p, and $q = 2^p - 1$ are prime, then the number $n = 2^{p-1}(2^p - 1)$ is perfect. This can be readily checked, since its proper divisors correspond to the set $\{1, 2, \ldots, 2^{p-1}, q, 2q, \ldots, 2^{p-2}q\}$, and hence their sum is:

$$\begin{aligned}\sigma^*(n) &= 1 + 2 + \cdots + 2^{p-1} + q + 2q + \cdots + 2^{p-2}q \\ &= 2^p - 1 + q(2^{p-1} - 1) = (2^p - 1)2^{p-1} = n.\end{aligned}$$

Euler proved conversely, that any even perfect number must be of this form. The proof is the same as that found in [33].

Theorem 1.1. *An even number n is perfect if and only if $n = 2^{p-1}q$, where p, and $q = 2^p - 1$ are both prime.*

Proof. We need only prove the only if part. Suppose n is an even perfect number. Write $n = 2^{k-1}m$, where $2^{k-1}||n$, and so $k \geq 2$, and m is odd. Since $\sigma(n) = 2n$, we have:

$$2n = 2^k m = \sigma(2^{k-1}m) = (2^k - 1)\sigma(m).$$

Hence, $2^k - 1$ divides m, so we may write $m = (2^k - 1)\ell$. Using this value of m, we get

$$2^k \ell = \sigma((2^k - 1)\ell).$$

If $\ell > 1$, then 1, ℓ, and $(2^k - 1)\ell$ are distinct divisors of m, and so

$$2^k \ell = \sigma((2^k - 1)\ell) \geq 1 + \ell + (2^k - 1)\ell = 2^k \ell + 1,$$

a contradiction. Thus $\ell = 1$ and so

$$2^k = \sigma(2^k - 1) = 1 + (2^k - 1) + \sum_{\substack{d \mid (2^k - 1) \\ 1 < d < 2^k - 1}} d,$$

so clearly, $2^k - 1$ must be prime. Furthermore, it is trivial to see that if k is composite, then $2^k - 1$ factors, thus q is also prime. □

It is still not known if there are infinitely many even perfect numbers, or equivalently, if there are infinitely many Mersenne primes (primes of the form $2^p - 1$). It is also unknown whether or not there exists an odd perfect number.

The equivalent definition for a perfect number, i.e. n is perfect if $\sigma(n) = 2n$ motivates one generalization, that of multiperfect numbers. A number n is said to be k-multiperfect if $\sigma(n) = kn$. For instance, $\sigma(120) = 360$, so 120 is 3-multiperfect. As already stated, there are numerous variations of the notion of perfect number.

The sequence $\{\sigma^{*(k)}(n)\}_{k=0}^{\infty}$, or the sequence $\{\sigma^{*(k)}(n)\}_{k=0}^{t}$ if it begins to cycle after $t + 1$ iterations, is called the aliquot sequence of n. That is, the aliquot sequence of n is obtained from n by iterating the sum of proper divisors function σ^*. The aliquot sequence of a number n can fluctuate up and down. The unsolved *Catalan-Dickson problem* is to determine whether for every n, the sequence $\{\sigma^{*(k)}(n)\}_{k=0}^{\infty}$ is eventually periodic. Instances are known when the shortest period is longer than 1. For instance:

$$12496 \xrightarrow{\sigma^*} 14288 \xrightarrow{\sigma^*} 15472 \xrightarrow{\sigma^*} 14536 \xrightarrow{\sigma^*} 14264 \xrightarrow{\sigma^*} 12496.$$

The case when the period is of length 2 has a special name. If $\sigma^{*(2)}(n) = n \neq \sigma^*(n)$, then the pair $\{n, \sigma^*(n)\}$ is called amicable. One such amicable pair is $\{220, 284\}$.

Chapter 2

Prime Symmetric Divisor Functions

The prime symmetric divisor functions are defined in terms of the *elementary* prime symmetric functions. Write $\mathbb{N}_0$, $\mathbb{N}_\infty$ and $\mathbb{N}_{0,\infty}$ for the sets $\mathbb{N} \cup \{0\}$, $\mathbb{N} \cup \{\infty\}$ and $\mathbb{N} \cup \{0, \infty\}$ respectively, and denote the set of primes by $\mathbb{P}$.

Definition 2.1. *Let $k \in \mathbb{N}_0$. Define $s_k : \mathbb{N}_0 \to \mathbb{N}_0$ as follows: if $n = 0$, then $s_k(0) = 0$, for all k. For $n > 0$, if $k = 0$, $s_k(n) = 1$. If $k > 0$, and $n = p_1 \cdots p_r$, where $r = \Omega(n)$ is the number of prime factors (with multiplicity) of n, then*

$$s_k(n) = \sum p_{i_1} \cdots p_{i_k},$$

where the sum is taken over all products of k prime factors from the multi-set $\{p_1, \ldots, p_r\}$. We say s_k is the k^{th} elementary prime symmetric function.

Note that if $k > r$, then the sum is empty, so $s_k(n) = 0$.

Definition 2.2. *A function $f : \mathbb{N}_0 \to \mathbb{Z}$ is called a prime symmetric function if it can be expressed as a polynomial over $\mathbb{Q}$ in the elementary prime symmetric functions $s_0, s_1, \ldots$.*

The elementary prime symmetric functions s_k are a special case of a larger class of functions:

Definition 2.3. *Let $k, \ell \in \mathbb{N}_0$. We define $s_{k,\ell}(0) = 0$. If $n = p_1 \cdots p_r \in \mathbb{N}$, where the p_i are primes, not necessarily distinct, then*

$$s_{k,\ell}(n) = \sum_{1 \le i_1 < \ldots < i_k \le r} (p_{i_1} \cdots p_{i_k})^\ell.$$

We write the function $s_{1,\ell}$ as e_ℓ, and call e_ℓ the ℓ-th "power prime symmetric function." Oftentimes we shall use the subscript k in place of ℓ in the power prime symmetric functions, so it is important to keep watch for that.

Note that $e_0(n) = \Omega(n)$, for $n \in \mathbb{N}$, and that $s_{k,1} = s_k$. Also observe that by the fundamental theorem of symmetric functions, the functions $s_{k,\ell}$ are indeed prime symmetric. In fact, in Chapter 5, we will express the functions e_ℓ in terms of the elementary prime symmetric functions and vice versa. In Chapter 6, we shall find the average order of $s_{k,\ell}$.

In the standard theory of symmetric polynomials, s is usually used to indicate Schur functions, and e the elementary symmetric polynomials, so please bear in mind the difference in notation.

Now we will extend the notion of perfection, aliquot sequences, etc. to more general classes of functions.

Definition 2.4. *Let $S \subset \mathbb{Z}$, and suppose $f : S \to \mathbb{Z}$. If $n \in S$ satisfies $f(n) = n$, then we say that n is f-perfect. If $f(n) > n$, then we say that n is f-excessive. If $f(n) < n$, then we say that n is f-defective. If $f(n) \geq n$, then we say that n is f-special.*

If the sequence $\{f^{(i)}(n)\}_{i=0}^{\infty}$ is well defined, then it is called the f-sequence of n. If the sequence $\{f^{(i)}(n)\}_{i=0}^{\ell}$ is well-defined, we say it is an f-sequence of length ℓ.

A finite sequence $\{n_0, \ldots, n_\ell\}$ is an f-cycle of length ℓ if the following conditions are satisfied:

1. $\ell > 1$,

2. $n_0, \ldots, n_{\ell-1}$ *are distinct and* $n_\ell = n_0$, *and*

3. $f(n_i) = n_{i+1}$, *for* $i = 0, 1, \ldots, \ell - 1$.

2.1 Elementary Prime Symmetric Divisor Functions

Observe that if $\Omega(n) < k$, we have $s_k(n) = 0$.

There is an alternate way of defining the functions s_k. Given $n = p_1 \cdots p_r \in \mathbb{N}$, set

$$S_n(x) = \prod_{i=1}^{r}(1 + p_i x).$$

Then $s_k(n)$ is the coefficient of x^k in $S_n(x)$. The empty product is taken to be 1.

Example 2.5. $s_0(12) = 1$, $s_1(12) = 2+2+3 = 7$, $s_2(12) = 2\cdot2+2\cdot3+2\cdot3 = 16$, $s_3(12) = 12$, *and* $s_4(12) = 0$.

If $\Omega(n) = k$, then we trivially have that n is s_k-perfect. This is rather uninteresting, so we define a related version of s_k-perfection to avoid this.

Definition 2.6. *Let $n \in \mathbb{N}$ satisfy $\Omega(n) > k$. If $s_k(n) = n$, then we say that n is s_k^*-perfect. If $s_k(n) \geq n$, then we say that n is s_k^*-special.*

Example 2.7. *If p is prime, then p^p is an s_{p-1}^*-perfect number, since*

$$s_{p-1}(p^p) = \binom{p}{p-1} p^{p-1} = p^p.$$

In fact, this example has a form of converse, first proved in [46]:

Theorem 2.8. *The prime power p^α is s_k^*-perfect if and only if $\alpha = p$ and $k = p - 1$.*

Proof. We have seen that this is sufficient, now suppose $k < \alpha$, and $s_k(p^\alpha) = p^\alpha$. Then

$$\binom{\alpha}{k} = p^{\alpha-k}. \tag{2.1}$$

For $1 < k < \alpha - 1$, $\binom{\alpha}{k}$ is divisible by two distinct prime factors, hence we must have $k = 1$ or $\alpha - 1$. Now $4 = 2^2$, is the only s_1^*-perfect number, and corresponds to the case where $k = 1 = \alpha - 1$. Hence we may assume $k = \alpha - 1$, which from (2.1) implies that $\alpha = p$ and $k = p - 1$. This proves the theorem. □

The next natural question to ask is: when is $p^\alpha q^\beta$ an s_k^*-perfect number for some k? This turns out to be substantially more difficult than the previous question. However, we will come up with some necessary conditions and also look at some special cases.

Theorem 2.9. *Suppose that p and q are distinct primes and that $\alpha, \beta > 0$. If $p^\alpha q^\beta$ is s_k^*-perfect, then $\alpha \neq k$, and $\beta \neq k$.*

Proof. Suppose $\alpha = k$. Then $s_k(p^\alpha q^\beta) = p^\alpha q^\beta$ is equivalent to

$$p^k + \binom{k}{k-1} p^{k-1} \binom{\beta}{1} q + \cdots + \binom{k}{1} p \binom{\beta}{k-1} q^{k-1} + \binom{\beta}{k} q^k = p^k q^\beta.$$

This is impossible as only the right side of this equation is divisible by q. We get a similar contradiction if $\beta = k$. □

From computer searches, there are two numbers of the form $p^\alpha q^\beta$ known to be s_k^*-perfect for some k. The number $48 = 2^4 \cdot 3$ is s_2^*-perfect, and $46875 = 3 \cdot 5^6$ is s_5^*-perfect. Note that these are both of the form $p^\alpha q$. It is to this form that we will specialize our next theorem.

Theorem 2.10. *Let p and q be distinct primes.*

(i) Suppose that $p > q$. If $p \geq k+1$, and $q \geq (\alpha+1)/(\alpha-k+1)$, then $p^\alpha q$ is not s_k^*-perfect.

(ii) Suppose that $p < q$. If $p \geq k+1$, then $p^\alpha q$ is not s_k^*-perfect.

Proof. (i) In the first instance, suppose that $p^\alpha q$ is s_k^*-perfect. We have

$$s_k(p^\alpha q) = \binom{\alpha}{k} p^k + \binom{\alpha}{k-1} p^{k-1} q = p^\alpha q.$$

This implies that

$$\binom{\alpha}{k} p + \binom{\alpha}{k-1} q = p^{\alpha-k+1} q,$$

or

$$\binom{\alpha}{\alpha-k} p + \binom{\alpha}{\alpha-k+1} q = p^{\alpha-k+1} q.$$

Letting $\ell = \alpha - k + 1$, we have

$$\binom{\alpha}{\ell-1} p + \binom{\alpha}{\ell} q = p^\ell q.$$

Since $p > q$, we have that

$$\binom{\alpha}{\ell-1} p + \binom{\alpha}{\ell} q < \binom{\alpha}{\ell-1} p + \binom{\alpha}{\ell} p = \binom{\alpha+1}{\ell} p.$$

If we show that the assumption in (i) implies that

$$\binom{\alpha+1}{\ell} \leq p^{\ell-1} q, \tag{2.2}$$

then we are done. But

$$\binom{\alpha+1}{\ell} = \left(\frac{\alpha+1}{\ell}\right)\left(\frac{\alpha}{\ell-1}\right)\cdots\left(\frac{\alpha-\ell+3}{2}\right)\left(\frac{\alpha-\ell+2}{1}\right).$$

There are ℓ terms in this product. The first term $\frac{\alpha+1}{\ell}$ satisfies $\frac{\alpha+1}{\ell} \leq q$ by assumption. Since

$$\frac{\alpha}{\ell-1} \leq \frac{\alpha-1}{\ell-2} \leq \cdots \leq \frac{\alpha-\ell+2}{1} \leq p,$$

by assumption, (2.2) holds and we have proved (i).

(ii) The proof of (ii) is similar. In this case

$$s_k(p^\alpha q) < \binom{\alpha+1}{\ell} q,$$

where ℓ is as before. Hence it suffices to prove that

$$\binom{\alpha+1}{\ell} \leq p^\ell.$$

Since $p \geq \alpha - \ell + 2 = k + 1$, this indeed holds, so (ii) is proved. □

2.2 Basic Properties of the Elementary Prime Symmetric Functions

The following proposition from [46] is an immediate consequence of the definition.

Proposition 2.11. *If $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, then*

$$s_k(n) = \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r \geq 0}} \binom{\alpha_1}{i_1} \cdots \binom{\alpha_r}{i_r} p_1^{i_1} \cdots p_r^{i_r}.$$

Proposition 2.12. *For nonnegative integers m and n, the function s_k satisfies*

$$s_k(mn) = \sum_{i=0}^{k} s_{k-i}(m) s_i(n).$$

Proof. If $m = 1$, or $n = 1$, the result is immediate, as it is if $k = 0$. If $k > 0$, $m = p_1 \ldots p_r$, and $n = q_1 \ldots q_s$, let

$$S = \{p_1, \ldots, p_r, q_1, \ldots, q_s\}.$$

Then

$$s_k(mn) = \sum_{\{r_1,\ldots,r_k\} \subset S} r_1 \cdots r_k.$$

We collect the terms of this sum having $k - i$ factors from m, and i factors from n. The sum of these is equal to $s_{k-i}(m)s_i(n)$. Summing as i ranges from 0 to k gives the desired result. □

Corollary 2.13. *Let n, $k \in \mathbb{N}$, and let p and q be primes, with $p < q$, and suppose $\Omega(n) \geq k$. If $pn - s_k(pn) \geq 0$, then $pn - s_k(pn) < qn - s_k(qn)$.*

Proof. Since $\Omega(n) \geq k$, we have that $s_k(n) > 0$. Therefore, since

$$pn - s_k(pn) \geq 0,$$

it follows that

$$n \geq s_{k-1}(n) + s_k(n)/p > s_{k-1}(n).$$

Hence,

$$\begin{aligned} qn - s_k(qn) &= qn - qs_{k-1}(n) - s_k(n) \\ &> pn - ps_{k-1}(n) - s_k(n) \\ &= pn - s_k(pn). \end{aligned}$$

□

In searching for s_k-cycles and s_k^*-perfect numbers, it is essential to know when $s_k(n) \geq n$. We search by fixing $\Omega(n)$, and systematically checking all products of $\Omega(n)$ primes. The corollary tells us that if $s_k(pn) < pn$, then for any $q > p$, qn is also s_k-defective.

Lemma 2.14. *Let k, $n \in \mathbb{N}$. Then there exists an $r > k$ such that if $\Omega(n) \geq r$, then n is s_k-defective. Let $r(k)$ denote the least such r. Then*

$$r(k) = \min\left\{ r \,:\, r \geq k, \binom{r}{k} < 2^{r-k} \right\}.$$

Proof. There is an $r > k$ such that the function

$$f(t) = \binom{t}{k}$$

satisfies $f(t) < g(t)$ for all $t \geq r$, where

$$g(t) = 2^{t-k},$$

since f is a polynomial, and g is an exponential function. Now suppose $t \geq r$, and let $p_1, \ldots, p_t$ be t primes. Then

$$\binom{t}{k} = \binom{t}{t-k} < 2^{t-k}, \text{ which implies}$$
$$\sum \frac{1}{p_{i_1} \cdots p_{i_{t-k}}} \leq \binom{t}{t-k} \frac{1}{2^{t-k}} < 1,$$

where the sum is taken over all $i_1, \ldots, i_{t-k}$ such that $1 \leq i_1 < \cdots < i_{t-k} \leq t$. This implies that

$$\sum p_{i_1} \cdots p_{i_k} < p_1 \cdots p_t,$$

where the sum is taken over all $i_1, \ldots, i_k$ such that $1 \leq i_1 < \cdots < i_k \leq t$. In other words,

$$s_k(p_1 \cdots p_t) < p_1 \cdots p_t.$$

Now we prove the second statement. First note that if $n = 2^{r(k)-1}$, then

$$s_k(n) = \binom{r(k)-1}{k} 2^k \geq 2^{r(k)-1-k+k} = 2^{r(k)-1} = n.$$

In other words there are numbers satisfying $\Omega(n) = r(k) - 1$ that are s_k-excessive.

Now let us see by induction that if $k < r < 2k$, then

$$2^{r-k} \leq \binom{r}{k}. \tag{2.3}$$

Note that this case does not apply when $k = 1$, so we may assume $k \geq 2$. Clearly then, (2.3) holds when $r = k + 1$. Now, for $k + 1 \leq r \leq 2k - 2$, suppose $2^{r-k} \leq \binom{r}{k}$. Then

$$2^{r+1-k} \leq 2\binom{r}{k} = 2\left(1 - \frac{k}{r+1}\right)\binom{r+1}{k} \leq \binom{r+1}{k}.$$

So (2.3) holds by induction.

The inequality

$$\binom{2k}{k} \geq 2^k$$

holds for all $k \geq 1$, and so we have $r(k) > 2k$. With this in mind, let $r(k)$ be as claimed in the last part of the statement of the lemma, for some $k \geq 1$. We argue inductively. Let $t > r(k)$, and suppose that

$$\binom{t-1}{k} < 2^{t-1-k}.$$

Then

$$2\binom{t-1}{k} < 2^{t-k}.$$

Since $t > 2k$, we have $t < 2(t - k)$, and so

$$\binom{t}{k} = \frac{t(t-1)\cdots(t-k+1)}{k!} < \frac{2(t-1)(t-2)\cdots(t-k)}{k!} = 2\binom{t-1}{k}.$$

Hence

$$\binom{t}{k} < 2\binom{t-1}{k} < 2^{t-k},$$

and the proof is complete by induction. □

Remark 2.15. *An instructive way of viewing Lemma 2.14 is the following equality of sets:*

$$\left\{ r : \binom{r}{k} < 2^{r-k} \right\} = \{r(k), r(k)+1, \ldots\}. \tag{2.4}$$

The first few values of $r(k)$ are given in the following table:

Table 1. Initial Values of $r(k)$

k	$r(k)$
1	3
2	6
3	10
4	14
5	19
6	23
7	27
8	31
9	36
10	40

From the table it appears as though $r(k)$ is approximately linear. This is indeed the case as was observed by Dr. Greg Martin who suggested the following. We will make use of Stirling's Formula to compute an asymptotic formula for $r(k)$:

$$\Gamma(x+1) = \left(\frac{x}{e}\right)^x \sqrt{2\pi x}(1 + O(1/x)), \tag{2.5}$$

as $x \to \infty$ (cf. [32], p.503). Note that $r = r(k)$ if and only if

$$\binom{r}{k} < 2^{r-k} \leq 2\left(\frac{r-k}{r}\right)\binom{r}{k}. \tag{2.6}$$

A simple induction can be used to show that $\binom{5k}{k} < 2^{4k}$ for all $k \in \mathbb{N}$, and hence that $r(k) \leq 5k$. Thus, $r(k) \asymp k$.

The inequalities (2.6) are equivalent to

$$\frac{1}{\sqrt{2\pi}} \frac{r^{r+\frac{1}{2}}}{k^{k+\frac{1}{2}}(r-k)^{r-k+\frac{1}{2}}} \left(1 + O\left(\frac{1}{k}\right)\right) < 2^{r-k}$$
$$\leq 2\left(\frac{r-k}{r}\right) \frac{1}{\sqrt{2\pi}} \frac{r^{r+\frac{1}{2}}}{k^{k+\frac{1}{2}}(r-k)^{r-k+\frac{1}{2}}} \left(1 + O\left(\frac{1}{k}\right)\right). \tag{2.7}$$

Write $c_k = r(k)/k = r/k$. Clearly $2 < c_k \leq 5$. Now (2.7) becomes

$$\frac{1}{\sqrt{2\pi}} \frac{k^{-\frac{1}{2}} c_k^{c_k k+\frac{1}{2}}}{(c_k-1)^{(c_k-1)k+\frac{1}{2}}} \left(1 + O\left(\frac{1}{k}\right)\right) < 2^{(c_k-1)k}$$
$$\leq 2\left(\frac{c_k-1}{c_k}\right) \frac{1}{\sqrt{2\pi}} \frac{k^{-\frac{1}{2}} c_k^{c_k k+\frac{1}{2}}}{(c_k-1)^{(c_k-1)k+\frac{1}{2}}} \left(1 + O\left(\frac{1}{k}\right)\right). \tag{2.8}$$

Taking logarithms and dividing through by k, we see that (2.8) is equivalent to

$$(c_k - 1)\log 2 = c_k \log c_k - (c_k - 1)\log(c_k - 1) + O\left(\frac{\log k}{k}\right). \tag{2.9}$$

Let c be the root, $c \approx 4.403497879$, of the equation

$$(c-1)\log 2 = c\log c - (c-1)\log(c-1),$$

and let $f(x) = (x-1)\log 2 - x\log x + (x-1)\log(x-1)$, on $[4,5]$, so that $f(c) = 0$. It is easily verified that $f'(c) \neq 0$, and that in an interval about c, f' is bounded, and bounded away from 0. By (2.9), $f(c_k) = O(\log k/k)$, so by the Mean Value Theorem we can conclude that

$$c_k = c + O\left(\frac{\log k}{k}\right).$$

Thus we have proved the following theorem:

Theorem 2.16. *As $k \to \infty$,*

$$r(k) = ck + O(\log k), \tag{2.10}$$

where c is the root of the equation $(c-1)\log 2 = c\log c - (c-1)\log(c-1)$.

The properties of s_1-perfection etc., corresponding to the first elementary prime symmetric function s_1 are easily characterized. The s_1-perfect numbers are the primes and 4, the latter being the only s_1^*-perfect number. All other numbers are s_1-defective. Clearly there are no s_1-cycles. We now investigate these properties in the second elementary prime symmetric function.

Chapter 3

The Second Elementary Prime Symmetric Function

To describe all s_2^*-perfect numbers and all s_2-cycles we need to find all numbers n such that $2 < \Omega(n) < 6$, with $s_2(n) \geq n$, since $r(2) = 6$. To do this we use the algorithm mentioned after Corollary 2.13, the approach taken in [46]. That is, we shall temporarily fix $\Omega(n)$, and increase the prime divisors until we cease to obtain s_k^*-special numbers. We begin with $\Omega(n) = 3$.

3.1 $\Omega(n) = 3$

Applying s_2 we have:

$$\begin{aligned} s_2(2 \cdot 2 \cdot p) &= 4p + 4 > 4p, \\ s_2(2 \cdot 3 \cdot p) &= 5p + 6 < 6p, \text{ when } p > 6. \end{aligned}$$

Below we find all s_2^*-special numbers not of the forms listed above

$$\begin{aligned} s_2(2 \cdot 3 \cdot 3) &= 21 > 18, \\ s_2(2 \cdot 3 \cdot 5) &= 31 > 30, \\ s_2(2 \cdot 3 \cdot 7) &= 41 < 42, \\ s_2(3 \cdot 3 \cdot 3) &= 27, \\ s_2(3 \cdot 3 \cdot 5) &= 39 < 45. \end{aligned}$$

By Corollary 2.13 there are no other s_2^*-special numbers satisfying $\Omega(n) = 3$ than those already mentioned. Thus 27 is the only s_2^*-perfect number satisfying $\Omega(n) = 3$. Iterating on all the s_2-excessive numbers above, save those of the form $4p$, shows that none belong to an s_2-cycle. For example

$$18 \xrightarrow{s_2} 21 \xrightarrow{s_2} 21 \cdots .$$

3.2 $\Omega(n) = 4$

Applying s_2 to numbers with $\Omega(n) = 4$ we obtain

$$s_2(2 \cdot 2 \cdot 2 \cdot p) = 6p + 12 < 8p, \text{ when } p > 6.$$

From this together with Corollary 2.13 we can conclude that there are only finitely many s_2^*-special numbers with $\Omega(n) = 4$. Iterating on $8p$ for $p = 2$, 3, 5, shows that none belongs to an s_2-cycle. Checking other cases:

$$\begin{aligned} s_2(2 \cdot 2 \cdot 3 \cdot 3) &= 37 > 36, \\ s_2(2 \cdot 2 \cdot 3 \cdot 5) &= 51 < 60, \\ s_2(2 \cdot 3 \cdot 3 \cdot 3) &= 45 < 54. \end{aligned}$$

Hence there are no s_2^*-perfect numbers satisfying $\Omega(n) = 4$. Iterating on the above s_2-excessive numbers shows that none belong to an s_2-cycle.

3.3 $\Omega(n) = 5$

Applying s_2 to numbers with $\Omega(n) = 5$ we obtain

$$s_2(2 \cdot 2 \cdot 2 \cdot 2 \cdot p) = 8p + 24 < 16p, \text{ when } p > 3.$$

Thus by Corollary 2.13 there are only finitely many s_2^*-special numbers with $\Omega(n) = 5$. Iterating on $16p$ for $p = 2$, 3, shows that 48 is in fact s_2^*-perfect, and 32, which is s_2-excessive, does not belong to an s_2-cycle. Checking other cases:

$$s_2(2 \cdot 2 \cdot 2 \cdot 3 \cdot 3) = 57 < 72,$$

Hence 48 is the only s_2^*-perfect number satisfying $\Omega(n) = 5$. We have proved the following theorem.

Theorem 3.1. *The numbers 27 and 48 are the only s_2^*-perfect numbers.*

Theorem 3.2. *There are no s_2-cycles.*

Proof. An s_2-cycle must have a least element that is s_2-excessive. It is easily verified (as it was above in the case $n = 18$), that all of the finitely many s_2-excessive numbers not of the form $4p$ do not belong to s_2-cycles. Thus any such element must belong to the family of numbers of the form $4p$.

We will show that in all but a few trivial cases $s_2(s_2(4p)) < 4p$, giving a contradiction. Now, $s_2(4p) = 8((p+1)/2)$. We may assume that p is odd, and set $m = (p+1)/2$. Thus we will have a contradiction if the following holds:

$$s_2(8m) < 8m - 4.$$

This is equivalent to

$$12 + 6s_1(m) + s_2(m) < 8m - 4, \tag{3.1}$$

which is equivalent to

$$\frac{16}{p_1 \cdots p_s} + 6\sum_{i=1}^{s} \frac{1}{p_1 \cdots \hat{p_i} \cdots p_s} + \sum_{1 \le i < j \le s} \frac{1}{p_1 \cdots \hat{p_i} \cdots \hat{p_j} \cdots p_s} < 8,$$

where $m = p_1 \cdots p_s$. Here $p_1 \cdots \hat{p_i} \cdots p_s$ is defined to be $p_1 \cdots p_s/p_i$, and $p_1 \cdots \hat{p_i} \cdots \hat{p_j} \cdots p_s$ is defined to be $p_1 \cdots p_s/p_i p_j$.

Since $p_i \ge 2$, this expression is implied by:

$$\frac{16}{2^s} + \frac{6s}{2^{s-1}} + \frac{s(s-1)}{2}\frac{1}{2^{s-2}} < 8,$$

which holds for all $s \ge 4$. If $s = 1$, then m is prime, and so condition (3.1) becomes:

$$12 + 6m < 8m - 4,$$

which holds for all $m > 8$. It is easily verified for $m = 2$, 3, 5 and 7, that $8m$ does not belong to an s_2-cycle.

For $s = 2$, we can write $m = pq$. The only values of m for which (3.1) fails are determined by the prime pairs $(p, q) = (2, 2)$, $(2, 3)$. In both cases, $8m$ does not belong to an s_2-cycle.

Finally for $s = 3$, if $m = pqr$, only for the triple $(p, q, r) = (2, 2, 2)$ does m fail to satisfy (3.1). Again, in this case, $8m$ does not belong to an s_2-cycle. □

Remark 3.3. *The longest increasing s_2-sequence is*

$$8 \xrightarrow{s_2} 12 \xrightarrow{s_2} 16 \xrightarrow{s_2} 24 \xrightarrow{s_2} 30 \xrightarrow{s_2} 31.$$

3.4 Inverse Problems

For a given prime symmetric function f, and an integer n, we are interested in the number of solutions to the equation $f(m) = n$. This leads to a new class of arithmetic functions. We state the following definition in full generality, then specialize it.

Definition 3.4. *Let $S \subseteq \mathbb{Z}$, and suppose $f : S \to \mathbb{Z}$. Let $r_f : \mathbb{N} \to \mathbb{N}_{0,\infty}$ be the function whose value at n is*

$$r_f(n) = |f^{-1}[\{n\}]|.$$

In case there are infinitely many solutions to the equation $f(m) = n$, we say $r_f(n) = \infty$. Moreover, let

$$r_{s_{k,\ell}} = r_{k,\ell}, \text{ and}$$
$$r_{s_k} = r_k.$$

Example 3.5. *$r_1(1) = 0$, but for all $n \geq 2$, $r_1(n) \geq 1$. In fact, $\lim_{n\to\infty} r_1(n) = \infty$. To see this, simply set*

$$n = s_1(2^a 3^b) = 2a + 3b,$$

and observe that the number of pairs (a, b) satisfying this equation can be made arbitrarily large for all n sufficiently large.

We now prove a weaker result for r_2, as found in [46].

Theorem 3.6. *There exists an $N \in \mathbb{N}$ such that for all $m \geq N$, $r_2(m) \geq 1$.*

Proof. It suffices to show that for m sufficiently large, $m = s_2(2^a 3^b 5^c 7^d)$, for some a, b, c, and $d \geq 0$. In general,

$$\begin{aligned} s_2(2^a 3^b 5^c 7^d) &= 4\binom{a}{2} + 9\binom{b}{2} + 25\binom{c}{2} + 49\binom{d}{2} \\ &+ 6\binom{a}{1}\binom{b}{1} + 10\binom{a}{1}\binom{c}{1} + 14\binom{a}{1}\binom{d}{1} \\ &+ 15\binom{b}{1}\binom{c}{1} + 21\binom{b}{1}\binom{d}{1} + 35\binom{c}{1}\binom{d}{1} \\ &= \frac{1}{2}\left[(2a + 3b + 5c + 7d)^2 - (4a + 9b + 25c + 49d)\right] \end{aligned}$$

So, given m, we need only find solutions to the equations:

$$2a + 3b + 5c + 7d = R,$$
$$4a + 9b + 25c + 49d = R^2 - 2m,$$

with nonnegative integers a, b, c, d, and $R \in \mathbb{N}$. These equations are equivalent to:

$$2a - 10c - 28d = 3R - R^2 + 2m, \tag{3.2}$$

$$3b + 15c + 35d = R^2 - 2R - 2m, \tag{3.3}$$

Since a and b must be nonnegative integers, we have the following necessary and sufficient conditions for a solution to (3.2) and (3.3):

1. $2m \equiv R^2 + R + d \pmod 3$,
2. $R^2 - 3R - 10c - 28d \leq 2m$,
3. $2m \leq R^2 - 2R - 15c - 35d$.

Note that equation (3.2) is always satisfied modulo 2. Condition 1 results from taking equation (3.3) modulo 3, and conditions 2 and 3 are derived from the fact that $a, b \geq 0$.

Consider the interval

$$I_R(c, d) = [R^2 - 3R - 10c - 28d, R^2 - 2R - 15c - 35d].$$

For fixed d, let $c_R(d)$ be the least c such that $\ell(I_R(c, d)) < 15$, where $\ell(I)$ denotes the length of an interval I. We use the notation $L(I)$ and $R(I)$ to denote the left and right endpoints of an interval I, respectively. Since $R(I_R(c, d)) = R(I_R(c + 1, d)) + 15$, when they exist, we have that

$$\bigcup_{c=0}^{c_R(d)} I_R(c, d) = [R^2 - 3R - 10c_R(d) - 28d, R^2 - 2R - 35d].$$

Denote the above interval by $\mathcal{I}_R(d)$. By definition of $c_R(d)$,

$$\begin{aligned} \ell(I_R(c_R(d), d)) &= R - 5c_R(d) - 7d \\ &< 15, \text{ so} \\ -10c_R(d) &< -2R + 30 + 14d, \end{aligned}$$

and $c_R(d)$ is the least such c. Consider the interval $\bigcap_{d=0}^{2} \mathcal{I}_R(d)$. Clearly $R(\bigcap_{d=0}^{2} \mathcal{I}_R(d)) = R^2 - 2R - 70$. We now wish to find an upper bound for $L(\bigcap_{d=0}^{2} \mathcal{I}_R(d))$. From the above inequality, we have that

$$L(\mathcal{I}_R(d)) = R^2 - 3R - 10c_R(d) - 28d < R^2 - 5R - 14d + 30.$$

Thus

$$\begin{aligned} L\left(\bigcap_{d=0}^{2} \mathcal{I}_R(d)\right) &= \max\{R^2 - 3R - 10c_R(d) - 28d | d = 0, 1, 2\} \\ &< \max\{R^2 - 5R - 14d + 30 | d = 0, 1, 2\} \\ &= R^2 - 5R + 30. \end{aligned}$$

Let $J_R = [R^2 - 5R + 30, R^2 - 2R - 70]$. Then $J_R \subset \bigcap_{d=0}^{2} \mathcal{I}_R(d)$. Now

$$\begin{gathered} L(J_{R+1}) \leq R(J_R), \text{ if and only if} \\ R^2 - 3R + 26 \leq R^2 - 2R - 70, \end{gathered}$$

which holds for all $R \geq 96$. So if $2m \geq L(J_{96}) = 8766$, then there is an $R \geq 96$ such that $2m \in J_R \subset \bigcap_{d=0}^{2} \mathcal{I}_R(d)$. Choose $d \in \{0, 1, 2\}$ such that condition 1 is satisfied. Since $2m \in \mathcal{I}_R(d)$, there is a $c \geq 0$ such that $2m \in I_R(c, d)$. For these values of R, c, and d, conditions 2 and 3 are satisfied. In other words, there exists an n such that $m = s_2(n)$. This completes the proof. □

Let us try and improve on this result.

Theorem 3.7. $\lim_{m\to\infty} r_2(m) = \infty$

Proof. Let

$$A_k = \sum_{i=1}^{\infty} p_i^k x_i,$$

where p_i denotes the i^{th} prime, $x_i \geq 0$, and $x_i = 0$ for all but finitely many values of i. If

$$n = \prod_{i=1}^{\infty} p_i^{x_i},$$

then

$$s_2(n) = \frac{1}{2}[A_1^2 - A_2].$$

Set $s_2(n) = m$ and then set $A_1 = R$, as before. This gives two equations in the unknowns $x_1, x_2, \ldots$, and R, namely:

$$2x_1 + 3x_2 + \cdots = R \tag{3.4}$$
$$2^2 x_1 + 3^2 x_2 + \cdots = R^2 - 2m. \tag{3.5}$$

Eliminating x_2 from the first and x_1 from the second gives

$$2x_1 - 10x_3 - \cdots - (p_r^2 - 3p_r)x_R - \cdots = -R^2 + 3R + 2m$$
$$3x_2 + 15x_3 + \cdots + (p_r^2 - 2p_r)x_R + \cdots = R^2 - 2R - 2m.$$

If $(R, x_3, x_4, \ldots)$, satisfy the conditions

(1) $2m \geq R^2 - 3R - 10x_3 - 28x_4 - \cdots - (p_r^2 - 3p_r)x_r - \cdots$,

(2) $2m \leq R^2 - 2R - 15x_3 - 35x_4 - \cdots - (p_r^2 - 2p_r)x_r - \cdots$,

(3) $15x_3 + 35x_4 + \cdots + (p_r^2 - 2p_r)x_r + \cdots \equiv R^2 + R + m \pmod 3$,

then they uniquely determine a solution $(R, x_1, x_2, \ldots)$ to equations (3.4) and (3.5).

Observe that given R and m, for any choice of $x_3, x_5, x_6, \ldots$, we can choose $d = x_4$ uniquely from the set $\{0, 1, 2\}$ so that the congruence (3) holds. We will do this in the following way.

Let $q_1, q_2, \ldots$ be the odd primes congruent to 2 modulo 3, listed in increasing order. Then for $y_i \geq 0$, and $r \geq 0$, we have that

$$15y_1 + 99y_2 + \cdots + (q_r^2 - 2q_r)y_r \equiv 0 \pmod 3.$$

Denote this sum as $D = D(y_1, \ldots, y_r)$, and let

$$C = \sum_{i=0}^{r} (q_i^2 - 3q_i)y_i$$

Let q be any odd prime congruent to 2 modulo 3, $k = q^2 - 3q$, $\ell = q^2 - 2q$, and suppose $\alpha \geq 0$. Consider intervals of the form

$$I_R(\alpha) = [R^2 - 3R - 28d - C - \alpha k, R^2 - 2R - 35d - D - \alpha\ell].$$

We will handle the case when R, $m \equiv 0 \pmod 3$. The other cases are virtually identical. In this case, choosing $d = 0$, we have that if $2m$ lies in an interval $I_R(\alpha) = [R^2 - 3R - C - \alpha k, R^2 - 2R - D - \alpha \ell]$, then m satisfies conditions (1), (2), and (3), with the primes q_i, q. We will show that the number of such intervals containing $2m$ tends to infinity as m does. This implies the theorem.

If $m' \equiv 0 \pmod 3$ and if $2m'$ lies in the interval

$$I_{R-1}(0) = [R^2 - 5R + 4 - C, R^2 - 4R + 3 - D],$$

then since $R \equiv 0 \pmod 3$, condition (3) is satisfied with respect to the primes occurring in the terms of C and D, $R - 1$ instead of R, and m' instead of m.

Given R sufficiently large, C, D, k, ℓ and q, let $\alpha_R \geq 0$ be the largest number such that (using notation as in the previous theorem):
(i) $L(I_R(\alpha_R)) \leq R(I_R(\alpha_R))$, and
(ii) $R(I_R(\alpha_R + 1)) \geq L(I_R(\alpha_R))$.
These conditions ensure that

$$\bigcup_{\alpha=0}^{\alpha_R} I_R(\alpha)$$

is an interval, and of longest possible length.

Since we always have $R(I_R(\alpha + 1)) \leq R(I_R(\alpha))$, then condition (ii) implies condition (i). Condition (ii) holds if

$$R^2 - 2R - D - (\alpha_R + 1)\ell \geq R^2 - 3R - C - \alpha_R k,$$

which in turn holds if and only if

$$R \geq (D - C) + \alpha_R(\ell - k) + \ell.$$

Thus we have

$$\alpha_R = \left\lfloor \frac{R - \ell - (D - C)}{\ell - k} \right\rfloor.$$

Define

$$J_R = \bigcup_{\alpha=0}^{\alpha_R} I_R(\alpha).$$

We wish to show that $J_R \cap I_{R-1}(0)$ is nonempty. That is that $L(J_R) \leq R(I_{R-1}(0))$. This is true if and only if $L(I_R(\alpha_R)) \leq R(I_{R-1}(0))$, i.e. if and only if

$$R^2 - 3R - C - \alpha_R k \leq R^2 - 4R + 3 - D.$$

Substituting in the value for α_R into the above gives

$$R^2 - 3R - C - \left\lfloor \frac{R - \ell - (D - C)}{\ell - k} \right\rfloor k \leq R^2 - 4R + 3 - D.$$

This inequality is implied by the inequality

$$R^2 - 3R - C + k - \left(\frac{R - \ell - (D - C)}{\ell - k} \right) k \leq R^2 - 4R + 3 - D,$$

which is in turn equivalent to the inequality

$$R + (D - C) - 3 + k \leq \left(\frac{R - \ell - (D - C)}{\ell - k} \right) k.$$

The coefficient of R on the left hand side of this inequality is 1. On the right hand side it is $k/(\ell - k) = q - 3 > 1$, and so for R large enough, this inequality holds.

Since q, and each q_i could be taken to be arbitrary odd primes congruent to 2 (mod 3), we can vary the primes in the terms of C and D. For each such choice, and each choice of q say with $q > q_r$, we obtain distinct solutions for every m such that 2m lies in J_R, in this case where R, $m \equiv 0 \pmod 3$.

The remaining cases, in which R and m need not be congruent to 0 (mod 3) are handled similarly, by first selecting a different value of $d \in \{0, 1, 2\}$, and showing that the corresponding interval J_R reaches low enough to intersect the interval $I_{R-1}(0)$. □

We end this section with a conjecture.

Conjecture 3.8. *For each $k, \ell \in \mathbb{N}$, $\lim_{n \to \infty} r_{k,\ell}(n) = \infty$.*

The cases when $k = 1$ are answered in this thesis. It indeed transpires that $r_{1,\ell}(n) \to \infty$. In fact, we shall do better and derive an asymptotic formula for $\log r_{1,\ell}(n)$.

Evidence for the conjecture lies in the fact that the average orders of $s_{k,\ell}$ are, as we shall see, only greater in magnitude than that of $s_{1,\ell}$ by a power of $\log \log x$, which proves to be relatively insignificant. As n increases there are more and more values of m such that $s_{1,\ell}(m) = m$, so it seems natural that it be similarly represented for $k > 1$.

Chapter 4

Higher Elementary Prime Symmetric Functions

Let $n > 1$ be an integer. By a family $E_k(n,r)$ of s_k^*-special numbers we mean a set

$$E_k(n,r) = \{np_1 \cdots p_r | p_1, \ldots, p_r \in \mathbb{P}\}$$

such that if $m \in E_k(n,r)$, then m is s_k^*-special. The family $E_2(4,1)$ of numbers of the form $4p$, where p is prime, is one such set, since the elements satisfy $s_2(4p) = 4p + 4 > 4p$. $E_k(n,0)$ merely denotes the singleton set of the number n satisfying $s_k(n) \geq n$, and $\Omega(n) > k$.

Introducing the definition $E_k(n,r)$, is productive for classifying all s_k^*-special numbers for a fixed k. We will see later that with finitely many exceptions, all such numbers are in fact s_k-excessive, and that the families $E_{k+1}(n, r+1)$ can be determined completely from the s_k^*-special numbers. As well, all but finitely many s_k^*-special numbers belong to such families.

Theorem 4.1. *(1) Let $n \in \mathbb{N}$. If n is s_k^*-special then pn is s_{k+1}-excessive for every prime p.*

(2) If pn is s_{k+1}^*-special for infinitely many primes p, then n is s_k^*-special, and hence by (1), pn is s_{k+1}-excessive for every prime p.

Proof. (1) Suppose n is s_k^*-special. Then since $\Omega(n) > k$, we have $s_{k+1}(n) > 0$. So

$$\begin{aligned} s_{k+1}(pn) &= ps_k(n) + s_{k+1}(n) \\ &\geq pn + s_{k+1}(n) \\ &> pn. \end{aligned}$$

(2) If $s_{k+1}(pn) = ps_k(n) + s_{k+1}(n) \geq pn$, for infinitely many primes p, then $s_k(n) \geq n - s_{k+1}(n)/p$. Letting $p \to \infty$, we have $s_k(n) \geq n$. □

Corollary 4.2. *For $k \in \mathbb{N}$, there are only finitely many s_k^*-perfect numbers.*

Proof. By the previous theorem, any infinite family $E_k(n,r)$, $r>0$ contains only s_k-excessive numbers. We claim that there are only finitely many s_k^*-special numbers not belonging to any such family. Indeed if there were infinitely many, then by Lemma 2.14, there would be an $r>k$ such that there are infinitely many such s_k^*-special numbers n satisfying $\Omega(n)=r$. Write them as follows:

$$S=\{p_1^{(1)}\cdots p_r^{(1)},p_1^{(2)}\cdots p_r^{(2)},p_1^{(3)}\cdots p_r^{(3)},\ldots\}$$

and assume that $p_i^{(j)}\leq p_{i+1}^{(j)}$. Clearly we must have that $\lim_{j\to\infty}p_r^{(j)}=\infty$, otherwise, we must have the same number occurring twice in the sequence. If $p_{r-1}^{(j)}$ contains a bounded subsequence, then there must be an identical product $p_1^{(\ell)}\cdots p_{r-1}^{(\ell)}$ occurring for infinitely many values of ℓ. This gives a contradiction, since it implies that $E_k(p_1^{(\ell_0)}\cdots p_{r-1}^{(\ell_0)},1)$ intersects S, where ℓ_0 is one such value of ℓ. Thus $\lim_{j\to\infty}p_{r-1}^{(j)}=\infty$. Inductively, by similar arguments, we conclude that $\lim_{j\to\infty}p_m^{(j)}=\infty$, for $m=1,2,\ldots,r$. But then

$$\lim_{j\to\infty}\frac{s_k(p_1^{(j)}\cdots p_r^{(j)})}{p_1^{(j)}\cdots p_r^{(j)}}=0.$$

This implies that the set S contains s_k-defective elements, a contradiction. □

The preceding results of this section were first proved in [46], by the author.

Corollary 4.3. *Let $k\in\mathbb{N}$. There are only finitely many $d>0$ such that the equation $s_k(n)=n+d$ has infinitely many solutions.*

Proof. Suppose $s_k(n)=n+d$ has infinitely many solutions n for some $d\in\mathbb{N}$. Of these solutions, infinitely many must belong to an infinite family of s_k-excessive numbers of the form $E_k(m,r)$. It is easily seen that we must have $r=1$. Thus there are infinitely many primes p such that $s_k(m)+ps_{k-1}(m)=pm+d$, and so $s_k(m)=d$. But there can only be finitely many such m, as $d>0$, so the Corollary holds. □

Thus the infinite families of s_3-excessive numbers are: $E_3(4,2)$, $E_3(16,1)$, $E_3(18,1)$, $E_3(24,1)$, $E_3(27,1)$, $E_3(30,1)$, $E_3(32,1)$, $E_3(36,1)$, $E_3(40,1)$, $E_3(48,1)$. That is, those corresponding to the s_2^*-special numbers $E_2(4,1)\cup\{16,18,24,27,30,32,36,40,48\}$. By exhaustive search (as was done with

$k = 2$), all other s_3^*-special numbers can be found. They constitute the following set:

$$\{42p|p = 7, 11, \ldots, 41\} \cup \{56p|p = 7, 11, \ldots, 43\} \cup \{64p|p = 2, 3, \ldots, 37\} \cup$$
$$\{726, 858, 250, 350, 225, 315, 968, 1144, 300, 420, 162, 270, 378, 243, 400, 560,$$
$$216, 360, 504, 324, 288, 480, 672, 432, 256, 384, 640, 576, 512, 768\}.$$

None of the elements in the above sets are s_3^*-perfect, hence there are no s_3^*-perfect numbers. The diversity of possible increasing s_3-sequences makes it difficult to rule out the existence of s_3-cycles as we did s_2-cycles. This is illustrated in the following example.

Example 4.4. *If p_1, q_1 are odd primes, then $s_3(4p_1q_1) = 4(p_1q_1 + p_1 + q_1)$. It is possible that $p_1q_1 + p_1 + q_1 = p_2q_2$, where p_2, q_2 are again odd primes, and so on. Several such sequences exist the longest one with $p_1q_1 < 50000$, and $p_i, q_i > 3$ is:*

$$184892 = 4 \cdot 17 \cdot 2719 \xrightarrow{s_3} 195836 = 4 \cdot 173 \cdot 283 \xrightarrow{s_3} 197660 = 4 \cdot 5 \cdot 9883$$
$$\xrightarrow{s_3} 237212 = 4 \cdot 31 \cdot 1913 \xrightarrow{s_3} 244988 = 4 \cdot 73 \cdot 839 \xrightarrow{s_3} 248636 = 4 \cdot 61 \cdot 1019$$
$$\xrightarrow{s_3} 252956 = 4 \cdot 11 \cdot 5749 \xrightarrow{s_3} 275996 = 4 \cdot 7 \cdot 9857 \xrightarrow{s_3} 315452 = 4 \cdot 17 \cdot 4639$$
$$\xrightarrow{s_3} 334076 = 4 \cdot 47 \cdot 1777 \xrightarrow{s_3} 341372 = 4 \cdot 31 \cdot 2753 \xrightarrow{s_3} 352508 = 4 \cdot 13 \cdot 6779$$
$$\xrightarrow{s_3} 379676 = 4 \cdot 11 \cdot 8629 \xrightarrow{s_3} 414236 = 4 \cdot 29 \cdot 3571 \xrightarrow{s_3} 428636 = 4 \cdot 13 \cdot 8243$$
$$\xrightarrow{s_3} 461660 = 4 \cdot 5 \cdot 41 \cdot 563.$$

It seems highly unlikely however that an increasing s_3-sequence be infinite. This is part of the following conjecture:

Conjecture 4.5. *Any increasing s_k-sequence is finite.*

A related conjecture is given by:

Conjecture 4.6. *For any $n \in \mathbb{N}_0$, the s_k-sequence of n is ultimately periodic.*

The evidence for these conjectures lies principally in the fact that for a fixed k, with finitely many exceptions, all s_k^*-excessive numbers belong to one of the families $E_k(n, r)$. The union of all such families has asymptotic density 0 in the positive integers.

Below is a complete list of all s_k^*-perfect numbers for $k = 1, 2, 3$, and 4, found by exhaustive search.

Table 2. s_k^*-Perfect Numbers

k	s_k^*-perfect numbers
1	4
2	27, 48
3	none
4	3125, 9315, 31280

4.1 The Hunt for s_3-cycles

As just noted, ruling out the existence of s_3-cycles is not a trivial matter, since there seems to be no apparent maximum length on increasing s_3-sequences. Nonetheless, there are a number of things which can be done to get a better picture of what such a sequence can look like.

Using Maple, computer searches were done to verify that many numbers do not belong to an s_3-cycle. These include:

$$\{42p|p = 7, 11, \ldots, 41\} \cup \{56p|p = 7, 11, \ldots, 43\} \cup \{64p|p = 2, 3, \ldots, 37\} \cup \\ \{250, 350, 225, 315, 968, 1144, 300, 420, 162, 270, 378, 243, 400, 560, 216, 360, \\ 504, 324, 288, 480, 672, 432, 256, 384, 640, 576, 512, 768\},$$

that is, all those s_3^*-special numbers not belonging to any of the infinite families.

It was also shown for each of the following sets: $E_3(16, 1)$, $E_3(18, 1)$, $E_3(24, 1)$, $E_3(27, 1)$, $E_3(30, 1)$, $E_3(32, 1)$, $E_3(36, 1)$, $E_3(40, 1)$, $E_3(48, 1)$, which are all of the form $\{np|p \text{ is prime}\}$, for all $p < 1000$, that none of the numbers belong to an s_3-cycle. Finally, the same was shown for the set $E_3(4, 2) = \{4pq|p, q \text{ are prime}\}$ in the following ranges: $2 \le p \le 30, 2 \le q \le 1000$, and $2 \le p \le 100, 2 \le q \le 100$.

Thus an s_3-cycle must have a least element belonging to one of the 10 infinite sets $E_3(n, r)$. We label these sets in the following way: we say m is type 1 if $m \in E_3(4, 2)$. Similarly we identify elements in the following sets:

- $E_3(16, 1)$ - type 2,
- $E_3(18, 1)$ - type 3,
- $E_3(24, 1)$ - type 4,
- $E_3(27, 1)$ - type 5,
- $E_3(30, 1)$ - type 6,
- $E_3(32, 1)$ - type 7,
- $E_3(36, 1)$ - type 8,
- $E_3(40, 1)$ - type 9,

•$E_3(48,1)$ - type 10.

To analyze what increasing sequences begin with elements of types 1 to 10, we will consider the effect of applying s_3 to such a number, and consider what type the resulting number may be, if any type. We will use the notation type $a \longrightarrow$ type b to indicate the following: it is not impossible for divisibility reasons for to have numbers m and $s_3(m)$ of type a and b respectively. We have noted that type 1 $\longrightarrow$ type 1 is possible. Below is a complete list of all the possibilities. Note that we are assuming that the numbers do not lie in the ranges ruled out by the computer searches.

•type 1 $\longrightarrow$ type 1, type 8,

•type 2 $\longrightarrow$ type 9,

•type 7 $\longrightarrow$ type 9,

•type 10 $\longrightarrow$ type 9.

This list was arrived at by simply ruling out possibilities. For example, if $m = 18p$ is type 3, then $s_3(18p) = 21p + 18 = 3(7p + 6)$. Since we may assume $p > 1000$, p is clearly not 2 or 3, so $3||s_3(18p)$, and $2 \nmid s_3(18p)$. Thus it is impossible for $s_3(18p)$ to be of any of types 1 through 10. The list arrived at above includes all those cases that could not be ruled out in like manner.

Next we will rule out the possibility that numbers of the types that can not be of another type upon application of s_3, are not least elements of any s_3-cycles.

Theorem 4.7. *If n is type 3, 4, 5, 6, 8, or 9, then n is not the least element of an s_3-cycle.*

Proof. First suppose $n = 18p$ is type 3. By the computer searches mentioned earlier in this section, we may assume that $p > 1000$. Applying s_3 we get $s_3(18p) = 21p + 18 = 3(7p + 6)$. Let $m = 7p + 6$. If m is prime, then the s_3-sequence terminates, as it does if m is the product of 2 primes. If $m = q_1q_2q_3$ is the product of 3 primes, then we have

$$s_3^{(2)}(n) < n \text{ if and only if}$$

$$3(q_1q_2 + q_1q_3 + q_2q_3) + q_1q_2q_3 < 18\left(\frac{q_1q_2q_3 - 6}{7}\right) \text{ if and only if}$$

$$108 + 21(q_1q_2 + q_1q_3 + q_2q_3) < 11q_1q_2q_3. \tag{4.1}$$

Since $p > 1000$, we have that $q_1q_2q_3 > 7006$. Also, clearly $q_i \neq 2, 3$, or 7. Under these constraints, the inequality (4.1) holds, and so n cannot be the least element of an s_3-cycle.

Hence we may assume $m = q_1 \cdots q_r$, where $r \geq 4$. We may again assume that $m > 7006$, and $q_i \neq 2, 3$, or 7. In this case,

$$s_3^{(2)}(n) < n \text{ if and only if}$$

$$108 + 21(q_1q_2 + \cdots + q_{r-1}q_r) + 7(q_1q_2q_3 + \cdots + q_{r-2}q_{r-1}q_r) < 18q_1 \cdots q_r.$$

Upon dividing each side by m, and using the facts that $m > 7006$, and $q_i \geq 5$, we have that the latter inequality is implied by the inequality

$$\frac{108}{7006} + \frac{21}{5^{r-2}}\binom{r}{2} + \frac{7}{5^{r-3}}\binom{r}{3} < 18.$$

This indeed holds for $r \geq 4$. Thus numbers of type 3 cannot be least elements of s_3-cycles.

The remaining cases, types 4, 5, 6, 8, and 9 are handled in virtually identical fashion, but are omitted to avoid tediousness. □

Remark 4.8. *A consequence of the above proof, is that all but finitely many maximal increasing s_3-sequences of length greater than 2 have as least element a type 1 number, and have as second last element either a type 1 or a type 8 number, with all intermediate numbers type 1. That is, the sequence must be of the form*

$$\text{type 1} \to \text{type 1} \to \cdots \to \text{type 1} \to m,$$

or of the form

$$\text{type 1} \to \text{type 1} \to \cdots \to \text{type 1} \to \text{type 8} \to m,$$

where m satisfies $s_3(m) < m$.

We have accumulated enough information to prove the non-existence of the simplest case of s_3-cycles.

Theorem 4.9. *There are no s_3-cycles of length 2. That is, if $s_3^{(2)}(n) = n$, then $s_3(n) = n$.*

Proof. Any such s_3-cycle must have a least element n of type 1, 2, 7, or 10. First assume that $n = 4pq$ is type 1. Because of the computer searches, we may assume that the pair (p, q) lies outside of the ranges $2 \leq p \leq 30$, $2 \leq q \leq 1008$, and $2 \leq p \leq 100$, $2 \leq q \leq 100$, and that $p \leq q$.

Let $m = pq + p + q$. In the allowable range, the minimum value for m occurs when $p = 2$, and $q = 1009$, which gives $m = 3029$. Note that $2 \nmid m$.

We have $s_3^{(2)}(4pq) = s_3(4m) = 4s_1(m) + 4s_2(m) + s_3(m)$. If $s_3^{(2)}(n) = n$, then

$$4s_1(m) + 4s_2(m) + s_3(m) = 4pq. \tag{4.2}$$

We will take cases depending on $\Omega(m)$. In the following, q_i always indicates a prime dividing m. If $m = q_1$ is prime, then we are done since then $s_2(4m) = 4m$.

If $m = q_1q_2$, then (4.2) implies

$$q_1 + q_2 + q_1q_2 = pq.$$

Substituting $q_1q_2 = pq + p + q$, and canceling gives $q_1 + q_2 + p + q = 0$, an impossibility.

If $m = q_1q_2q_3$, then (4.2) implies that

$$4s_1(m) + 4s_2(m) + m = 4pq.$$

This is impossible, since $2 \nmid m$.

In case $m = q_1 \cdots q_4$, we will show that (4.2) should be replaced with a strict inequality. Dividing (4.2) through by m gives

$$4\left(\frac{1}{q_1q_2q_3} + \cdots + \frac{1}{q_2q_3q_4}\right) + 4\left(\frac{1}{q_1q_2} + \cdots + \frac{1}{q_3q_4}\right) + \left(\frac{1}{q_1} + \cdots + \frac{1}{q_4}\right) = \frac{4pq}{pq + p + q}.$$

We can minimize the right hand side of this equation. In the range in question, $4pq/(pq+p+q)$ is minimized when $p = 2, q = 1009$, and this gives us that $4pq/(pq + p + q) > 2.6649$. It is an easy check that if $q_i \geq 3$, and $m = q_1 \cdots q_4 \geq 3029$, then we always have

$$4\left(\frac{1}{q_1q_2q_3} + \cdots + \frac{1}{q_2q_3q_4}\right) + 4\left(\frac{1}{q_1q_2} + \cdots + \frac{1}{q_3q_4}\right) + \left(\frac{1}{q_1} + \cdots + \frac{1}{q_4}\right) < 2.6649 < \frac{4pq}{pq + p + q}.$$

A similar approach can be used when $\Omega(m) = 5$. If $r \geq 6$, then

$$\begin{aligned} 4\left(\frac{1}{q_1q_2q_3} + \cdots + \frac{1}{q_{r-2}q_{r-1}q_r}\right) &+ 4\left(\frac{1}{q_1q_2} + \cdots + \frac{1}{q_{r-1}q_r}\right) + \left(\frac{1}{q_1} + \cdots + \frac{1}{q_r}\right) \\ &< \binom{r}{1}\frac{4}{3^{r-1}} + \binom{r}{2}\frac{4}{3^{r-2}} + \binom{r}{3}\frac{1}{3^{r-3}} \\ &< 2.6649 < \frac{4pq}{pq + p + q}, \end{aligned}$$

and we are done the type 1 case.

Next suppose that the least element $n = 16p$ is type 2. We may assume that $p \geq 1009$. Now $s_3^{(2)}(16p) = s_3(8(3p+4))$. Let $m = 3p+4$, so $16p = 16(m-4)/3$. Then $2, 3 \nmid m$, and $m \geq 3031$. So $s_3^{(2)}(16p) = 8 + 12s_1(m) + 6s_2(m) + s_3(m)$. The supposition that $s_3^{(2)}(n) = n$ is equivalent to

$$8 + 12s_1(m) + 6s_2(m) + s_3(m) = 16\left(\frac{m-4}{3}\right),$$

which is equivalent to

$$88 + 36s_1(m) + 18s_2(m) + 3s_3(m) = 16m, \tag{4.3}$$

We will consider cases depending on the value of $\Omega(m)$. In the following, each q_i is assumed to be prime. If $m = q_1$ is prime, then (4.3) becomes $88 + 36q_1 = 16q_1$, which is impossible.

If $m = q_1q_2$, then (4.3) is also impossible for m in the allowable range. If $m = q_1q_2q_3$, then (4.3) is equivalent to

$$88 + 36(q_1 + q_2 + q_3) + 18(q_1q_2 + q_1q_3 + q_2q_3) = 13q_1q_2q_3.$$

Dividing through by m, and using the facts that $m \geq 3031$, and $q_i \geq 5$, we have the inequality

$$\frac{88}{3031} + 36\left(\frac{1}{q_1q_2} + \frac{1}{q_1q_3} + \frac{1}{q_2q_3}\right) + 18\left(\frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{q_3}\right) < 13.$$

This contradicts (4.3). A similar contradiction occurs when $\Omega(m) > 3$.

The proofs of the cases when n is type 7 or 10 follow the same lines as the proof for case 2, and are omitted. $\square$

Chapter 5

The Power Prime Symmetric Divisor Functions

We now begin our study of the power prime symmetric functions. Recall we were able to classify when a prime power p^α was s_k^*-perfect for some k. We now present an analogous result pertaining to e_k-perfection.

Theorem 5.1. *Let $n = p^\alpha$ be a prime power. Then there exists a $k \in \mathbb{N}$ such that n is e_k-perfect if and only if $k = p^\beta - \beta$ and $\alpha = p^\beta$ for some $\beta \geq 0$.*

Proof. $e_k(p^\alpha) = p^\alpha$ if and only if $\alpha p^k = p^\alpha$. Hence we may write $\alpha = p^\beta$ for some $\beta \geq 0$. This implies that $p^{\beta+k} = p^{p^\beta}$, and so $k = p^\beta - \beta$.

Conversely, it is trivial to see that p^{p^β} is always $e_{p^\beta-\beta}$-perfect. □

Remark 5.2. *The set $\{p^\beta - \beta | p \text{ is prime and } \beta \geq 0\}$ has asymptotic density 0 in $\mathbb{N}$. So, for "most" values of k, there is no prime power p^α that is e_k-perfect.*

With the functions e_k, we can also characterize which numbers are e_k-perfect for some k for products of two prime powers.

Theorem 5.3. *Let $n = p^\alpha q^\beta$, where p and q are primes, and $\alpha, \beta > 0$. Then n is not e_k-perfect for any k.*

Proof. Without loss of generality, $p < q$. Furthermore, since $e_1 = s_1$, and the statement has been proved for s_1, we may assume that $k \geq 2$. If $e_k(p^\alpha q^\beta) = p^\alpha q^\beta$, then

$$\alpha p^k + \beta q^k = p^\alpha q^\beta. \tag{5.1}$$

First assume that $\alpha, \beta \geq k$. Then since

$$\beta q^k = p^k (p^{\alpha-k} q^\beta - \alpha),$$

we have that p^k divides β. Similarly q^k divides α. Write $\beta = bp^k$, and $\alpha = aq^k$. Clearly $a, b > 0$, and equation (5.1) becomes

$$(a+b)p^k q^k = p^{aq^k} q^{bp^k}.$$

We will show that we instead have a strict inequality

$$(a+b)p^k q^k < p^{aq^k} q^{bp^k}. \tag{5.2}$$

Since $a, b > 0$, $p \geq 2$, and $q \geq 3$, it is easily seen that $bp^{k-1} \leq bp^k - k$, and $aq^{k-1} \leq aq^k - k$, and hence the inequality (5.2) is implied by

$$a + b < p^{aq^{k-1}} q^{bp^{k-1}}.$$

This is in turn implied by

$$a + b < ab(pq)^{k-1}$$

Now $pq \geq 6$, and $k \geq 2$, so the latter inequality holds if $a + b < 6ab$. This is indeed true for $a, b > 0$.

The second case we consider is when $\alpha, \beta < k$. Then in similar fashion to the previous case, we can write $\beta = bp^\alpha$, and $\alpha = aq^\beta$, where $a, b > 0$. This however gives an immediate contradiction, since then $\alpha = aq^\beta = aq^{bp^\alpha} > \alpha$.

Next suppose that $\alpha < k$, and $\beta \geq k$. Then for some $a > 0$, we have that

$$\alpha = aq^k > aq^\alpha > \alpha.$$

This is a contradiction. A similar contradiction results in the remaining case, thus, the theorem is proved. □

We can relate the functions s_k and e_k using the Newton-Girard formulas (see [44]). These formulas imply that

$$(-1)^k k s_k + \sum_{i=1}^{k} (-1)^{i+k} e_i s_{k-i} = 0.$$

The function e_k can be expressed strictly in terms of the elementary prime symmetric functions $s_1, \ldots, s_k$ as a determinant (see also [45]):

$$e_k = (-1)^k \begin{vmatrix} s_1 & 1 & 0 & 0 & \cdots & 0 \\ 2s_2 & s_1 & 1 & 0 & \cdots & 0 \\ 3s_3 & s_2 & s_1 & 1 & \cdots & 0 \\ 4s_4 & s_3 & s_2 & s_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ ks_k & s_{k-1} & s_{k-2} & s_{k-3} & \cdots & s_1 \end{vmatrix}$$

Similarly, we can express s_k in terms of e_k as follows:

$$s_k = \begin{vmatrix} e_1 & 1 & 0 & 0 & \cdots & 0 \\ \frac{e_2}{2} & \frac{e_1}{2} & 1 & 0 & \cdots & 0 \\ \frac{e_3}{3} & \frac{e_2}{3} & \frac{e_1}{3} & 1 & \cdots & 0 \\ \frac{e_4}{4} & \frac{e_3}{4} & \frac{e_2}{4} & \frac{e_1}{4} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{e_{k-1}}{k-1} & \frac{e_{k-2}}{k-1} & \frac{e_{k-3}}{k-1} & \frac{e_{k-4}}{k-1} & \cdots & 1 \\ \frac{e_k}{k} & \frac{e_{k-1}}{k} & \frac{e_{k-2}}{k} & \frac{e_{k-3}}{k} & \cdots & \frac{e_1}{k} \end{vmatrix}$$

5.1 The Average Order of e_k

Analysis in Sections 5.1 and 5.2 and Chapter 6 are taken from the paper [50] of the author, previously published in Integers: The Electronic Journal of Combinatorial Number Theory.

In this section we shall prove the following asymptotic formula for the average order of $e_k(n)$:

$$\sum_{n\le x} e_k(n) = \frac{\zeta(k+1)x^{k+1}}{(k+1)\log x} + O\left(\frac{x^{k+1}\log\log x}{\log^2 x}\right).$$

The asymptotic alone (without the error term included) has been proved by Kerawala [24] and LeVan [27] each using different techniques. The case when $k = 1$ was treated by Alladi and Erdős [1]. We shall make precise LeVan's sketch of the proof. First we require the following Lemma:

Lemma 5.4. *For $x \ge 2$, and $k, \ell \in \mathbb{N}_0$ we have*

$$\sum_{p\le x} \frac{p^k}{\log^\ell p} = \frac{x^{k+1}}{(k+1)\log^{\ell+1} x} + O\left(\frac{x^{k+1}}{\log^{\ell+2} x}\right).$$

Proof. The sum may be expressed as a Riemann-Stieltjes integral:

$$\sum_{p\le x} \frac{p^k}{\log^\ell p} = \int_{3/2}^{x} \frac{t^k}{\log^\ell t}\, d\pi(t).$$

Integrating by parts we have:

$$\int_{3/2}^{x} \frac{t^k}{\log^\ell t}\, d\pi(t) = \frac{x^k\pi(x)}{\log^\ell x} - \int_{3/2}^{x} \left(\frac{kt^{k-1}}{\log^\ell t} - \frac{\ell t^{k-1}}{\log^{\ell+1} t}\right)\pi(t)\, dt.$$

By the prime number theorem, this gives

$$\begin{aligned}
\int_{3/2}^{x} \frac{t^k}{\log^\ell t}\, d\pi(t) =& \frac{x^k}{\log^\ell x}\left(\frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right)\right) \\
& - \int_{3/2}^{x} \left(\frac{kt^{k-1}}{\log^\ell t} - \frac{\ell t^{k-1}}{\log^{\ell+1} t}\right)\left(\frac{t}{\log t} + O\left(\frac{t}{\log^2 t}\right)\right) dt \\
=& \frac{x^{k+1}}{\log^{\ell+1} x} + O\left(\frac{x^{k+1}}{\log^{\ell+2} x}\right) \\
& - k\int_{3/2}^{x} \frac{t^k}{\log^{\ell+1} t}\, dt + O\left(\int_{3/2}^{x} \frac{t^k}{\log^{\ell+2} t}\, dt\right) \\
=& \frac{x^{k+1}}{\log^{\ell+1} x} - k\left[\frac{x^{k+1}}{(k+1)\log^{\ell+1} x} + O(1) + \frac{\ell+1}{k+1}\int_{3/2}^{x} \frac{t^k}{\log^{\ell+2} t}\, dt\right] \\
& + O\left(\frac{x^{k+1}}{\log^{\ell+2} x}\right) \\
=& \frac{x^{k+1}}{(k+1)\log^{\ell+1} x} + O\left(\frac{x^{k+1}}{\log^{\ell+2} x}\right),
\end{aligned}$$

and the proof is complete. □

Theorem 5.5. *For $k \in \mathbb{N}$,*

$$\sum_{n \le x} e_k(n) = \frac{\zeta(k+1)x^{k+1}}{(k+1)\log x} + O\left(\frac{x^{k+1}\log\log x}{\log^2 x}\right).$$

Proof. We have

$$\begin{aligned}
\sum_{n \le x} e_k(n) &= \sum_{p \le x}\sum_{i=1}^{\infty} p^k \left\lfloor \frac{x}{p^i} \right\rfloor \\
&= \sum_{p \le x} p^k \left\lfloor \frac{x}{p} \right\rfloor + \sum_{i=2}^{\infty}\sum_{p \le x^{1/i}} p^k \left\lfloor \frac{x}{p^i} \right\rfloor. \qquad (5.3)
\end{aligned}$$

As we shall see, the first term contributes the greater portion to the sum:

$$\sum_{p\le x} p^k \left\lfloor \frac{x}{p} \right\rfloor = \sum_{i\le x/2} \sum_{\frac{x}{i+1}<p\le \frac{x}{i}} ip^k$$
$$= \sum_{i\le x/2} \sum_{p\le x/i} p^k$$
$$= \sum_{i\le x/2} \left(\frac{(x/i)^{k+1}}{(k+1)\log{(x/i)}} + O\left(\frac{(x/i)^{k+1}}{\log^2{(x/i)}} \right) \right). \tag{5.4}$$

Let

$$\Sigma_1 = \sum_{i\le \log^2 x} \frac{1}{i^{k+1}\log{(x/i)}},$$

and

$$\Sigma_2 = \sum_{\log^2 x<i\le x/2} \frac{1}{i^{k+1}\log{(x/i)}},$$

so that

$$\sum_{i\le x/2} \frac{1}{i^{k+1}\log{(x/i)}} = \Sigma_1 + \Sigma_2.$$

We have that

$$\Sigma_1 \ge \frac{1}{\log x} \sum_{i\le \log^2 x} \frac{1}{i^{k+1}}$$
$$= \frac{1}{\log x} \left(\zeta(k+1) - \sum_{i>\log^2 x} \frac{1}{i^{k+1}} \right)$$
$$= \frac{1}{\log x} \left(\zeta(k+1) + O\left(\frac{1}{\log^{2k} x} \right) \right)$$
$$= \frac{\zeta(k+1)}{\log x} + O\left(\frac{1}{\log^{2k+1} x} \right).$$

On the other hand,

$$\begin{aligned}
\Sigma_1 &\leq \sum_{i \leq \log^2 x} \frac{1}{i^{k+1}(\log x - 2\log\log x)} \\
&= \frac{1}{\log x}\left(1 + O\left(\frac{\log\log x}{\log x}\right)\right) \sum_{i \leq \log^2 x} \frac{1}{i^{k+1}} \\
&= \frac{1}{\log x}\left(1 + O\left(\frac{\log\log x}{\log x}\right)\right)\left(\zeta(k+1) + O\left(\frac{1}{\log^{2k} x}\right)\right) \\
&= \frac{\zeta(k+1)}{\log x} + O\left(\frac{\log\log x}{\log^2 x}\right).
\end{aligned}$$

Combining these results we have that

$$\Sigma_1 = \frac{\zeta(k+1)}{\log x} + O\left(\frac{\log\log x}{\log^2 x}\right). \tag{5.5}$$

The sum Σ_2 is negligible by comparison:

$$\begin{aligned}
\Sigma_2 &= \sum_{\log^2 x < i \leq x/2} \frac{1}{i^{k+1}\log(x/i)} \\
&\ll \sum_{i > \log^2 x} \frac{1}{i^{k+1}} \\
&\ll \int_{\log^2 x}^{\infty} \frac{1}{t^{k+1}}\, dt \\
&\ll \frac{1}{\log^{2k} x}.
\end{aligned} \tag{5.6}$$

Now we need to bound the error term in (5.4):

$$\begin{aligned}
\sum_{i \leq x/2} \frac{1}{i^{k+1}\log^2(x/i)} &= O\left(\int_1^{x/2} \frac{dt}{t^2 \log^2 x/t}\right) \\
&= O\left(\int_1^{\sqrt{x}} \frac{dt}{t^2\log^2 x/t} + \int_{\sqrt{x}}^{x/2} \frac{dt}{t^2\log^2 x/t}\right) \\
&= O\left(\frac{4}{\log^2 x}\int_1^{\sqrt{x}} \frac{dt}{t^2} + \int_{\sqrt{x}}^{x/2} \frac{dt}{t^2}\right) \\
&= O\left(\frac{1}{\log^2 x}\right)
\end{aligned} \tag{5.7}$$

Hence by (5.5), (5.6), and (5.7), we have that

$$\sum_{p\le x} p^k \left\lfloor \frac{x}{p} \right\rfloor = \frac{\zeta(k+1)x^{k+1}}{(k+1)\log x} + O\left(\frac{x^{k+1}\log\log x}{\log^2 x}\right).$$

To conclude the proof, we need to bound the second term in (5.3).

$$\begin{aligned}
\sum_{i=2}^{\infty}\sum_{p\le x^{1/i}} p^k \left\lfloor \frac{x}{p^i} \right\rfloor &\le x \sum_{i=2}^{\infty}\sum_{p\le x^{1/i}} \frac{p^k}{p^i} \\
&\le x \sum_{p\le\sqrt{x}} \frac{p^{k-1}}{p-1} \\
&\le 2x \sum_{p\le\sqrt{x}} p^{k-2} \\
&= \begin{cases} O(x\log\log x), & \text{if } k=1; \\ O\left(\frac{x^{\frac{k+1}{2}}}{\log x}\right), & \text{if } k>1. \end{cases}
\end{aligned}$$

Note that for the $k=1$ case,

$$x\log\log x = O\left(\frac{x^2\log\log x}{\log^2 x}\right),$$

and for the $k>1$ case,

$$\frac{x^{\frac{k+1}{2}}}{\log x} = O\left(\frac{x^{k+1}\log\log x}{\log^2 x}\right).$$

Hence, these error terms may be absorbed to yield the theorem. □

The case when $k=0$, i.e. the average order of $\Omega(n)$, will be covered in the next section.

5.2 A Statistical Look at e_k

In the last section we found the average order of $e_k(n)$. In this section we will take a statistical look at its range of values. First, we need some notation and context for the functions $r_{1,k}$.

We have looked at the functions r_k, which counted the number of solutions to the equation $s_k(m) = n$ for a fixed n. The analogue for the functions e_k is none other than the function which counts the number of partitions into k-th powers of primes.

Definition 5.6. *Let $A \subseteq \mathbb{N}$. The number of partitions of n into parts taken from A is denoted by $p_A(n)$.*

Note that $p_A(0) = 1$, corresponding to the empty partition. Writing the set of k-th powers of primes as $\mathbb{P}^{(k)}$, we have

$$r_{1,k}(n) = p_{\mathbb{P}^{(k)}}(n) = |e_k^{(-1)}\{n\}| \tag{5.8}$$

It is much easier to work with the functions $p_{\mathbb{P}^{(k)}}(n)$, than with the analogues $r_k(n)$ with the elementary prime symmetric functions. For instance, it is a trivial matter to see that $p_{\mathbb{P}^{(k)}}(n) \to \infty$ with n:

Theorem 5.7. *For $k \in \mathbb{N}$, we have that $\lim_{n\to\infty} p_{\mathbb{P}^{(k)}}(n) = \infty$.*

Proof. Let $\ell \in \{0, 1, \ldots, 2^k - 1\}$, and consider the sequence $\{n2^k + \ell\}_{n=1}^{\infty}$. Let q be any prime congruent to 1 modulo 2^k. Numbers of the form $2^n q^\ell$, satisfy $e_k(2^n q^\ell) = n2^k + \ell q^k$, and $\ell q^k \equiv \ell \pmod{2^k}$. Since there are infinitely many such q, we have that $\lim_{n\to\infty} p_{\mathbb{P}^{(k)}}(n2^k + \ell) = \infty$. Since this holds for any $\ell \in \{0, 1, \ldots, 2^k - 1\}$, we have the stated result. □

From the theory of partitions (see for instance [2]), we can construct generating functions useful in calculating $p_{\mathbb{P}^{(k)}}(n)$. Indeed, we have

$$\prod_p \frac{1}{1 - x^{p^k}} = \sum_{n=0}^{\infty} p_{\mathbb{P}^{(k)}}(n) x^n.$$

The proof of this fact is by analogy with [2] pp.308-310. Generating functions provide easy expressions for the j-th difference functions for $p_{\mathbb{P}^{(k)}}$, which we write as $p_{\mathbb{P}^{(k)}}^{(j)}$:

$$\sum_{n=0}^{\infty} p_{\mathbb{P}^{(k)}}^{(j)}(n) x^n = (1-x)^j \sum_{n=0}^{\infty} p_{\mathbb{P}^{(k)}}(n) x^n = (1-x)^j \prod_p \frac{1}{1 - x^{p^k}}. \tag{5.9}$$

In particular we have that

$$p_{\mathbb{P}^{(k)}}^{(-1)}(n) = \sum_{m=0}^{n} p_{\mathbb{P}^{(k)}}(m).$$

The following trivial identity turns out to be rather useful in deriving some statistical properties of the exponent prime symmetric functions. We denote by $P(n)$ the largest prime factor dividing n. Then:

$$P(n)^k \leq e_k(n) \leq P(n)^k \Omega(n). \tag{5.10}$$

Definition 5.8. *Let*

$$b_k(x,y) = \#\{n \le x : e_k(n) \le y\}, \tag{5.11}$$

$$\Psi(x,y) = \#\{n \le x : P(n) \le y\}, \tag{5.12}$$

Erdős and Pomerance [16] show that for every $\epsilon > 0$ there is a $\delta > 0$ such that for x sufficiently large there is at least $(1-\epsilon)x$ choices for $n \le x$ such that $P(n) < e_1(n) < (1 + x^{-\delta})P(n)$.

We will use our comparison 5.10 to relate the quantities $b_k(x,y)$ and $\Psi(x,y)$, but first we require some background material from analytic number theory. The proof of the first lemma that follows can be found in [33] pp. 278-279.

Lemma 5.9. *There is a constant b_1 such that*

$$\sum_{p \le x} \frac{1}{p} = \log\log x + b_1 + O\left(\frac{1}{\log x}\right).$$

Lemma 5.10. *There is a constant b_2 such that*

$$\sum_{n \le x} \Omega(n) = x \log\log x + b_2 x + O\left(\frac{x}{\log x}\right).$$

Proof. By the additivity of $\Omega(n)$, and de Polignac's formula,

$$\begin{aligned}
\sum_{n \le x} \Omega(n) &= \sum_{p \le x} \sum_{j=1}^{\infty} \left\lfloor \frac{x}{p^j} \right\rfloor \\
&= x \sum_{p \le x} \sum_{j=1}^{\infty} \left(\frac{1}{p^j}\right) - \sum_{p \le x} \sum_{j=1}^{\infty} \left\{\frac{x}{p^j}\right\} \\
&= x \sum_{p \le x} \frac{1}{p-1} + O\left(\sum_{p \le x} \frac{\log x}{\log p}\right) \\
&= x \sum_{p \le x} \left(\frac{1}{p} + \frac{1}{p(p-1)}\right) + O\left(\frac{x}{\log x}\right).
\end{aligned}$$

The last step followed from Lemma 5.4. Hence by Lemma 5.9,

$$\begin{aligned}
\sum_{n \le x} \Omega(n) &= x\left(\log\log x + b_1 + O\left(\frac{1}{\log x}\right)\right) + x \sum_{p \in \mathbb{P}} \frac{1}{p(p-1)} + O\left(\frac{x}{\log x}\right) \\
&= x \log\log x + b_2 x + O\left(\frac{x}{\log x}\right),
\end{aligned}$$

where

$$b_2 = b_1 + \sum_{p \in \mathbb{P}} \frac{1}{p(p-1)}. \tag{5.13}$$

□

A similar result holds for $\omega(n)$, and can be found in [33] p. 283. We also need to compute the average order for $\Omega(n)^2$. Doing so will constitute the content of the next lemma.

Lemma 5.11. *The average order of $\Omega(n)^2$ is given by*

$$\sum_{n \le x} \Omega(n)^2 = x(\log\log x)^2 + O(x \log\log x). \tag{5.14}$$

Proof.

$$\begin{aligned}
\sum_{n \le x} \Omega(n)^2 &= \sum_{n \le x} \left(\sum_{p^\ell | n} 1 \right)^2 \\
&= \sum_{n \le x} \left(\sum_{p^\ell | n} 1 \right) \left(\sum_{q^m | n} 1 \right) \\
&= \sum_{p^\ell \le x} \sum_{q^m \le x} \sum_{\substack{n \le x \\ p^\ell, q^m | n}} 1 \\
&= \sum_{p \le x} \sum_{\ell, m \le \log x / \log p} \sum_{\substack{n \le x \\ p^{\max\{\ell, m\}} | n}} 1 + \sum_{\substack{p^\ell, q^m \le x \\ p \ne q}} \left\lfloor \frac{x}{p^\ell q^m} \right\rfloor \\
&= \Sigma_1 + \Sigma_2,
\end{aligned}$$

where the obvious designations are made for Σ_1, and Σ_2. We shall first show that the triple sum Σ_1 is $O(x \log\log x)$. Indeed, letting $R_p = \log x / \log p$, we have

$$\begin{aligned}
\Sigma_1 &= \sum_{p \le x} \sum_{\ell, m \le R_p} \left\lfloor \frac{x}{p^{\max\{\ell,m\}}} \right\rfloor \\
&= x \sum_{p \le x} \sum_{\ell, m \le R_p} \frac{1}{p^{\max\{\ell,m\}}} + O\left(\sum_{p \le x} \left\lfloor \frac{\log x}{\log p} \right\rfloor^2 \right) \\
&= x \sum_{p \le x} \sum_{k \le R_p} \frac{2k-1}{p^k} + O\left(\log^2 x \sum_{p \le x} \frac{1}{\log^2 p} \right) \\
&= O\left(x \sum_{p \le x} \sum_{k \le R_p} \frac{k}{p^k} \right) + O\left(\frac{x}{\log x} \right), \text{ using Lemma 5.4} \\
&= O\left(x \sum_{p \le x} \left[\frac{1}{p} \left(\frac{1 - 1/p^{\lfloor R \rfloor}}{(1-1/p)^2} \right) - \frac{\lfloor R \rfloor}{p^{\lfloor R \rfloor + 1}(1-1/p)} \right] \right) + O\left(\frac{x}{\log x} \right) \\
&= O\left(x \sum_{p \le x} \left[\frac{p}{(p-1)^2} \left(1 - \frac{1}{x} \right) - \frac{\log x}{(p-1) x \log p} \right] \right) + O\left(\frac{x}{\log x} \right) \\
&= O\left(x \sum_{p \le x} \frac{1}{p} \right) + O\left(\frac{x}{\log x} \right) \\
&= O(x \log \log x).
\end{aligned}$$

We now turn our attention to Σ_2:

$$\begin{aligned}
\Sigma_2 &= x \sum_{\substack{p^\ell, q^m \le x \\ p \ne q}} \frac{1}{p^\ell q^m} + O\left(\sum_{\substack{p^\ell, q^m \le x \\ p \ne q}} 1 \right) \\
&= x \sum_{\substack{p, q \le x \\ p \ne q}} \frac{1}{pq} + x \sum_{\substack{p^\ell, q^m \le x \\ p \ne q \\ \max\{\ell, m\} \ge 2}} \frac{1}{p^\ell q^m} + O\left(\frac{x}{\log x} \right).
\end{aligned}$$

But

$$\sum_{\substack{p,q \le x \\ p \ne q}} \frac{1}{pq} = \sum_{q \le x} \frac{1}{q} \sum_{\substack{p \le x \\ p \ne q}} \frac{1}{p}$$
$$= \sum_{q \le x} \frac{1}{q} (\log\log x + O(1))$$
$$= (\log\log x)^2 + O(\log\log x),$$

and

$$\sum_{\substack{p^\ell, q^m \le x \\ p \ne q \\ \max\{\ell, m\} \ge 2}} \frac{1}{p^\ell q^m} = O\left(\sum_{\substack{p^\ell \le x \\ \ell \ge 2}} \frac{1}{p^\ell} \left[\sum_{q \le x} \frac{1}{q} + O(1) \right] \right) = O(\log\log x). \quad (5.15)$$

Hence

$$\Sigma_2 = x(\log\log x)^2 + O(x \log\log x), \quad (5.16)$$

and therefore

$$\sum_{n \le x} \Omega(n)^2 = x(\log\log x)^2 + O(x \log\log x) \quad (5.17)$$

as claimed. □

The function $\Omega(n)$ is "close" to $\log\log n$ for almost all values of n. This statement is made more precise by the following theorem and its subsequent remark, for which we have Turán [43] to thank:

Theorem 5.12. *The number of $n \le x$ such that*

$$|\Omega(n) - \log\log x| > (\log\log x)^{3/4}$$

is $O\left(\frac{x}{\sqrt{\log\log x}}\right)$.

Proof. Denote by S the set $\{n \in \mathbb{N} : |\Omega(n) - \log\log x| > (\log\log x)^{3/4}\}$, and let $S(x) = \#\{n \leq x : n \in S\}$, be the counting function of S. Then

$$\begin{aligned}
S(x) &= \sum_{n \leq x,\ n \in S} 1 \\
&\leq \sum_{n \leq x} \frac{(\Omega(n) - \log\log x)^2}{(\log\log x)^{3/2}} \\
&\ll (\log\log x)^{-3/2} \left(\sum_{n \leq x} \Omega(n)^2 - 2\log\log x \sum_{n \leq x} \Omega(n) + x(\log\log x)^2 \right) \\
&\ll (\log\log x)^{-3/2} (x(\log\log x)^2 + O(x \log\log x) \\
&- 2\log\log x(x \log\log x + O(x)) + x(\log\log x)^2) \\
&= O\left(\frac{x}{\sqrt{\log\log x}} \right).
\end{aligned}$$

□

Remark 5.13. *It is also true that*

$$\#\{n \leq x : |\Omega(n) - \log\log n| > (\log\log n)^{3/4}\} = O\left(\frac{x}{\sqrt{\log\log x}} \right).$$

The proof of this fact is similar to that of theorem 5.12.

We are now equipped to relate the quantities $b_k(x, y)$, and $\Psi(x, y)$.

Theorem 5.14. *There is an absolute positive constant c_1 such that the inequalities*

$$\Psi\left(x, \left(\frac{y}{\log\log x + (\log\log x)^{3/4}} \right)^{1/k} \right) - \frac{c_1 x}{\sqrt{\log\log x}} \leq b_k(x, y) \leq \Psi(x, y^{1/k}) \tag{5.18}$$

hold.

Proof. By equation (5.10), we have that

$$\#\{n \leq x : P(n)^k \Omega(n) \leq y\} \leq b_k(x, y) \leq \#\{n \leq x : P(n)^k \leq y\}.$$

This implies the second inequality. By remark 5.13, there is an absolute positive constant c_1 such that

$$\#\{n \leq x : |\Omega(n) - \log\log n| \leq (\log\log n)^{3/4}\} \geq x\left(1 - \frac{c_1}{\sqrt{\log\log x}} \right).$$

Given this, we have the following chain of inequalities:

$$\begin{aligned}
\#\{n \le x : P(n)^k\Omega(n) \le y\} \ge& \#\{n \le x : P(n)^k\Omega(n) \le y, \text{ and} \\
&\Omega(n) \le \log\log n + (\log\log n)^{3/4}\} \\
\ge& \#[\{n \le x : P(n)^k(\log\log n + (\log\log n)^{3/4}) \le y\} \\
&\cap \{n \le x : \Omega(n) \le \log\log n + (\log\log n)^{3/4}\}] \\
\ge& \#\{n \le x : P(n)^k(\log\log x + (\log\log x)^{3/4}) \le y\} \\
&+ \#\{n \le x : |\Omega(n) - \log\log n| \le (\log\log n)^{3/4}\} - x \\
\ge& \Psi\left(x, \left(\frac{y}{\log\log x + (\log\log x)^{3/4}}\right)^{1/k}\right) - \frac{c_1 x}{\sqrt{\log\log x}},
\end{aligned}$$

and the theorem is proved. □

There is an interesting connection between $b_k(x, y)$, and $p_{\mathbb{P}(k)}(n)$. For x sufficiently large relative to n,

$$b_k(x, n) - b_k(x, n-1) = p_{\mathbb{P}(k)}(n).$$

A telescoping sum yields

$$b_k(x, n) = \sum_{m=0}^{n} p_{\mathbb{P}(k)}(n) = p_{\mathbb{P}(k)}^{(-1)}(n).$$

Unfortunately, for x in this range, the best estimates for $\Psi(x, n)$ give error terms that are too large to be useful. However, we will see that for x in ranges such as $y^{1/\alpha}$, where $0 < \alpha < 1$, Theorem 5.14 is useful in describing the behaviour of $b_k(x, y)$.

The "Dickman function" $\rho(u)$ is defined to be the unique continuous solution to the differential-difference equation

$$u\rho'(u) = -\rho(u-1) \quad (u > 1),$$

satisfying the initial condition

$$\rho(u) = 1 \quad (0 \le u \le 1).$$

The Dickman function is nonnegative for $u > 0$, and decreasing for $u > 1$. This definition and description is taken from [22], which is a comprehensive survey of work on the function $\Psi(x, y)$.

It is also true that $\rho(u)$ is convex on $(1, \infty)$. By the functional equation, it is apparent that

$$\lim_{u \to 1^+} \rho'(u) = -1.$$

It follows that

$$\rho(a+b) \geq \rho(a) - b, \text{ for } a \geq 1,\ b \geq 0. \tag{5.19}$$

De Bruijin [9] proved that

$$\Psi(x, y) = x\rho(u) \left[1 + O\left(\frac{\log(u+1)}{\log y}\right)\right] \tag{5.20}$$

holds uniformly for $u = \log x / \log y$ in the range

$$y \geq 2,\ 1 \leq u \leq (\log y)^{3/5-\epsilon}. \tag{5.21}$$

This has since been improved by Hildebrand [21] to the range

$$y \geq 2,\ 1 \leq u \leq \exp\Big((\log y)^{3/5-\epsilon}\Big). \tag{5.22}$$

For the next theorem, we are interested in the range $y = x^{\alpha/k}$, and hence $u = k/\alpha$.

Theorem 5.15. *Let $0 < \alpha < 1$, and let $k \in \mathbb{N}$ be fixed constants. Then*

$$b_k(x, x^\alpha) \sim \Psi(x, x^{\alpha/k}) \sim \rho\left(\frac{k}{\alpha}\right) x.$$

In fact, we have

$$b_k(x, x^\alpha) = \rho\left(\frac{k}{\alpha}\right) x + O\left(\frac{x}{\sqrt{\log \log x}}\right).$$

Proof. Note that for $x \geq 2^{k/\alpha}$, we are within the range (5.22). Consequently (5.20) applies:

$$\Psi(x, x^{\alpha/k}) = x\rho\left(\frac{k}{\alpha}\right)\left[1 + O\left(\frac{1}{\log x}\right)\right].$$

Furthermore, the second inequality of Theorem 5.14 gives us:

$$b_k(x, x^\alpha) \leq \Psi(x, x^{\alpha/k}).$$

Similarly, the first inequality from theorem 5.14 yields:

$$\Psi\left(x, \left(\frac{x^{\alpha}}{2\log\log x}\right)^{1/k}\right) - \frac{c_1 x}{\sqrt{\log\log x}} \leq b_k(x, x^{\alpha}),$$

since $\Psi(x, y)$ is increasing in y.

For x sufficiently large, (5.20) gives

$$\Psi\left(x, \left(\frac{x^{\alpha}}{2\log\log x}\right)^{1/k}\right) = x\rho\left(\frac{\log x}{(\alpha/k)\log x - (1/k)\log(2\log\log x)}\right) \times \left[1 + O\left(\frac{\log(u+1)}{\log y}\right)\right],$$

where in this instance,

$$y = \left(\frac{x^{\alpha}}{2\log\log x}\right)^{1/k}$$

and

$$\begin{aligned} u = \frac{\log x}{\log y} &= \frac{\log x}{(\alpha/k)\log x - (1/k)\log(2\log\log x)} \\ &= \left(\frac{k}{\alpha}\right)\frac{1}{1 - (1/\alpha)\log(2\log\log x)/\log x}. \end{aligned}$$

The identity

$$\frac{1}{1-a} < 1 + 2a$$

holds for $0 < a < 1/2$. So if x is chosen sufficiently large such that

$$0 < \frac{\log(2\log\log x)}{\alpha\log x} < \frac{1}{2},$$

then

$$u < \frac{k}{\alpha} + \frac{2k\log(2\log\log x)}{\alpha^2\log x}.$$

Hence, since $\rho(u)$ decreases on $(1, \infty)$, we have

$$\begin{aligned}
\Psi\left(x, \left(\frac{x^{\alpha}}{2\log\log x}\right)^{1/k}\right) &\geq x\rho\left(\frac{k}{\alpha} + \frac{2k\log(2\log\log x)}{\alpha^2 \log x}\right)\left[1 + O\left(\frac{1}{\log x}\right)\right] \\
&= x\rho\left(\frac{k}{\alpha} + \frac{2k\log(2\log\log x)}{\alpha^2 \log x}\right) + O\left(\frac{x}{\log x}\right) \\
&\geq x\rho\left(\frac{k}{\alpha}\right) - \frac{2kx\log(2\log\log x)}{\alpha^2 \log x} + O\left(\frac{x}{\log x}\right) \\
&= x\rho\left(\frac{k}{\alpha}\right) + O\left(\frac{x\log\log\log x}{\log x}\right),
\end{aligned}$$

using equation (5.19). Combining this information with theorem 5.14 we obtain the following inequalities:

$$\begin{aligned}
x\rho\left(\frac{k}{\alpha}\right) + O\left(\frac{x}{\sqrt{\log\log x}}\right) &\leq \Psi\left(x, \left(\frac{x^{\alpha}}{2\log\log x}\right)^{1/k}\right) \\
&\leq b_k(x, x^{\alpha}) \\
&\leq \Psi(x, x^{\alpha/k}) \\
&= x\rho\left(\frac{k}{\alpha}\right) + O\left(\frac{x}{\log x}\right),
\end{aligned}$$

which proves the theorem. □

Chapter 6

The Average Order of $s_{k,\ell}$

In the last chapter, we looked at the power prime symmetric functions, and were able to compute asymptotic expressions for their average orders. This computation proves to be rather more straightforward than the analogous computation for s_k, primarily due to the fact that e_k is an additive arithmetic function. We will now, however, do even better and compute the average order of $s_{k,\ell}$.

A useful tool in the following computation will be Lemma 5.4. First we need to get a handle on the sum $\sum_{n\le x} s_{k,\ell}(n)$. Throughout this section, k will be a positive integer, and p_i shall always denote a prime, not necessarily the i-th prime. Our main result is Theorem 6.3, which states that

$$\sum_{n\le x} s_{k,\ell}(n) = \frac{\zeta(\ell+1)x^{\ell+1}(\log\log x)^{k-1}}{(\ell+1)(k-1)!\log x} + O\left(\frac{x^{\ell+1}(\log\log x)^{k-2}}{\log x}\right),$$

for $k \ge 2$, and $\ell \ge 1$. The $k = 1$ case was handled in the last chapter. For it, the error term is slightly better.

Theorem 6.1. *The average order of $s_{k,\ell}(n)$ has the following expression:*

$$\sum_{n\le x} s_{k,\ell}(n) = \sum_{r=1}^{k} \sum_{p_1<\ldots<p_r} \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r>0}} (p_1^{i_1}\cdots p_r^{i_r})^\ell \times \left(\sum_{j_1,\ldots,j_r=1}^{\infty} \binom{j_1-1}{i_1-1}\cdots\binom{j_r-1}{i_r-1}\left\lfloor \frac{x}{p_1^{j_1}\cdots p_r^{j_r}}\right\rfloor\right).$$

Proof. The terms in the sum $\sum_{n\le x} s_{k,\ell}(n)$ are products of k ℓ-th powers of primes, not necessarily distinct. In other words, an arbitrary term is of the form $(p_1^{i_1}\cdots p_r^{i_r})^\ell$, where $r \le k$, $p_1 < \ldots < p_r$, $i_1, \ldots, i_r > 0$, and $i_1 + \cdots + i_r = k$. Fix $(p_1^{i_1}\cdots p_r^{i_r})^\ell$. We shall count the number of times this expression occurs in the sum.

By the inclusion-exclusion principle, the number of $n \leq x$ such that $p_1^{j_1}, \ldots, p_r^{j_r} || n$ is

$$\left\lfloor \frac{x}{p_1^{j_1} \cdots p_r^{j_r}} \right\rfloor$$
$$- \left(\left\lfloor \frac{x}{p_1^{j_1+1} p_2^{j_2} \cdots p_r^{j_r}} \right\rfloor + \left\lfloor \frac{x}{p_1^{j_1} p_2^{j_2+1} \cdots p_r^{j_r}} \right\rfloor + \cdots + \left\lfloor \frac{x}{p_1^{j_1} p_2^{j_2} \cdots p_r^{j_r+1}} \right\rfloor \right)$$
$$+ \cdots + (-1)^r \left\lfloor \frac{x}{p_1^{j_1+1} \cdots p_r^{j_r+1}} \right\rfloor,$$

which we write as $\beta(j_1, \ldots, j_r)$. Each such n contributes $\binom{j_1}{i_1} \cdots \binom{j_r}{i_r}$ copies of $(p_1^{i_1} \cdots p_r^{i_r})^\ell$ to the sum $\sum_{n \leq x} s_{k,\ell}(n)$. Thus $(p_1^{i_1} \cdots p_r^{i_r})^\ell$ occurs

$$\sum_{j_1, \ldots, j_r = 1}^{\infty} \binom{j_1}{i_1} \cdots \binom{j_r}{i_r} \beta(j_1, \ldots, j_r)$$

times. We make the following claim:

$$\sum_{j_1, \ldots, j_r = 1}^{\infty} \binom{j_1}{i_1} \cdots \binom{j_r}{i_r} \beta(j_1, \ldots, j_r) = \sum_{j_1, \ldots, j_r = 1}^{\infty} \binom{j_1 - 1}{i_1 - 1} \cdots \binom{j_r - 1}{i_r - 1} \left\lfloor \frac{x}{p_1^{j_1} \cdots p_r^{j_r}} \right\rfloor. \tag{6.1}$$

To prove (6.1), first note that $\left\lfloor \frac{x}{p_1^{j_1} \cdots p_r^{j_r}} \right\rfloor$ occurs

$$\binom{j_1}{i_1} \cdots \binom{j_r}{i_r} - \left(\binom{j_1 - 1}{i_1} \binom{j_2}{i_2} \cdots \binom{j_r}{i_r} + \cdots + \binom{j_1}{i_1} \binom{j_2}{i_2} \cdots \binom{j_r - 1}{i_r} \right)$$
$$+ \cdots + (-1)^r \binom{j_1 - 1}{i_1} \binom{j_2 - 1}{i_2} \cdots \binom{j_r - 1}{i_r}$$

times in the left hand side. But an induction on r, with the identity

$$\binom{j}{i} - \binom{j-1}{i} = \binom{j-1}{i-1}, \tag{6.2}$$

being the $r = 1$ case gives us that

$$\binom{j_1}{i_1}\cdots\binom{j_r}{i_r} - \left(\binom{j_1-1}{i_1}\binom{j_2}{i_2}\cdots\binom{j_r}{i_r} + \cdots + \binom{j_1}{i_1}\binom{j_2}{i_2}\cdots\binom{j_r-1}{i_r}\right)$$
$$+\cdots+(-1)^r\binom{j_1-1}{i_1}\binom{j_2-1}{i_2}\cdots\binom{j_r-1}{i_r} = \binom{j_1-1}{i_1-1}\cdots\binom{j_r-1}{i_r-1}. \tag{6.3}$$

Indeed, suppose that (6.3) holds for $r-1$. Denote the left-hand side of (6.3) by C_r. We need to show that

$$C_r = \binom{j_1-1}{i_1-1}\cdots\binom{j_r-1}{i_r-1}.$$

But factoring, the identity (6.2) and the induction hypothesis give us that

$$\begin{aligned} C_r &= \binom{j_r}{i_r}C_{r-1} - \binom{j_r-1}{i_r}C_{r-1} \\ &= C_{r-1}\binom{j_r-1}{i_r-1} \\ &= \binom{j_1-1}{i_1-1}\cdots\binom{j_{r-1}-1}{i_{r-1}-1}\binom{j_r-1}{i_r-1}. \end{aligned}$$

Thus the claim (6.1) is proved. The Theorem follows by summing over all values of r from 1 to k, all possible r-tuples of primes, and for each such r-tuple, all r-tuples $(i_1,\ldots,i_r)$ satisfying $i_j > 0$ for $j = 1,\ldots,r$, and $i_1+\cdots+i_r = k$. □

We will first investigate the part of the sum in Theorem 6.1 corresponding to $r = k$, and hence $i_1 = \ldots = i_k = 1$. To do so, we require some generalizations of the prime number theorem. These are taken from Nathanson [33], pp.313-319, however we also include precise error terms.

Let

$$\begin{aligned} \pi_k(x) &= \#\{n \le x : \Omega(n) = \omega(n) = k\}; \text{ and} \\ \pi_k^*(x) &= \#\{n \le x : \Omega(n) = k\}. \end{aligned}$$

That is, $\pi_k(x)$ counts the number of $n \le x$ such that are products of exactly k distinct prime factors, and $\pi_k^*(x)$ counts the number of $n \le x$ which have k prime factors with repetition. Note that $\pi_1(x) = \pi_1^*(x) = \pi(x)$. For $k = 1$, the prime number theorem gives us

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right). \tag{6.4}$$

For $k \geq 2$, we also have that

$$\pi_k(x) = \frac{x(\log\log x)^{k-1}}{(k-1)!\log x} + O\left(\frac{x(\log\log x)^{k-2}}{\log x}\right), \tag{6.5}$$

and

$$0 \leq \pi_k^*(x) - \pi_k(x) \ll \frac{x(\log\log x)^{k-2}}{\log x}. \tag{6.6}$$

We shall use the first result to prove a Lemma:

Lemma 6.2. *Let $u \geq 0$, and $k \geq 2$. Then*

$$\sum_{\substack{n\leq x \\ \omega(n)=\Omega(n)=k}} n^u = \frac{x^{u+1}(\log\log x)^{k-1}}{(u+1)(k-1)!\log x} + O\left(\frac{x^{u+1}(\log\log x)^{k-2}}{\log x}\right).$$

Proof. Writing the sum as a Riemmann-Stieltjes integral, we have:

$$\begin{aligned}\sum_{\substack{n\leq x \\ \omega(n)=\Omega(n)=k}} n^u &= \int_{2^k-}^{x} t^u \, d\pi_k(t) \\ &= x^u\pi_k(x) - u\int_{2^k}^{x} t^{u-1}\pi_k(t)\, dt \\ &= \frac{x^{u+1}(\log\log x)^{k-1}}{(k-1)!\log x} + O\left(\frac{x^{u+1}(\log\log x)^{k-2}}{\log x}\right) \\ &\quad - u\int_{2^k}^{x} \frac{t^u(\log\log t)^{k-1}}{(k-1)!\log t}\, dt + O\left(\int_{2^k}^{x} \frac{t^u(\log\log t)^{k-2}}{\log t}\, dt\right)\end{aligned}$$

It is easy to see via a straightforward application of integration by parts that

$$\int_{2^k}^{x} \frac{t^u(\log\log t)^{k-1}}{\log t}\, dt = \frac{x^{u+1}(\log\log x)^{k-1}}{(u+1)\log x} + O\left(\frac{x^{u+1}(\log\log x)^{k-1}}{\log^2 x}\right).$$

Combining this information, we have

$$\sum_{\substack{n\leq x \\ \omega(n)=\Omega(n)=k}} n^u = \frac{x^{u+1}(\log\log x)^{k-1}}{(u+1)(k-1)!\log x} + O\left(\frac{x^{u+1}(\log\log x)^{k-2}}{\log x}\right),$$

and the Lemma is proved. □

The case when $k = 1$ was handled in an earlier chapter. Hence for the remainder, we shall assume that $k \geq 2$. We will also assume that $\ell \geq 1$.

For $r = k$, we have the following:

$$\sum_{p_1<\ldots<p_k} (p_1 \cdots p_k)^\ell \sum_{j_1,\ldots,j_k=1}^{\infty} \left\lfloor \frac{x}{p_1^{j_1} \cdots p_k^{j_k}} \right\rfloor = \sum_{p_1<\ldots<p_k} (p_1 \cdots p_k)^\ell \left\lfloor \frac{x}{p_1 \cdots p_k} \right\rfloor + \sum_{p_1<\ldots<p_k} (p_1 \cdots p_k)^\ell \sum_{\substack{j_1,\ldots,j_k=1 \\ j_1\cdots j_k>1}}^{\infty} \left\lfloor \frac{x}{p_1^{j_1} \cdots p_k^{j_k}} \right\rfloor. \tag{6.7}$$

We further focus by looking at the first term on the right-hand side of (6.7), that is, the term corresponding to $j_1 = \ldots = j_k = 1$. Making use of Lemma 6.2, we have:

$$\begin{aligned}
\sum_{p_1<\ldots<p_k} (p_1 \cdots p_k)^\ell \left\lfloor \frac{x}{p_1 \cdots p_k} \right\rfloor &= \sum_{m\leq x/2^k} \sum_{\substack{p_1<\ldots<p_k \\ \frac{x}{m+1}<p_1\cdots p_k\leq \frac{x}{m}}} m(p_1 \cdots p_k)^\ell \\
&= \sum_{m\leq x/2^k} \sum_{\substack{p_1<\ldots<p_k \\ p_1\cdots p_k\leq x/m}} (p_1 \cdots p_k)^\ell \\
&= \frac{x^{\ell+1}}{(\ell+1)(k-1)!} \sum_{m\leq x/2^k} \frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} \\
&\quad + O\left(x^{\ell+1} \sum_{m\leq x/2^k} \frac{(\log\log(x/m))^{k-2}}{m^{\ell+1}\log(x/m)} \right).
\end{aligned} \tag{6.8}$$

Now

$$\sum_{m\leq x/2^k} \frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} = \sum_{m\leq \log^2 x} \frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} + \sum_{\log^2 x<m\leq x/2^k} \frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)}. \tag{6.9}$$

For $m \in [1, \log^2 x]$,

$$\begin{aligned}\frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} &\leq \frac{(\log\log x)^{k-1}}{m^{\ell+1}(\log x - 2\log\log x)} \\ &= \frac{(\log\log x)^{k-1}}{m^{\ell+1}\log x}\left(1+O\left(\frac{\log\log x}{\log x}\right)\right),\end{aligned}$$

which implies that

$$\begin{aligned}\sum_{m\leq \log^2 x}\frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} \leq & \frac{(\log\log x)^{k-1}}{\log x}\left(1+O\left(\frac{\log\log x}{\log x}\right)\right)\sum_{m\leq\log^2 x}\frac{1}{m^{\ell+1}} \\ =& \frac{(\log\log x)^{k-1}}{\log x}\left(1+O\left(\frac{\log\log x}{\log x}\right)\right) \\ & \times\zeta(\ell+1)\left(1+O\left(\frac{1}{\log^{2\ell} x}\right)\right) \\ =& \frac{\zeta(\ell+1)(\log\log x)^{k-1}}{\log x}+O\left(\frac{(\log\log x)^k}{\log^2 x}\right).\end{aligned} \tag{6.10}$$

On the other hand, for $m \in [1, \log^2 x]$ we have

$$\begin{aligned}\frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} \geq & \frac{(\log(\log x - 2\log\log x))^{k-1}}{m^{\ell+1}\log x} \\ =& \frac{\left(\log\log x + \log\left(1-\frac{2\log\log x}{\log x}\right)\right)^{k-1}}{m^{\ell+1}\log x} \\ =& \frac{\left(\log\log x + O\left(\frac{\log\log x}{\log x}\right)\right)^{k-1}}{m^{\ell+1}\log x} \\ =& \frac{(\log\log x)^{k-1}+O\left(\frac{(\log\log x)^{k-1}}{\log x}\right)}{m^{\ell+1}\log x},\end{aligned}$$

and so

$$\sum_{m\le\log^2 x}\frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} \ge \frac{(\log\log x)^{k-1}}{\log x}\left(\zeta(\ell+1)+O\left(\frac{1}{\log^{2\ell}x}\right)\right)$$
$$+O\left(\frac{(\log\log x)^{k-1}}{\log^2 x}\right)$$
$$=\frac{\zeta(\ell+1)(\log\log x)^{k-1}}{\log x}+O\left(\frac{(\log\log x)^{k-1}}{\log^2 x}\right). \tag{6.11}$$

Combining (6.10) and (6.11) we have that

$$\sum_{m\le\log^2 x}\frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)}=\frac{\zeta(\ell+1)(\log\log x)^{k-1}}{\log x}+O\left(\frac{(\log\log x)^{k}}{\log^2 x}\right). \tag{6.12}$$

Now we must bound the second term on the right-hand side of (6.9).

$$\sum_{\log^2 x<m\le x/2^k}\frac{(\log\log(x/m))^{k-1}}{m^{\ell+1}\log(x/m)} \ll \sum_{\log^2 x<m\le x/2^k}\frac{(\log\log x)^{k-1}}{m^{\ell+1}}$$
$$\ll\frac{(\log\log x)^{k-1}}{\log^2 x}.$$

A similar argument shows that

$$\sum_{m\le x/2^k}\frac{(\log\log(x/m))^{k-2}}{m^{\ell+1}\log(x/m)}\ll\frac{(\log\log x)^{k-2}}{\log x}. \tag{6.13}$$

This we use to bound the error term in (6.8).

Applying this information to (6.8) we have that

$$
\begin{aligned}
\sum_{p_1<\ldots<p_k} (p_1\cdots p_k)^\ell \left\lfloor \frac{x}{p_1\cdots p_k} \right\rfloor =& \frac{x^{\ell+1}}{(\ell+1)(k-1)!} \times \\
&\left(\frac{\zeta(\ell+1)(\log\log x)^{k-1}}{\log x} + O\left(\frac{(\log\log x)^k}{\log^2 x} \right)\right) \\
&+ O\left(\frac{x^{\ell+1}(\log\log x)^{k-2}}{\log x} \right) \\
=& \frac{\zeta(\ell+1)x^{\ell+1}(\log\log x)^{k-1}}{(\ell+1)(k-1)!\log x} \\
&+ O\left(\frac{x^{\ell+1}(\log\log x)^{k-2}}{\log x} \right). \qquad (6.14)
\end{aligned}
$$

This is the main term in $\sum_{n\le x} s_{k,\ell}(n)$. To complete the computation of the sum, we need only bound all that remains. We will first complete the case when $r = k$, by bounding the second term in the right-hand side of (6.7):

$$\begin{aligned}
\sum_{p_1<\ldots<p_k} (p_1\cdots p_k)^\ell \sum_{\substack{j_1,\ldots,j_k=1\\ j_1\cdots j_k>1}}^{\infty} \left\lfloor \frac{x}{p_1^{j_1}\cdots p_k^{j_k}} \right\rfloor
&\ll x \sum_{\substack{p_1<\ldots<p_k\\ p_1^2p_2\cdots p_k\le x}} (p_1\cdots p_k)^\ell \sum_{\substack{j_1,\ldots,j_k=1\\ j_1\cdots j_k>1}}^{\infty} \frac{1}{p_1^{j_1}\cdots p_k^{j_k}}\\
&\ll x \sum_{\substack{p_1<\ldots<p_k\\ p_1^2p_2\cdots p_k\le x}} (p_1\cdots p_k)^\ell \left(\frac{1}{(p_1-1)\cdots(p_k-1)} - \frac{1}{p_1\cdots p_k} \right)\\
&\ll x \sum_{\substack{p_1<\ldots<p_k\\ p_1^2p_2\cdots p_k\le x}} \frac{(p_1\cdots p_k)^\ell}{p_1^2p_2\cdots p_k}\\
&\ll x \sum_{p\le x^{\frac{1}{k+1}}} \left(p^{\ell-2} \sum_{\substack{n\le x/p\\ \omega(n)=\Omega(n)=k-1}} n^{\ell-1} \right)\\
&\ll x \sum_{p\le x^{\frac{1}{k+1}}} \left(p^{\ell-2} \frac{(x/p)^\ell(\log\log(x/p))^{k-2}}{\log(x/p)} \right)\\
&= x^{\ell+1} \sum_{p\le x^{\frac{1}{k+1}}} \frac{(\log\log(x/p))^{k-2}}{p^2\log(x/p)}\\
&\ll \frac{x^{\ell+1}(\log\log x)^{k-2}}{\log x}. \qquad (6.15)
\end{aligned}$$

Let us now bound the terms of Theorem 6.1 corresponding to $r<k$. We require the following power series identity:

$$\sum_{n=j}^{\infty} \binom{n-1}{j-1} x^n = \left(\frac{x}{1-x}\right)^j,$$

which holds for $|x|<1$. We have

$$\begin{aligned}
&\sum_{r=1}^{k-1} \sum_{p_1<\ldots<p_r} \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r>0}} (p_1^{i_1}\cdots p_r^{i_r})^\ell \times \\
&\quad \left(\sum_{j_1,\ldots,j_r=1}^{\infty} \binom{j_1-1}{i_1-1} \cdots \binom{j_r-1}{i_r-1} \left\lfloor \frac{x}{p_1^{j_1}\cdots p_r^{j_r}} \right\rfloor \right) \\
&\quad \ll x \sum_{r=1}^{k-1} \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r>0}} \sum_{\substack{p_1<\ldots<p_r \\ p_1^{i_1}\cdots p_r^{i_r}\le x}} (p_1^{i_1}\cdots p_r^{i_r})^\ell \prod_{m=1}^{r} \sum_{j_m=i_m}^{\infty} \binom{j_m-1}{i_m-1} \frac{1}{p^{j_m}} \\
&\quad = x \sum_{r=1}^{k-1} \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r>0}} \sum_{\substack{p_1<\ldots<p_r \\ p_1^{i_1}\cdots p_r^{i_r}\le x}} (p_1^{i_1}\cdots p_r^{i_r})^\ell \prod_{m=1}^{r} \left(\frac{1}{p_m-1} \right)^{i_m} \\
&\quad \ll x \sum_{r=1}^{k-1} \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r>0}} \sum_{\substack{p_1<\ldots<p_r \\ p_1^{i_1}\cdots p_r^{i_r}\le x}} (p_1^{i_1}\cdots p_r^{i_r})^{\ell-1} \\
&\quad = x \sum_{\substack{n\le x \\ \omega(n)<k=\Omega(n)}} n^{\ell-1} \\
&\quad = x \int_{2^-}^{x} t^{\ell-1}\, d(\pi_k^*(t)-\pi_k(t)). \qquad (6.16)
\end{aligned}$$

Applying integration by parts to (6.16), and using the bound (6.6), we have that

$$\begin{aligned}
&\sum_{r=1}^{k-1} \sum_{p_1<\ldots<p_r} \sum_{\substack{i_1+\cdots+i_r=k \\ i_1,\ldots,i_r>0}} (p_1^{i_1}\cdots p_r^{i_r})^\ell \times \\
&\quad \left(\sum_{j_1,\ldots,j_r=1}^{\infty} \binom{j_1-1}{i_1-1} \cdots \binom{j_r-1}{i_r-1} \left\lfloor \frac{x}{p_1^{j_1}\cdots p_r^{j_r}} \right\rfloor \right) \\
&\quad \ll \frac{x^{\ell+1}(\log\log x)^{k-2}}{\log x}. \qquad (6.17)
\end{aligned}$$

Combining Theorem 6.1 with (6.14), (6.15), and (6.17), we have proved the following Theorem:

Theorem 6.3. *Let $k \geq 2$, and let $\ell \geq 1$. Then*

$$\sum_{n \leq x} s_{k,\ell}(n) = \frac{\zeta(\ell+1)x^{\ell+1}(\log\log x)^{k-1}}{(\ell+1)(k-1)!\log x} + O\left(\frac{x^{\ell+1}(\log\log x)^{k-2}}{\log x}\right).$$

Chapter 7

Modular Distribution of Prime Symmetric Functions

The material found in this chapter is an outgrowth of work found in the paper of the author entitled "On the modular distribution of divisor functions" submitted for publication to the Journal of Number Theory.

The problems we tackle here are motivated by a result of Alladi and Erdős. They showed [1], that the sum of prime factors with repetition function, e_1 is uniformly distributed modulo 2. More precisely, they used an elementary argument to demonstrate that there is a constant $c_0 > 0$ such that

$$\sum_{n\leq x}(-1)^{e_1(n)} \ll xe^{-c_0\sqrt{\log x \log\log x}}.$$

They also proved the related result:

$$\sum_{n=1}^{\infty}\frac{(-1)^{e_1(n)}}{n} = 0.$$

We will prove look at some generalizations of this result, though with weaker error terms. Our main tool will be Perron's formula.

7.1 Perron's Formula

$$\left|{\sum_{n\leq x}}' a_n - \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{A(s)x^s}{s}\,ds\right| \ll \left(\sum_{n=1}^{\infty}|a_n|n^{-c}\right)\frac{x^c}{T} + \sum_{\frac{x}{2}<n<\frac{3x}{2}}|a_n|\min\left\{1, \frac{x}{T|x-n|}\right\} + \left(\frac{a_x}{T} \text{ if } x\in\mathbb{N}\right).$$

The version in this form was first shown to us by Dr. Greg Martin from his lecture notes. It can be derived from [32], pp. 137-142. The summation symbol Σ' indicates that if x is an integer, then "the last term is to be counted with weight 1/2" ([32] p.138). Here $A(s)$ is the Dirichlet series for the sequence a_n, and the constant c is a real number greater than the maximum of 0 and the abscissa of convergence of $A(s)$.

In the cases we shall be concerned with, a_n will always satisfy $|a_n| \leq 1$. Consequently, we can take $c = 1+1/\log x$. We will also assume that $T \to \infty$ but that $T \leq x$.

The third error term is $O(1/T)$. The first error term satisfies:

$$\left(\sum_{n=1}^{\infty} |a_n| n^{-c}\right) \frac{x^c}{T} \ll \zeta(c)\frac{x^c}{T} \ll \frac{x}{T(c-1)} \ll \frac{x \log x}{T}.$$

For the second error term, we break up the sum:

$$\begin{aligned}\sum_{\frac{x}{2}<n<\frac{3x}{2}} |a_n| \min\left\{1, \frac{x}{T|x-n|}\right\} = &\sum_{\frac{x}{2}<n<\left(1-\frac{1}{T}\right)x} \frac{x}{(x-n)T} \\ &+ \sum_{\left(1-\frac{1}{T}\right)x \leq n \leq \left(1+\frac{1}{T}\right)x} 1 + \sum_{\left(1+\frac{1}{T}\right)x<n<\frac{3x}{2}} \frac{x}{(n-x)T} \\ &= \Sigma_1 + \Sigma_2 + \Sigma_3.\end{aligned}$$

Now $\Sigma_2 \ll x/T$, and

$$\Sigma_1 \ll \frac{x}{T} \int_{\frac{x}{2}}^{\left(1-\frac{1}{T}\right)x} \frac{dt}{x-t} = \frac{x}{T}\left[-\log\left(x-t\right)\right]_{t=\frac{x}{2}}^{t=\left(1-\frac{1}{T}\right)x} \ll \frac{x \log T}{T}.$$

A similar result holds for Σ_3. We have proved the following Lemma:

Lemma 7.1. *If $|a_n| \leq 1$, $T \leq x$ and $c = 1 + 1/\log x$, then*

$$\left| \sum_{n \leq x} a_n - \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{A(s)x^s}{s}\, ds \right| \ll \frac{x \log x}{T}.$$

Lemma 7.2. *Suppose that $A(s)$ is analytic on a region of the form*

$$D = \left\{\sigma + it : \sigma \geq 1 - \frac{K}{\log\left(|t|+4\right)}\right\} \setminus \{y \in \mathbb{R} : y \leq 1\}.$$

Suppose further that there are real constants $w > 0$ and α, $0 \leq \alpha < 1$ such that

$$\begin{aligned} A(s) &\ll \log^w |t| \;\; \textit{if } |t| \geq 4 \\ A(s) &\ll \frac{1}{|s-1|^\alpha} \;\; \textit{if } |t| \leq 4. \end{aligned}$$

If $c = 1 + 1/\log x$, and $T = \exp\left(\log^\delta x\right)$, where $\delta = (1+\alpha^2)/2$, then

$$\int_{c-iT}^{c+iT} \frac{A(s)x^s}{s}\, ds \ll \frac{x}{(\log x)^{(1-\alpha)/2}}.$$

Furthermore, if $A(s)$ is the Dirichlet series for a_n, and $|a_n| \le 1$, then

$$\sum_{n\le x} a_n \ll \frac{x}{(\log x)^{(1-\alpha)/2}}.$$

Proof. The last statement follows from the first together with Lemma 7.1. The idea of the proof is to pull back the integral where we can, and integrate around the branch cut. Let $c' = 1 - \frac{K}{\log(T+4)}$, let $\beta = (1+\alpha)/2$, and let $c'' = 1/\log^\beta x$.

Shift the line of integration to the union of the following line segments in sequence:

$$\begin{aligned}
&\gamma_1 : c - iT \text{ to } c' - iT\\
&\gamma_2 : c' - iT \text{ to } c' - i4\\
&\gamma_3 : c' - i4 \text{ to } c' - ic''\\
&\gamma_4 : c' - ic'' \text{ to } c - ic''\\
&\gamma_5 : c - ic'' \text{ to } c + ic''\\
&\gamma_6 : c + ic'' \text{ to } c' + ic''\\
&\gamma_7 : c' + ic'' \text{ to } c' + i4\\
&\gamma_8 : c' + i4 \text{ to } c' + iT\\
&\gamma_9 : c' + iT \text{ to } c + iT.
\end{aligned}$$

Assume that $T \ge 4$. Denote the integral of $A(s)x^s/s$ over γ_i as $\int_i$. Then

$$\begin{aligned}
\int_9 &\ll \int_{c'+iT}^{c+iT} \frac{|A(s)||x^s|}{|s|}\, ds\\
&\ll \frac{x^c \log^w T(c-c')}{T}\\
&\ll \frac{x \log^{w-1} T}{T}\\
&\ll \frac{x(\log x)^{\delta(w-1)}}{\exp(\log^\delta x)}
\end{aligned}$$

The integral over γ_1 can be likewise bounded.

Next,

$$\begin{aligned}\int_8 &\ll \int_{c'+i4}^{c'+iT} \frac{|A(s)||x^s|}{|s|}\,ds \\ &\ll x^{c'} \int_2^T \frac{\log^w t}{t}\,dt \\ &\ll \frac{x\log^{w+1} T}{\exp\left(\frac{K\log x}{\log T}\right)} \\ &\ll \frac{x(\log x)^{(w+1)\delta}}{\exp\left(K(\log x)^{1-\delta}\right)}\end{aligned}$$

The integral over γ_2 can be likewise bounded. Furthermore,

$$\int_7 \ll x^{c'} \int_{c'+ic''}^{c'+i4} \frac{ds}{|s-1|^\alpha} \ll \frac{x\log^\alpha T}{\exp\left(\frac{K\log x}{\log T}\right)} \ll \frac{x\log^\alpha T}{\exp\left(K(\log x)^{1-\delta}\right)},$$

and the integral over γ_3 can be likewise bounded. Over γ_6 we have

$$\int_6 \ll x^c \int_{c+ic''}^{c'+ic''} \frac{ds}{|s-1|^\alpha} \ll \frac{x}{(c'')^\alpha \log T} \ll \frac{x}{(\log x)^{\delta-\alpha\beta}} = \frac{x}{(\log x)^{(1-\alpha)/2}},$$

with a similar result holding for γ_4.

Finally,

$$\int_5 \ll x^c \int_{c-ic''}^{c+ic''} \frac{ds}{|s-1|^\alpha} \ll xc'' \log^\alpha x = \frac{x}{(\log x)^{\beta-\alpha}} = \frac{x}{(\log x)^{(1-\alpha)/2}}.$$

In each case, it is clear that the integral is bounded as the Lemma requires. □

7.2 Bounding $L(s,\chi)$ and $\zeta(s)$

Fix a modulus q. We wish to bound $\zeta(s)$ and $L(s,\chi)$ for a non-principal character χ modulo q in ways that will be useful to us.

We have the following consequences of [32] pp.362-363. There is a positive constant K_q such that if $\sigma \geq 1 - K_q/\log(|t|+4)$, then for each non-principal $\chi \pmod q$, $L(s,\chi)$ is nonzero, and satisfies

$$\begin{aligned}|\log L(s,\chi)| &\leq \log\log(|t|+4) + O_q(1); \\ \frac{1}{L(s,\chi)} &\ll_q \log(|t|+4). \end{aligned} \tag{7.1}$$

These inequalities in turn imply that:

$$L(s,\chi) \ll_q \log(|t|+4); \tag{7.2}$$
$$|\arg L(s,\chi)| \leq 2\log\log(|t|+4) + O_q(1). \tag{7.3}$$

Note that we need not concern ourselves with whether or not q is exceptional, since we need only take the constant K_q to be smaller to produce the same results.

In the case when $\chi = \chi_0$ is the principal character modulo q, we have

$$L(s,\chi_0) = \zeta(s)\prod_{p|q}(1-p^{-s}). \tag{7.4}$$

From [5], p.188, 197-198, 208, there is a constant $K > 0$ such that for $\sigma \geq 1 - K/\log(|t|+4)$, $\zeta(s)$ is nonzero, and satisfies the following:

$$\frac{1}{\zeta(s)} \ll \log(|t|+4), \tag{7.5}$$
$$\zeta(s) \ll \log|t|, \text{ if } |t| \geq 4 \tag{7.6}$$
$$\frac{\zeta'}{\zeta}(s) \ll \log|t|, \text{ if } |t| \geq 4. \tag{7.7}$$

From equations (7.5) and (7.6), it follows that

$$|\log|\zeta(s)|| \leq \log\log|t| + O(1), \text{ for } |t| \geq 4. \tag{7.8}$$

Let D denote the region

$$D = \left\{s : 1 - \frac{K}{\log(|t|+4)} \leq \sigma \leq 2\right\} \setminus \{y \in \mathbb{R} : y \leq 1\}.$$

Lemma 7.3. *For $s \in D$, with $|t| \geq 4$,*

$$\log\zeta(s) \ll \log\log t$$

Proof. We may assume that $t \geq 4$. For any δ, $0 < \delta \leq 1$, if $1+\delta \leq \varsigma \leq 2$, then

$$\begin{aligned}\left|\frac{\zeta'}{\zeta}(\varsigma+it)\right| &\le \sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^{1+\delta}}\\ &= \sum_{p}\frac{\log p}{p^{1+\delta}}+O(1)\\ &= \int_{2^-}^{\infty}\frac{\log u}{u^{1+\delta}}\,d\pi(u)+O(1)\\ &\ll \int_{2}^{\infty}\frac{du}{u^{1+\delta}}+O(1)\\ &= O\left(\frac{1}{\delta}\right). \end{aligned} \tag{7.9}$$

For a given $N\in\mathbb{N}$, σ must lie in one of the intervals

$$\left[1-\frac{K}{\log(t+4)},1+\frac{1}{\log t}\right],\left[1+\frac{1}{\log t},1+\frac{1}{(\log t)^{1-1/N}}\right],$$
$$\left[1+\frac{1}{(\log t)^{1-1/N}},1+\frac{1}{(\log t)^{1-2/N}}\right],\ldots,\left[1+\frac{1}{(\log t)^{1/N}},2\right].$$

Let $\sigma_n=1+1/(\log t)^{n/N}$, for $n=0,1,\ldots,N$, and let $\sigma_{N+1}=1-K/\log(t+4)$. Now

$$\begin{aligned}&|\log\zeta(s)| = \left|\int_{\infty}^{\sigma}\frac{\zeta'}{\zeta}(\varsigma+it)\,d\varsigma\right|\\ &\le \int_{\infty}^{2}\left|\frac{\zeta'}{\zeta}(\varsigma+it)\right|d\varsigma+\sum_{n=0}^{N-1}\int_{\sigma_n}^{\sigma_{n+1}}\left|\frac{\zeta'}{\zeta}(\varsigma+it)\right|d\varsigma+\int_{\sigma_N}^{\sigma_{N+1}}\left|\frac{\zeta'}{\zeta}(\varsigma+it)\right|d\varsigma\\ &= O(1)+\sum_{n=0}^{N-1}O\left(\frac{\sigma_n-1}{\sigma_{n+1}-1}\right)+O(1)\\ &= O\left(N(\log t)^{1/N}\right)+O(1).\end{aligned}$$

Putting $N=\lfloor\log\log t\rfloor$, we have the desired result. □

We use Lemma 7.3 to prove the following:

Lemma 7.4. *Let $\omega\in\mathbb{C}$ satisfy $|\omega|\le 1$. There is an absolute positive constant C_0 such that for $s\in D$ with $|t|\ge 4$,*

$$\exp(\omega\log\zeta(s))\ll\log^{C_0}|t|.$$

Proof. Write $\omega = \alpha + i\beta$. Then for s as required,

$$\begin{aligned}|\exp(\omega\log\zeta(s))| =& |\zeta(s)|^{\alpha}\exp(-\beta\arg\zeta(s)) \\ &\ll \log^{|\alpha|}|t|\exp(|\log\zeta(s)| + |\log|\zeta(s)||), \text{ by (7.5) and (7.6),} \\ &\ll \exp(O(\log\log t)), \text{ by (7.8) and Lemma 7.3.}\end{aligned}$$

□

Lemma 7.5. *Let $\omega = \alpha + i\beta$. For $s \in D$, $|t| \leq 4$,*

$$\exp(\omega\log\zeta(s)) \ll \frac{1}{|s-1|^{\alpha}}.$$

Proof. Since $\zeta(s) = 1/(s-1) + O(1)$, on the specified region we have that $\arg\zeta(s) \ll 1$. Hence

$$\begin{aligned}|\exp(\omega\log\zeta(s))| &= |\zeta(s)|^{\alpha}\exp(-\beta\arg\zeta(s)) \\ &\ll |\zeta(s)|^{\alpha} \\ &= \frac{1}{|s-1|^{\alpha}}(1 + O(s-1)) \ll \frac{1}{|s-1|^{\alpha}}.\end{aligned}$$

□

Lemma 7.6. *Let $\omega = \alpha + i\beta$. For $\sigma \geq 1 - K_q/\log(|t|+4)$, with χ non-principal, we have that*

$$\exp(\omega\log L(s,\chi)) \ll_q (\log(|t|+4))^{|\alpha|+2|\beta|}.$$

Proof. For s as required, we have by (7.1), (7.2) and (7.3) that

$$\begin{aligned}|\exp(\omega\log L(s,\chi))| &= |L(s,\chi)|^{\alpha}\exp(-\beta\arg L(s,\chi)) \\ &\ll_q (\log(|t|+4))^{|\alpha|}\exp(2|\beta|\log\log(|t|+4)) \\ &= (\log(|t|+4))^{|\alpha|+2|\beta|}.\end{aligned}$$

□

Lemma 7.7. *Suppose $(a,q) = 1$, and $\omega \in \mathbb{C}$ satisfies $|\omega| \leq 1$. Then*

$$\prod_{p\equiv a\ (mod\ q)} \frac{1}{1-\omega p^{-s}} = \exp\left[\frac{\omega}{\phi(q)}\sum_{\chi\ (mod\ q)} \bar{\chi}(a)\log L(s,\chi)\right]\exp(\varphi(s)),$$

where $\varphi(s)$ is analytic on $\sigma > 1/2$, and satisfies $\varphi(s) \ll_{\epsilon} 1$ for $\sigma \geq 1/2 + \epsilon$.

Hence $\prod_{p\equiv a\,(q)} \frac{1}{1-\omega p^{-s}}$ *has an analytic continuation to*

$$D_q = \left\{ s : 1 - \frac{K_q}{\log\left(|t|+4\right)} \leq \sigma \right\} \setminus \{y \in \mathbb{R} : y \leq 1\}.$$

Proof. We shall make use of the result

$$\frac{1}{\phi(q)} \sum_{\chi(\text{mod } q)} \bar{\chi}(a) \log L(s,\chi) = \sum_{p^j \equiv a (\text{mod } q)} \frac{p^{-js}}{j},$$

(c.f. [5], p. 239), which holds for $\sigma > 1$. Let

$$\varphi(s) = \sum_{p\equiv a\,(\text{mod } q)} \sum_{j=2}^{\infty} \frac{\omega^j p^{-js}}{j} - \omega \sum_{\substack{p^j \equiv a (\text{mod } q) \\ j \geq 2}} \frac{p^{-js}}{j}$$

Thus $\varphi(s)$ satisfies the claims of the Lemma, because it is represented as an absolutely convergent series for $\sigma > 1/2$, and as the coefficients of p^{-js} are bounded, $\varphi(s)$ is bounded as required. Furthermore:

$$\begin{aligned}
\prod_{p\equiv a\,(\text{mod } q)} \frac{1}{1-\omega p^{-s}} &= \exp\left[\sum_{p\equiv a\,(\text{mod } q)} \sum_{j=1}^{\infty} \frac{\omega^j p^{-js}}{j} \right] \\
&= \exp\left[\omega \sum_{p^j \equiv a\,(\text{mod } q)} \frac{p^{-js}}{j} \right] \exp(\varphi(s)) \\
&= \exp\left[\frac{\omega}{\phi(q)} \sum_{\chi\,(\text{mod } q)} \bar{\chi}(a) \log L(s,\chi) \right] \exp(\varphi(s)).
\end{aligned}$$

The lemma then follows from the results on zero free regions of analyticity for L-functions. □

7.3 Additive Prime Symmetric Functions

In this section, we will put the previous results to work in generalizing Alladi and Erdős. Let $m, r, k_1, \ldots, k_r \in \mathbb{N}$, let b be an integer such that $1 \leq b \leq m$, and let $q_1, \ldots, q_r \geq 2$ be pairwise relatively prime integers. In general, we shall denote the real and imaginary parts of a complex number s by σ and t respectively, and for $n \in \mathbb{N}$, we shall write ζ_n for the primitive n-th root of unity $e^{2\pi i/n}$.

The following notations apply to Section 7.3:

Notation 7.8.

$$\begin{aligned}
\mathbf{e} &= (e_{k_1}, \ldots, e_{k_r}),\\
q &= q_1 \cdots q_r,\\
d &= gcd(m, b)\\
m' &= m/d\\
b' &= b/d\\
Q &= lcm(q, m')\\
S &= \{1 \le n < Q : (n, Q) = 1\}\\
e_{k\ell}(n) &= \begin{cases} 1, & \textit{if } n \equiv \ell \textit{ (mod } k\textit{);}\\ 0 & \textit{otherwise.}\end{cases}
\end{aligned}$$

For ease of explication, let $\mathbf{a} = (a_1, \ldots, a_r)$ be a fixed element of $\mathbb{Z}_{q_1} \times \cdots \times \mathbb{Z}_{q_r}$, and let $\mathbf{q} = (q_1, \ldots, q_r)$, and $\mathbf{k} = (k_1, \ldots, k_r)$. Congruences involving these vectors are taken coordinate-wise over the appropriate moduli.

We wish to show that the collection of functions $(e_{k_1}, \ldots, e_{k_r})$ taken in order is simultaneously uniformly distributed modulo $(q_1, \ldots, q_r)$, over the arithmetic sequence b (mod m). The main result of Section 7.3 is the following asymptotic formula:

$$\sum_{\substack{n \le x\\ n \equiv b (\text{mod } m)\\ \mathbf{e}(n) \equiv \mathbf{a} (\text{mod } \mathbf{q})}} 1 \sim \frac{x}{mq} \text{ as } x \to \infty. \tag{7.10}$$

Observe that

$$\frac{1}{q} \prod_{\ell=1}^{r} \left(\sum_{j=0}^{q_\ell - 1} \zeta_{q_\ell}^{(e_{k_\ell}(n) - a_\ell) j} \right) = \begin{cases} 1, & \text{if } \mathbf{e}(n) \equiv \mathbf{a}(\mathbf{q});\\ 0 & \text{otherwise.}\end{cases}$$

For $\mathbf{j} = (j_1, \ldots, j_r) \in \mathbb{Z}_{q_1} \times \cdots \times \mathbb{Z}_{q_r}$, let

$$f(n) = f(\mathbf{j}; n) = \zeta_{q_1}^{e_{k_1}(n) j_1} \cdots \zeta_{q_r}^{e_{k_r}(n) j_r}.$$

Note that $f(n)$ is completely multiplicative. Hence we have

$$\sum_{\substack{n\le x\\ n\equiv b(\mathrm{mod}\ m)\\ \mathbf{e}(n)\equiv\mathbf{a}(\mathrm{mod}\ \mathbf{q})}} 1 = \frac{1}{q}\sum_{j_1=0}^{q_1-1}\zeta_{q_1}^{-a_1j_1}\cdots\sum_{j_r=0}^{q_r-1}\zeta_{q_r}^{-a_rj_r}\sum_{\substack{n\le x\\ n\equiv b(m)}} f(\mathbf{j};n)$$

$$= \frac{1}{q}\sum_{j_1=0}^{q_1-1}\zeta_{q_1}^{-a_1j_1}\cdots\sum_{j_r=0}^{q_r-1}\zeta_{q_r}^{-a_rj_r}\sum_{0\le n\le(x-b)/m} f(\mathbf{j};mn+b)$$

$$= \frac{1}{q}\sum_{j_1=0}^{q_1-1}\zeta_{q_1}^{-a_1j_1}\cdots\sum_{j_r=0}^{q_r-1}\zeta_{q_r}^{-a_rj_r} f(\mathbf{j};d)\sum_{0\le n\le(x-b)/m} f(\mathbf{j};m'n+b')$$

$$= \frac{1}{q}\sum_{j_1=0}^{q_1-1}\zeta_{q_1}^{-a_1j_1}\cdots\sum_{j_r=0}^{q_r-1}\zeta_{q_r}^{-a_rj_r} f(\mathbf{j};d)\sum_{n\le x/d} e_{m'b'}(n)f(\mathbf{j};n).$$

Since m' and b' are relatively prime, we may write

$$e_{m'b'}(n) = \frac{1}{\phi(m')}\sum_{\chi(\mathrm{mod}\ m')}\bar{\chi}(b')\chi(n).$$

Thus,

$$\sum_{\substack{n\le x\\ n\equiv b(\mathrm{mod}\ m)\\ \mathbf{e}(n)\equiv\mathbf{a}(\mathrm{mod}\ \mathbf{q})}} 1 = \frac{1}{q\phi(m')}\sum_{j_1=0}^{q_1-1}\zeta_{q_1}^{-a_1j_1}\cdots\sum_{j_r=0}^{q_r-1}\zeta_{q_r}^{-a_rj_r}$$

$$\times f(\mathbf{j};d)\sum_{\chi(\mathrm{mod}\ m')}\bar{\chi}(b')\sum_{n\le x/d}\chi(n)f(\mathbf{j};n).$$

Note that $f(\mathbf{0};n) = 1$, and so

$$\begin{aligned}
\sum_{\substack{n\le x\\ n\equiv b(\text{mod } m)\\ \mathbf{e}(n)\equiv \mathbf{a}(\text{mod } \mathbf{q})}} 1 &= \frac{1}{q\phi(m')}\sum_{\chi(\text{mod } m')}\bar{\chi}(b')\sum_{n\le x/d}\chi(n)\\
&+\frac{1}{q\phi(m')}\sum_{\mathbf{j}\neq\mathbf{0}}\zeta_{q_1}^{-a_1j_1}\cdots\zeta_{q_r}^{-a_rj_r}f(\mathbf{j};d)\sum_{\chi(\text{mod } m')}\bar{\chi}(b')\sum_{n\le x/d}\chi(n)f(\mathbf{j};n)\\
=&\frac{x}{qm}+O(1)\\
&+\frac{1}{q\phi(m')}\sum_{\mathbf{j}\neq\mathbf{0}}\zeta_{q_1}^{-a_1j_1}\cdots\zeta_{q_r}^{-a_rj_r}f(\mathbf{j};d)\sum_{\chi(\text{mod } m')}\bar{\chi}(b')\sum_{n\le x/d}\chi(n)f(\mathbf{j};n)
\end{aligned} \tag{7.11}$$

To prove the asymptotic formula (7.10), it suffices, due to (7.11), to show that
$\sum_{n\le x}\chi(n)f(n)=o(x)$, for any fixed $\mathbf{j}\neq\mathbf{0}$ and character χ (mod m'). To do this, we shall use Perron's formula, and hence we must consider the Dirichlet series $\sum_{n=1}^{\infty}\chi(n)f(n)n^{-s}$. Fix such a $\mathbf{j}$ and χ (mod m'). Now, for $\sigma>1$, define

$$A(s)=A(\mathbf{j},\chi;s)=\sum_{n=1}^{\infty}\chi(n)f(n)n^{-s}.$$

Note that $A(s)$ is absolutely convergent for $\sigma>1$. Due to the complete multiplicativity of $\chi(n)f(n)$, $A(s)$ has the following Euler product for $\sigma>1$:

$$\begin{aligned}
A(s)&=\prod_{p}\frac{1}{1-\chi(p)f(p)p^{-s}}\\
&=\prod_{p\mid q}\frac{1}{1-\chi(p)f(p)p^{-s}}\prod_{a\in S}\prod_{p\equiv a(\text{mod } Q)}\frac{1}{1-\chi(a)\zeta_{q_1}^{a^{k_1}j_1}\cdots\zeta_{q_r}^{a^{k_r}j_r}p^{-s}}.
\end{aligned}$$

The first term in this product is a non-zero analytic function for $\sigma>0$, which we will denote by $F(s)$. That is,

$$F(s)=F(\mathbf{j},\chi;s)=\prod_{p\mid q}\frac{1}{1-\chi(p)f(p)p^{-s}}.$$

$A(s)$ continues analytically to the region D_Q as described in Lemma 7.7. By taking K_Q smaller if necessary, we may insist that $D_Q\subseteq D\cup\{s\in\mathbb{C}:$

$\sigma \geq 1\}$, and so the results of Lemmas 7.3 through 7.6 apply for $s \in D_Q$. Furthermore, there is a function $\varphi(s)$ satisfying the claims of Lemma 7.7 such that

$$A(s) = F(s)\exp(\varphi(s))\exp\left[\sum_{\chi^* (\mathrm{mod}\ q)} \omega(\chi^*)\log L(s,\chi^*)\right],$$

where

$$\omega(\chi^*) = \omega(\mathbf{j},\chi;\chi^*) = \frac{1}{\phi(Q)}\sum_{a\in S}\chi(a)\bar{\chi^*}(a)\zeta_{q_1}^{a^{k_1}j_1}\cdots\zeta_{q_r}^{a^{k_r}j_r}.$$

Note that $F(s) \ll_{\mathbf{q},\mathbf{k}} 1$ on $1/2 \leq \sigma \leq 2$, and also that $|\omega(\chi^*)| \leq 1$.

Write

$$\omega(\chi^*) = \alpha(\chi^*) + \beta(\chi^*)i,$$

for its real and imaginary parts, and for the principal character modulo Q, make the denotations:

$$\begin{aligned}\omega_0 &= \omega(\chi_0^*),\\ \alpha_0 &= \alpha(\chi_0^*),\\ \beta_0 &= \beta(\chi_0^*).\end{aligned}$$

Remark 7.9. *It transpires that* $\omega(\chi^*) \neq 1$. *Indeed, if* $\omega(\chi^*) = 1$, *then* $\chi(a)\bar{\chi^*}(a)\zeta_{q_1}^{a^{k_1}j_1}\cdots\zeta_{q_r}^{a^{k_r}j_r} = 1$ *for all* $a \in S$, *and in particular for* $a = 1$. *However, since* $q_1,\ldots,q_r$ *are pairwise relatively prime,* $\mathbf{j} \neq \mathbf{0}$ *by assumption, and* $\chi(1) = \chi^*(1) = 1$, *this yields a contradiction. It follows that* $\alpha(\chi^*) < 1$.

Define constants $w_1, w > 0$ as follows:

$$\begin{aligned}w_1 = w_1(\mathbf{j},\chi) &= \sum_{\chi^* \neq \chi_0^* (\mathrm{mod}\ Q)} (|\alpha(\chi^*)| + 2|\beta(\chi^*)|);\\ w &= w_1 + C_0.\end{aligned}$$

Proposition 7.10. *Let* $s \in D \cap D_Q$. *Then*

$$A(s) \ll_Q \log^{w_1}(|t|+4)\exp(\omega_0\log\zeta(s)). \tag{7.12}$$

Furthermore,

$$A(s) \ll_Q \frac{1}{|s-1|^{\alpha_0}} \text{ if } |t| \leq 4, \tag{7.13}$$

$$A(s) \ll_Q \log^w |t| \text{ if } |t| \geq 4. \tag{7.14}$$

Proof. The inequality (7.12) follows from the above results, together with (7.4) and Lemma 7.6. From this, (7.13) follows, via Lemma 7.5. The inequality (7.14) follows from (7.12) and Lemma 7.4. □

Theorem 7.11. *If $q_1, \dots, q_r \geq 2$ are pairwise relatively prime, then there is a number $\delta > 0$, such that*

$$\sum_{\substack{n \leq x \\ n \equiv b(m) \\ \mathbf{e}(n) \equiv \mathbf{a}(\mathbf{q})}} 1 = \frac{x}{mq} + O\left(\frac{x}{\log^{\delta} x}\right)$$

Proof. By Proposition 7.10, we may apply Lemma 7.2 to $A(\mathbf{j}; s)$, for each $\mathbf{j} \neq \mathbf{0}$, taking $\alpha = \max\{0, \alpha_0\}$, and $K = K_Q$, since $\alpha_0 < 1$. Note that α_0 depends on $\mathbf{j}$. The theorem then follows by (7.11). □

Corollary 7.12. *Let $s(n)$ denote the sum of prime factors with repetition function, and let q be a fixed prime. Then*

$$\sum_{\substack{n \leq x \\ s(n) \equiv a(q)}} 1 = \frac{x}{q} + O_q\left(\frac{x}{\sqrt{\log x}}\right).$$

Proof. In this case, we have $m = 1 = r$, and $k_1 = 1$. Since q is prime, we have:

$$\omega_0 = \frac{1}{q-1} \sum_{b=1}^{q-1} \zeta_q^{bj} = \alpha_0 = -\frac{1}{q-1},$$

for any j not congruent to 0 modulo q. Thus we may take $\alpha = \max\{-1/(q-1), 0\} = 0$, and so $(1-\alpha)/2 = 1/2$. The Corollary follows from Lemma 7.2. □

7.4 The Distribution of s_k Modulo q

We now switch our focus to the elementary prime symmetric functions. It is tempting to assume that s_k must also be uniformly distributed modulo any modulus, but this is not the case. Our goal then, is to compute an asymptotic formula for $\#\{n \leq x : s_k(n) \equiv a(\text{mod } q)\}$ for some $q \geq 2$. To do so, we will proceed as we did in the previous section, using Perron's formula in conjunction with Lemma 7.2. Let $\zeta = e^{2\pi i/q}$. Then

$$\sum_{\substack{n \leq x \\ s_k(n) \equiv a(q)}} 1 = \frac{1}{q} \sum_{j=0}^{q-1} \zeta^{-aj} \sum_{n \leq x} \zeta^{j s_k(n)} \tag{7.15}$$

Thus we must study the sum $\sum_{n\leq x}\zeta^{js_k(n)}$. With this intent let

$$A(j;s)=\sum_{n=1}^{\infty}\zeta^{js_k(n)}n^{-s}.$$

7.4.1 The Dirichlet Series

Definition 7.13. *Define*

$$\kappa=\kappa(q,k)=q\prod_{\substack{p|q\\ p^\beta||k!}}p^\beta.$$

Note that for any $n\in\mathbb{N}$, $(q,n)=1$ if and only if $(\kappa,n)=1$.

Lemma 7.14. *For any $n\in\mathbb{N}$, and $k'\leq k$,*

$$\binom{n}{k'}\equiv\binom{n+\kappa(q,k)}{k'}\ (\textit{mod } q).$$

Proof. Let $\kappa=\kappa(q,k)$. Then there is a polynomial $p(x)\in\mathbb{Z}[x]$ such that

$$\binom{n+\kappa}{k'}=\binom{n}{k'}+\frac{\kappa}{k'!}p(n)\equiv\binom{n}{k'}\ (\text{mod } q).$$

□

Definition 7.15. *For $i,q\in\mathbb{N}$, let*

$$\Omega_{q,i}(n)=\#\{p|n \textit{ repetitions allowed} : p\equiv i\ (\textit{mod } q)\}.$$

Corollary 7.16. *Suppose that for $n\in\mathbb{N}$, $\Omega_{q,i}(n)\equiv k_i$ (mod κ), for $i=0,\ldots,q-1$, where k_i is chosen to be the least nonnegative residue modulo κ. Let $\sigma:\mathbb{Z}_\kappa^q\longrightarrow\mathbb{Z}_q$ be defined by*

$$\begin{aligned}\sigma(\mathbf{k})&=\sigma(k_0,\ldots,k_{q-1})\\&=\sum_{j_1+\cdots+j_{q-1}=k}\binom{k_1}{j_1}\cdots\binom{k_{q-1}}{j_{q-1}}1^{j_1}\cdots(q-1)^{j_{q-1}},\end{aligned}$$

Then

$$\sigma(\mathbf{k})\equiv s_k(n)\,(\textit{mod } q).$$

Proof. By the definition of s_k, we have modulo q:

$$\begin{aligned} s_k(n) &\equiv \sum_{j_1+\cdots+j_{q-1}=k} \binom{\Omega_{q,1}(n)}{j_1} \cdots \binom{\Omega_{q,q-1}(n)}{j_{q-1}} 1^{j_1}2^{j_2}\cdots(q-1)^{j_{q-1}} \\ &\equiv \sum_{j_1+\cdots+j_{q-1}=k} \binom{k_1}{j_1} \cdots \binom{k_{q-1}}{j_{q-1}} 1^{j_1}2^{j_2}\cdots(q-1)^{j_{q-1}}, \text{ by Lemma 7.14} \\ &= \sigma(\mathbf{k}). \end{aligned}$$

□

Remark 7.17. *Corollary 7.16 implies that the value of $s_k(n)$ modulo q depends only on the values of $\Omega_{q,i}(n)$ modulo κ. From this it follows that for any $m, n \in \mathbb{N}$, we have that $s_k(m^\kappa n) \equiv s_k(n) \pmod q$.*

The Dirichlet series $A(j;s)$ cannot be nicely expressed as an Euler product, since except for the case when $k = 1$, $\zeta^{js_k(n)}$ is not multiplicative. The key however, is to observe that $A(j;s)$ can be expressed as a linear combination of Euler products.

Proposition 7.18. *There exist constants $\alpha(j; i_0, \dots, i_{q-1})$, where $0 \le i_b \le \kappa - 1$ for $b = 0, \dots, q-1$ such that for $\sigma > 1$,*

$$A(j;s) = \tag{7.16}$$

$$\sum_{0\le i_0\le\kappa-1} \cdots \sum_{0\le i_{q-1}\le\kappa-1} \alpha(j; i_0, \dots, i_{q-1}) \prod_{b=0}^{q-1} \prod_{p\equiv b (mod\ q)} \left(\frac{1}{1-\zeta_\kappa^{i_b} p^{-s}}\right). \tag{7.17}$$

Denote the vector $(i_0, \dots, i_{q-1})$ by $\mathbf{i}$, and by $\mathbf{k}$ the vector $(k_0, \dots, k_{q-1})$, $0 \le k_i \le \kappa - 1$ for $i = 0, \dots, q-1$, and let $\sigma(\mathbf{k})$ be as in Corollary 7.16. Then we have

$$\alpha(j;\mathbf{i}) = \frac{1}{\kappa^q} \sum_{\mathbf{k}\in\mathbb{Z}_\kappa^q} \zeta_\kappa^{-\mathbf{k}\cdot\mathbf{i}} \zeta^{j\sigma(\mathbf{k})}.$$

Proof. Order variables $\alpha(j; i_0, \dots, i_{q-1})$ according to the base κ expansion of $i_0 \dots i_{q-1}$; that is, let $\alpha(j; i_0, \dots, i_{q-1})$ be in the $i_0 + i_1\kappa + \cdots + i_{q-1}\kappa^{q-1}$-th position, beginning at 0. Sums over $\mathbf{i}$ range over all values in $\mathbb{Z}_\kappa^q$.

Now consider the κ^q equations in the κ^q unknowns $\alpha(j;\mathbf{i})$ where the $k_0 + k_1\kappa + \cdots + k_{q-1}\kappa^{q-1}$-th equation is

$$\zeta^{j\sigma(\mathbf{k})} = \sum_{\mathbf{i}} \alpha(j;\mathbf{i}) \zeta_\kappa^{\mathbf{i}\cdot\mathbf{k}}. \tag{7.18}$$

Let A_1 denote the $\kappa \times \kappa$ Vandermonde matrix whose ij-th entry (beginning the row and column count at 0) is ζ_κ^{ij}.

By construction, the coefficient matrix of this system, which we shall denote by A, is the q-fold Kronecker product of A_1 with itself. A_1 and hence A are symmetric Butson-Hadamard matrices (cf. [6] and [7]). Hence,

$$AA^* = \kappa^q I,$$

where A^* is the conjugate matrix of A (equal in this case to the conjugate transpose).

If we express this system in the form $A\mathbf{x} = \mathbf{b}$, where

$$\mathbf{x} = \begin{pmatrix} \alpha(j;0,0,\ldots,0) \\ \alpha(j;1,0,\ldots,0) \\ \vdots \\ \alpha(j;\kappa-1,\kappa-1,\ldots,\kappa-1) \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} \zeta^{j\sigma(0,0,\ldots,0)} \\ \zeta^{j\sigma(1,0,\ldots,0)} \\ \vdots \\ \zeta^{j\sigma(\kappa-1,\kappa-1,\ldots,\kappa-1)} \end{pmatrix},$$

then

$$\mathbf{x} = \frac{1}{\kappa^q} A^* \mathbf{b},$$

which implies that

$$\alpha(j;\mathbf{i}) = \frac{1}{\kappa^q} \sum_{\mathbf{k} \in \mathbb{Z}_\kappa^q} \zeta_\kappa^{-\mathbf{k}\cdot\mathbf{i}} \zeta^{j\sigma(\mathbf{k})}$$

Now fix an n and suppose that $\Omega_{q,i}(n) \equiv k_i \pmod{\kappa}$ for $i = 0, \ldots, q-1$. Then $\zeta^{js_k(n)} = \zeta^{j\sigma(\mathbf{k})}$. On the right-hand side of (7.17), we see that the coefficient of n^{-s} is

$$\sum_{\mathbf{i}} \alpha(j;\mathbf{i}) \zeta_\kappa^{\mathbf{i}\cdot\mathbf{k}},$$

so equating the coefficients in (7.17) we obtain equations of the type (7.18). Hence the $\alpha(j;\mathbf{i})$ exist as required. □

Let $S = \{n \in \mathbb{N} : 1 \le n \le q-1, (n,q) = 1\}$, and let

$$F(\mathbf{i};s) = \prod_{p|q} \frac{1}{1 - \zeta_\kappa^{i_p} p^{-s}}.$$

Then by the previous proposition and Lemma 7.7 we can write

$$\begin{aligned} A(j;s) &= \sum_{\mathbf{i}\in\mathbb{Z}_\kappa^q} \alpha(j;\mathbf{i})F(\mathbf{i};s)\prod_{b\in S}\prod_{p\equiv b \pmod q}\left(\frac{1}{1-\zeta_\kappa^{i_b}p^{-s}}\right) \\ &= \sum_{\mathbf{i}\in\mathbb{Z}_\kappa^q} \alpha(j;\mathbf{i})F(\mathbf{i};s)\exp\left[\sum_{\chi \pmod q}\omega(\mathbf{i};\chi)\log L(s,\chi)\right]\exp(\varphi(\mathbf{i};s)), \end{aligned} \tag{7.19}$$

where $\varphi(\mathbf{i};s) \ll_\epsilon 1$ for $\sigma \geq 1/2+\epsilon$, and

$$\omega(\mathbf{i};\chi) = \frac{1}{\phi(q)}\sum_{b\in S}\zeta_\kappa^{i_b}\bar{\chi}(b) = \alpha(\mathbf{i};\chi) + i\beta(\mathbf{i};\chi).$$

Remark 7.19. *Note that $|\omega(\mathbf{i};\chi)| \leq 1$. The important question though is whether or not $\omega(\mathbf{i};\chi_0) = 1$, since these will be the exponents of $\zeta(s)$ in the expression for $A(j;s)$. This occurs precisely when $i_b = 0$ for all $b \in S$. Thus there are $\kappa^{q-\phi(q)}$ values of $\mathbf{i}$ for which this takes place.*

Notation 7.20.

$$\begin{aligned} Z &= \{\mathbf{i}\in\mathbb{Z}_\kappa^q : i_b = 0 \textit{ for all } b\in S\}, \\ Z' &= \mathbb{Z}_\kappa^q\backslash Z, \\ \alpha &= \max\{0,\alpha(\mathbf{i};\chi_0) : \mathbf{i}\in Z'\}, \\ T &= \exp\left((\log x)^{(1+\alpha^2)/2}\right) \\ c &= 1+\frac{1}{\log x} \\ G(\mathbf{i};s) &= \prod_{p|q}\frac{1-p^{-s}}{1-\zeta_\kappa^{i_p}p^{-s}}. \end{aligned}$$

Note that $0 \leq \alpha < 1$. Now we can express $A(j;s)$ as follows:

$$\begin{aligned} A(j;s) =& \sum_{\mathbf{i}\in Z}\alpha(j;\mathbf{i})G(\mathbf{i};s)\zeta(s) \\ &+ \sum_{\mathbf{i}\in Z'}\alpha(j;\mathbf{i})F(\mathbf{i};s)\exp\left[\sum_{\chi \pmod q}\omega(\mathbf{i};\chi)\log L(s,\chi)\right]\exp(\varphi(\mathbf{i};s)). \end{aligned} \tag{7.20}$$

It is easily seen that for $\mathbf{i} \in Z'$ the function

$$F(\mathbf{i};s)\exp\left[\sum_{\chi \pmod q} \omega(\mathbf{i};\chi)\log L(s,\chi)\right]\exp(\varphi(\mathbf{i};s))$$

satisfies the conditions of Lemma 7.2 with respect to the given α. Hence by that same Lemma, we may conclude that

$$\int_{c-iT}^{c+iT}\left(\frac{x^s}{s}\right)F(\mathbf{i};s)\exp\left[\sum_{\chi \pmod q} \omega(\mathbf{i};\chi)\log L(s,\chi)\right]\exp(\varphi(\mathbf{i};s))\,ds \ll \frac{x}{(\log x)^{(1-\alpha)/2}}. \tag{7.21}$$

On the other hand, if $\mathbf{i} \in Z$, then

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\left(\frac{x^s}{s}\right)G(\mathbf{i};s)\zeta(s)\,ds = Res\left(\left(\frac{x^s}{s}\right)G(\mathbf{i};s)\zeta(s); s=1\right) + \frac{1}{2\pi i}\int_{\gamma}\left(\frac{x^s}{s}\right)G(\mathbf{i};s)\zeta(s)\,ds,$$

where γ is the union of the following three paths:

$$\gamma_1 : \text{line segment from } c-iT \text{ to } 1-\frac{K}{\log(T+4)}-iT$$

$$\gamma_2 : \text{Along } 1-\frac{K}{\log(|t|+4)}+it \text{ for } -T\le t\le T$$

$$\gamma_3 : \text{line segment from } 1-\frac{K}{\log(T+4)}+iT \text{ to } c+iT$$

Using arguments analogous to Lemma 7.2, together with the bounds established on $\zeta(s)$ and the fact that the path γ is bounded away from the pole at $s=1$, we have that

$$\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\left(\frac{x^s}{s}\right)G(\mathbf{i};s)\zeta(s)\,ds = xG(\mathbf{i};1)+O\left(\frac{x}{(\log x)^{(1-\alpha)/2}}\right). \tag{7.22}$$

The order of magnitude of the error term is in fact much better than this, but it is all we require. Note that the implicit constants here depend on q.

Now, by (7.20), (7.21), (7.22) and Lemma 7.1 we obtain

$$\begin{aligned}
\sum_{n\le x}\zeta^{js_k(n)} &= \frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{A(j;s)x^s}{s}\,ds+O\left(\frac{x\log x}{T}\right)\\
&=\frac{1}{2\pi i}\sum_{\mathbf{i}\in Z}\alpha(j;\mathbf{i})\int_{c-iT}^{c+iT}\left(\frac{x^s}{s}\right)G(\mathbf{i};s)\zeta(s)\,ds\\
&\quad+\frac{1}{2\pi i}\sum_{\mathbf{i}\in Z'}\alpha(j;\mathbf{i})\\
&\quad\times\int_{c-iT}^{c+iT}\left(\frac{x^s}{s}\right)F(\mathbf{i};s)\exp\left[\sum_{\chi \pmod q}\omega(\mathbf{i};\chi)\log L(s,\chi)\right]\exp(\varphi(\mathbf{i};s))\,ds\\
&\quad+O\left(\frac{x\log x}{T}\right)\\
&=\sum_{\mathbf{i}\in Z}\alpha(j;\mathbf{i})xG(\mathbf{i};1)+O\left(\frac{x}{(\log x)^{(1-\alpha)/2}}\right). \qquad (7.23)
\end{aligned}$$

Theorem 7.21. *There is a $\delta>0$ such that as $x\to\infty$,*

$$\sum_{\substack{n\le x\\ s_k(n)\equiv a\,(q)}}1=\frac{x}{q\kappa^q}\sum_{j=0}^{q-1}\sum_{\substack{\mathbf{i}\in\mathbb{Z}_\kappa^q\\ (b,q)=1\Rightarrow i_b=0}}\sum_{\mathbf{k}\in\mathbb{Z}_\kappa^q}\zeta^{(\sigma(\mathbf{k})-a)j}\zeta_\kappa^{-\mathbf{i}\cdot\mathbf{k}}\prod_{p|q}\left(\frac{1-p^{-1}}{1-\zeta_\kappa^{i_p}p^{-1}}\right)+O\left(\frac{x}{\log^\delta x}\right),$$

where if q is prime, we interpret i_q to be i_0.

Proof. Using Proposition 7.18, and combining equations (7.15) and (7.23) we have:

$$\begin{aligned}
\sum_{\substack{n\le x\\ s_k(n)\equiv a\,(q)}}1 &=\frac{1}{q}\sum_{j=0}^{q-1}\zeta^{-aj}\left[\sum_{\mathbf{i}\in Z}\alpha(j;\mathbf{i})xG(\mathbf{i};1)+O\left(\frac{x}{(\log x)^{(1-\alpha)/2}}\right)\right]\\
&=\frac{x}{q}\sum_{j=0}^{q-1}\zeta^{-aj}\sum_{\mathbf{i}\in Z}\alpha(j;\mathbf{i})G(\mathbf{i};1)+O\left(\frac{x}{(\log x)^{(1-\alpha)/2}}\right)\\
&=\frac{x}{q\kappa^q}\sum_{j=0}^{q-1}\sum_{\mathbf{i}\in Z}\sum_{\mathbf{k}\in\mathbb{Z}_\kappa^q}\zeta^{(\sigma(\mathbf{k})-a)j}\zeta_\kappa^{-\mathbf{i}\cdot\mathbf{k}}\prod_{p|q}\left(\frac{1-p^{-1}}{1-\zeta_\kappa^{i_p}p^{-1}}\right)\\
&\quad+O\left(\frac{x}{(\log x)^{(1-\alpha)/2}}\right).
\end{aligned}$$

□

7.5 Examples

The process outlined in the last section can be simplified (and elucidated) with examples.

7.5.1 s_2 modulo 3

It transpires that s_2 is uniformly distributed modulo 3 in the sense that

$$\sum_{n \le x} \zeta^{s_2(n)} = o(x),$$

where $\zeta = e^{2\pi i/3}$. Let

$$A(s) = \sum_{n=1}^{\infty} \zeta^{s_2(n)} n^{-s}.$$

Since $s_2(3^r n) \equiv s_2(n) \pmod 3$, we may write

$$A(s) = \frac{1}{1-3^{-s}} \sum_{3 \nmid n}^{\infty} \zeta^{s_2(n)} n^{-s}.$$

For $\mathbf{i} = (i_1, i_2) \in \mathbb{Z}_3^2$, there are constants $\alpha(\mathbf{i})$ such that

$$\sum_{3 \nmid n}^{\infty} \zeta^{s_2(n)} n^{-s} = \sum_{\mathbf{i} \in \mathbb{Z}_3^2} \alpha(\mathbf{i}) \prod_{p \equiv 1\,(3)} \frac{1}{1-\zeta^{i_1} p^{-s}} \prod_{p \equiv 2\,(3)} \frac{1}{1-\zeta^{i_2} p^{-s}}.$$

If $3 \nmid n$, then

$$s_2(n) \equiv \sum_{j_1+j_2=2} \binom{\Omega_{3,1}(n)}{j_1} \binom{\Omega_{3,2}(n)}{j_2} 2^{j_2} \pmod 3.$$

Equating the coefficient of n^{-s}, we obtain a system of equations in the

$\alpha(\mathbf{i})$, $A\mathbf{x} = \mathbf{b}$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 & 1 & \zeta & \zeta^2 & 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta & 1 & \zeta^2 & \zeta & 1 & \zeta^2 & \zeta \\ 1 & 1 & 1 & \zeta & \zeta & \zeta & \zeta^2 & \zeta^2 & \zeta^2 \\ 1 & \zeta & \zeta^2 & \zeta & \zeta^2 & 1 & \zeta^2 & 1 & \zeta \\ 1 & \zeta^2 & \zeta & \zeta & 1 & \zeta^2 & \zeta^2 & \zeta & 1 \\ 1 & 1 & 1 & \zeta^2 & \zeta^2 & \zeta^2 & \zeta & \zeta & \zeta \\ 1 & \zeta & \zeta^2 & \zeta^2 & 1 & \zeta & \zeta & \zeta^2 & 1 \\ 1 & \zeta^2 & \zeta & \zeta^2 & \zeta & 1 & \zeta & 1 & \zeta^2 \end{pmatrix} \begin{pmatrix} \alpha(0,0) \\ \alpha(1,0) \\ \alpha(2,0) \\ \alpha(0,1) \\ \alpha(1,1) \\ \alpha(2,1) \\ \alpha(0,2) \\ \alpha(1,2) \\ \alpha(2,2) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \zeta \\ 1 \\ \zeta^2 \\ \zeta^2 \\ \zeta \\ \zeta^2 \\ \zeta \end{pmatrix}.$$

The vector $\mathbf{b}$ is obtained by considering the values of $\zeta^{\sigma(\mathbf{k})}$, where $\mathbf{k} = (k_1, k_2) \in \mathbb{Z}_3^2$ corresponds to $(\Omega_{3,1}(n), \Omega_{3,2}(n))$, and

$$\sigma(\mathbf{k}) = \sum_{j_1+j_2=2} \binom{k_1}{j_1}\binom{k_2}{j_2} 2^{j_2}.$$

The coefficient matrix A satisfies

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta & \zeta^2 \\ 1 & \zeta^2 & \zeta \end{pmatrix}.$$

It is a symmetric Butson Hadamard matrix and hence satisfies $AA^* = 9I$, where A^* is its conjugate matrix. Using this it is easy to solve the system and obtain $\alpha(0,2) = \alpha(2,0) = \frac{1}{2} + \frac{i\sqrt{3}}{6}$, $\alpha(1,1) = \frac{-i\sqrt{3}}{3}$, and the rest to be 0.

This enables us to write

$$\begin{aligned} A(s) = {} & \left(\frac{1}{2} + \frac{i\sqrt{3}}{6}\right) \exp\left[\left(\frac{\zeta^2 - 1}{2}\right) \log\left(1 - 3^{-s}\right)\right] \exp\left[-\left(\frac{\zeta}{2}\right) \log \zeta(s)\right] \\ & \times \sum_{j=1}^{2} \exp\left[(-1)^j \left(\frac{\zeta^2 - 1}{2}\right) \log L(s, \chi_1)\right] \exp\left(\varphi_j(s)\right) \\ & - \frac{i\sqrt{3}}{3} \exp\left((\zeta - 1)\log(1 - 3^{-s})\right) \exp\left(\zeta \log \zeta(s)\right) \exp\left(\varphi_3(s)\right), \end{aligned}$$

where χ_1 is the non-principal character (mod 3), and the $\varphi_j(s)$ are bounded as in Lemma 7.7.

Since every term of the form $\exp(c \log \zeta(s))$ occurs only with the real part of c less than 1, everything can be suitably bounded, and we can prove that $\sum_{n \le x} \zeta^{s_2(n)} = o(x)$.

7.5.2 s_3 modulo 2

This example will be a case in which uniform distribution does not occur. Let $A(s) = \sum_{n=1}^{\infty}(-1)^{s_3(n)}n^{-s}$. It transpires that $s_3(n) \pmod 2$ is determined by $\Omega_{2,1}(n) \pmod 4$. In fact,

$$s_3(n) \equiv \binom{\Omega_{2,1}(n)}{3} \pmod 2.$$

There are constants $\alpha(j)$ for $j \in \mathbb{Z}_4$ such that

$$A(s) = \frac{1}{1-2^{-s}} \sum_{j=0}^{3} \alpha(j) \prod_{p \neq 2} \frac{1}{1 - i^j p^{-s}}.$$

Equating coefficients in the Dirichlet series yields the system

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix} \begin{pmatrix} \alpha(0) \\ \alpha(1) \\ \alpha(2) \\ \alpha(3) \end{pmatrix} = \begin{pmatrix} (-1)^{\binom{0}{3}} \\ (-1)^{\binom{1}{3}} \\ (-1)^{\binom{2}{3}} \\ (-1)^{\binom{3}{3}} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix},$$

with solutions

$$\begin{pmatrix} \alpha(0) \\ \alpha(1) \\ \alpha(2) \\ \alpha(3) \end{pmatrix} = \begin{pmatrix} 1/2 \\ -i/2 \\ 1/2 \\ i/2 \end{pmatrix}.$$

This enables us to write

$$\begin{aligned} A(s) =& \frac{1}{2}\zeta(s) - \frac{i}{2} \exp\left((i-1)\log(1-2^{-s})\right) \exp\left(i \log \zeta(s)\right) \exp(\varphi_1(s)) \\ &+ \left(\frac{1}{2}\right)\left(\frac{1+2^{-s}}{1-2^{-s}}\right) \frac{\zeta(2s)}{\zeta(s)} \\ &+ \frac{i}{2} \exp\left((-i-1)\log(1-2^{-s})\right) \exp\left(-i \log \zeta(s)\right) \exp(\varphi_3(s)). \end{aligned}$$

Note the term $\frac{1}{2}\zeta(s)$. Thus taking $T = \exp\left(\sqrt{\log x}\right)$, we can show

$$\begin{aligned}
\sum_{\substack{n\le x \\ s_3(n)\equiv 0\,(2)}} 1 &= \sum_{n\le x} \frac{1+(-1)^{s_3(n)}}{2} \\
&= \frac{\lfloor x \rfloor}{2} + \frac{1}{2}\sum_{n\le x}(-1)^{s_3(n)} \\
&\sim \frac{x}{2} + \left(\frac{1}{2}\right)\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{A(s)x^s}{s}\,ds \\
&\sim \frac{x}{2} + \left(\frac{1}{2}\right)\frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{\zeta(s)}{2}\frac{x^s}{s}\,ds \\
&\sim \frac{x}{2} + \left(\frac{1}{2}\right)\operatorname{Res}\left(\frac{\zeta(s)}{2}\frac{x^s}{s}; s=1\right) \\
&= \frac{x}{2} + \frac{x}{4} = \frac{3x}{4}.
\end{aligned}$$

This density result is confirmed by empirical data:

Table 3. $s_3(n)$-Modulo 2

x	$\#\{n \le x : s_3(n) \equiv 0 \pmod 2\}$	$\frac{1}{x}\#\{n \le x : s_3(n) \equiv 0 \pmod 2\}$
100	93	.93
1000	836	.836
5000	3954	.7908
10000	7743	.7743
20000	15232	.7616
30000	22654	.75513
50000	37350	.747

7.5.3 s_k modulo a prime q

The result of Theorem 7.21 simplifies considerably if we restrict q to being prime. In this case, we take $\kappa = q^{\beta+1}$, where $q^\beta || k!$, and we define $\sigma : \mathbb{Z}_\kappa^{q-1} \longrightarrow \mathbb{Z}_q$ by

$$\begin{aligned}\sigma(\mathbf{k}) &= \sigma(k_1, \dots, k_{q-1}) \\ &= \sum_{j_1+\cdots+j_{q-1}=k} \binom{k_1}{j_1} \cdots \binom{k_{q-1}}{j_{q-1}} 1^{j_1} \cdots (q-1)^{j_{q-1}}.\end{aligned}$$

If $\zeta = e^{2\pi i/q}$, then

$$\sum_{\substack{n \le x \\ s_k(n) \equiv a\,(q)}} 1 \sim \frac{x}{q\kappa^{q-1}} \sum_{j=0}^{q-1} \sum_{\mathbf{k} \in \mathbb{Z}_\kappa^{q-1}} \zeta^{(\sigma(\mathbf{k})-a)j}.$$

Chapter 8

Partitions into Prime Powers

8.1 Introduction

In this chapter we look at partitions of positive integers into fixed powers of prime numbers. For a given underlying set $A \subseteq \mathbb{N}$, and a positive integer $n \in \mathbb{N}$, we denote the number of partitions of n with parts taken from A by $p_A(n)$. For the most part, we will be taking $A = \mathbb{P}^{(r)}$, by which we mean the set of rth powers of prime numbers. The asymptotic formula

$$\log p_{\mathbb{P}^{(r)}}(n) \sim (r+1)\left[\Gamma\left(\frac{1}{r}+2\right)\zeta\left(\frac{1}{r}+1\right)\right]^{r/(r+1)} n^{1/(r+1)}(\log n)^{-r/(r+1)}, \tag{8.1}$$

where

$$\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}\,dt$$

is the gamma function, was first proved by Hardy and Ramanujan [19] in 1916. Further work has been done by Mitsui [31] and Grosswald [17]. Kerawala [25] attempts an elementary proof of (8.1) for $r = 1$, but his proof contains errors. We begin by adapting work of Bateman and Erdős on the eventual monotonicity of a large class of partition functions to a case of interest.

8.2 Monotonicity

In this section we follow the work of the author [49] previously published in the Journal of Integer Sequences, and that of Bateman and Erdős [4]. First let us revisit some important notation. The k-th difference function $p_A^{(k)}(n)$ can defined inductively as follows: for $k = 0$, $p_A^{(k)}(n) = p_A(n)$. If $k > 0$, then

$$p_A^{(k)}(n) = p_A^{(k-1)}(n) - p_A^{(k-1)}(n-1).$$

Let $f_A^{(k)}(x)$ be the generating function for $p_A^{(k)}(n)$. We have [4] the following power series identity:

$$f_A^{(k)}(x) = \sum_{n=0}^{\infty} p_A^{(k)}(n)x^n \tag{8.2}$$

$$= (1-x)^k \sum_{n=0}^{\infty} p_A(n)x^n \tag{8.3}$$

$$= (1-x)^k \prod_{a\in A} \frac{1}{1-x^a}. \tag{8.4}$$

This may be used to define $p_A^{(k)}(n)$ for all $k \in \mathbb{Z}$, including $k < 0$.

Bateman and Erdős [4] characterize the sets A for which $p_A^{(k)}(n)$ is ultimately nonnegative. Note that if $k < 0$, then the power series representation of $(1-x)^k$ has nonnegative coefficients so that $p_A^{(k)}(n) \geq 0$. For $k \geq 0$, they prove the following: if A satisfies the property that whenever k elements are removed from it, the remaining elements have greatest common divisor 1, then

$$\lim_{n\to\infty} p_A^{(k)}(n) = \infty.$$

A simple consequence of this is the fact that $p_A(n)$ is eventually monotonic if $k > 0$.

No explicit bounds for when $p_A^{(k)}(n)$ ceases to take on negative values are included with the result of Bateman and Erdős [4]. By following their approach but specializing to the case when $A = \mathbb{P}$, we shall find bounds for n depending on k which guarantee that $p_{\mathbb{P}}^{(k)}(n) \geq 0$.

In a subsequent paper, Bateman and Erdős [3] prove that $p_{\mathbb{P}}^{(1)}(n) \geq 0$ for all $n \geq 2$. That is, the sequence of partitions of n into primes is weakly increasing for $n \geq 1$. We have shown [49] that if $A = \mathbb{P}^{(\ell)}$ is the set of ℓ-th powers of primes, and $F(k,\ell) = \min\{N \in \mathbb{N} : n \geq N \text{ implies that } p_A^{(k)}(n) > 0\}$, then for a fixed $\ell \in \mathbb{N}$,

$$\log F(k,\ell) = o\left((k+2)^{8\ell k}\right),$$

as $k \to \infty$. For the sake of clarity, we have decided to restrict our attention to the set of prime numbers.

In a series of papers (c.f. [36]- [39]), L. B. Richmond studies the asymptotic behaviour of partition functions and their differences for sets satisfying certain stronger conditions. The results nonetheless apply in the cases of $A = \mathbb{P}^{(\ell)}$. Richmond finds [38] an asymptotic formula in n for $p_A^{(k)}(n)$. Unfortunately, his formula is not useful towards finding bounds for when $p_A^{(k)}(n)$ must be positive, since, as is customary, he does not include explicit constants in the error term. Furthermore, his asymptotic formula includes functions such as $\alpha = \alpha(n)$ defined by the solution to the equation

$$n = \sum_{a \in A} \frac{a}{e^{\alpha a} - 1} - \frac{k}{e^{\alpha} - 1}.$$

As we are seeking explicit constants, a direct approach will be cleaner than than attempting to adapt the aforementioned formula.

Another result worthy of comment from Richmond [38] pertains to a conjecture of Bateman and Erdős [4]. His result applies to $A = \mathbb{P}^{(\ell)}$, and states that

$$\frac{p_A^{(k+1)}(n)}{p_A^{(k)}(n)} = O(n^{-1/2}), \text{ as } n \to \infty.$$

We shall write ζ_n for the primitive n-th root of unity $e^{2\pi i/n}$, and the m-th prime by p_m. We will have cause to use the so-called “primorial” notation, that is, let

$$n\# = \prod_{p \le n} p,$$

where the product is taken over all primes less than or equal to n.

We have two main results. The first is Theorem 8.13 in which we demonstrate that for a given $k \in \mathbb{N}$, there exist h, $h_1 \in \mathbb{N}$ depending on k such that if $n \ge h + h_1 - 1$, then $p_{\mathbb{P}}^{(k)}(n) \ge 0$. In Notation 8.9 (p. 91) shall determine explicit formulae for h and h_1 in terms of k. Secondly, in Theorem 8.14 we show that if

$$N(k) = \min\{N \in \mathbb{N} : p_{\mathbb{P}}^{(k)}(n) \ge 0 \text{ for all } n \ge N\}, \text{ then as } k \to \infty,$$
$$\log N(k) \ll k^{k(4+\log 16)\left(1+O\left(\frac{\log\log k}{\log k}\right)\right)}.$$

Following Bateman and Erdős [4], we first tackle the finite case.

Lemma 8.1. *Let B be a finite subset of $\mathbb{P}$ of size r, and suppose that $k < r$. The function $p_B^{(k)}(n)$ can be decomposed as follows:*

$$p_B^{(k)}(n) = g_B^{(k)}(n) + \psi_B^{(k)}(n),$$

where $g_B^{(k)}(n)$ is a polynomial in n of degree $r-k-1$ with leading coefficient $\left((r-k-1)!\prod_{p\in B} p\right)^{-1}$, and $\psi_B^{(k)}(n)$ is periodic in n with period $\prod_{p\in B} p$.

Proof. We use partial fractions to decompose the generating function $f_B^{(k)}(x)$ as follows:

$$f_B^{(k)}(x) = \frac{1}{(1-x)^{r-k}} \prod_{p\in B} \prod_{j=1}^{p-1} \frac{1}{1-\zeta_p^j x} \tag{8.5}$$

$$= \frac{\alpha_1}{1-x} + \frac{\alpha_2}{(1-x)^2} + \ldots + \frac{\alpha_{r-k}}{(1-x)^{r-k}} + \sum_{p\in B} \sum_{j=1}^{p-1} \frac{\beta(\zeta_p^j)}{1-\zeta_p^j x}, \tag{8.6}$$

where the α_i, and $\beta(\zeta_p^j)$ are complex numbers depending on k and the set B. Note that

$$\alpha_{r-k} = \left(\prod_{p\in B} p\right)^{-1}.$$

The power series expansion for $(1-x)^{-h}$ is given by

$$\frac{1}{(1-x)^h} = \sum_{n=0}^{\infty} \binom{n+h-1}{h-1} x^n.$$

Hence, if

$$g_B^{(k)}(n) = \sum_{h=1}^{r-k} \alpha_h \binom{n+h-1}{h-1},$$

and

$$\psi_B^{(k)}(n) = \sum_{q\in B} \sum_{j=1}^{q-1} \beta(\zeta_q^j)\zeta_q^{jn},$$

then the lemma is proved. □

For the remainder of this section, $g_B^{(k)}(n)$ and $\psi_B^{(k)}(n)$ shall be as in Lemma 8.1, for a given finite set $B \subseteq \mathbb{P}$ which shall be clear from the context.

Remark 8.2. *Let B be as in Lemma 8.1. We wish to know the precise value of $\beta(\zeta_q^j)$. To simplify notation a little, write β_ζ instead, when ζ is clear from* *the context. In particular, suppose* $\eta = \zeta_q^j$*, where* $q \in B$*, and* $0 < j < q$*. Then*

$$\begin{aligned}(1-\eta x)f_B^{(k)}(x) &= (1-x)^k \frac{1}{1+\eta x+\cdots+(\eta x)^{q-1}} \prod_{\substack{p\in B\\ p\neq q}} \frac{1}{1-x^p} \\ &= \beta_\eta + (1-\eta x)\left(\frac{\alpha_1}{1-x}+\cdots+\frac{\alpha_{r-k}}{(1-x)^{r-k}}+\sum_{\zeta\neq\eta}\frac{\beta_\zeta}{1-\zeta x}\right),\end{aligned}$$

hence,

$$\beta_\eta = \frac{(1-\bar{\eta})^k}{q} \prod_{\substack{p\in B\\ p\neq q}} \frac{1}{1-\bar{\eta}^p}.$$

The following inequalities shall be useful:

$$1-\frac{\theta^2}{2} \le \cos\theta \le 1-\frac{\theta^2}{2}+\frac{\theta^4}{24},$$

which holds for all values of θ. Note that

$$|e^{i\theta}-1| = \sqrt{2(1-\cos\theta)},$$

so for $-2\sqrt{3} \le \theta \le 2\sqrt{3}$,

$$|\theta|\sqrt{1-\frac{\theta^2}{12}} \le |e^{i\theta}-1| \le |\theta|.$$

In particular, for $0 \le \theta \le \pi$,

$$\theta\sqrt{1-\frac{\theta^2}{12}} \le |e^{i\theta}-1| \le \theta. \tag{8.7}$$

Lemma 8.3. *Suppose that* $\zeta \neq 1$ *is a* q*-th root of unity,* $q \geq 2$*, and let* $b_0 = 2\pi\sqrt{1-\frac{\pi^2}{12}}$*. Then*

$$\frac{b_0}{q} \le |1-\zeta| \le 2.$$

Proof. Clearly $|1-\zeta| \le 2$. On the other hand, by the inequality (8.7),

$$\begin{aligned}|1-\zeta| &\ge |1-e^{2\pi i/q}| \\ &\ge \frac{2\pi}{q}\sqrt{1-\frac{4\pi^2}{12q^2}} \\ &\ge \frac{b_0}{q}.\end{aligned}$$

□

We shall continue to use the value for b_0 given in Lemma 8.3 throughout this section. Note that $b_0 \approx 2.647398833$.

Lemma 8.4. *Suppose that $k < r$, and $B \subseteq \mathbb{P}$, satisfies $|B| = r$. Suppose further that $\eta = \zeta_q^j$, for some $q \in B$, $j \in \{1, \ldots, q-1\}$. Then*

$$|\beta_\eta| \leq \begin{cases} \frac{2^k q^{r-2}}{b_0^{r-1}}, & \text{if } k \geq 0; \\ \frac{q^{r-k-2}}{b_0^{r-k-1}}, & \text{if } k < 0. \end{cases}$$

Proof. Making use of Remark 8.2 and Lemma 8.3 we have that for $k \geq 0$:

$$\begin{aligned} |\beta_\eta| &= \frac{|1-\zeta_q^{-j}|^k}{q} \prod_{\substack{p \in B \\ p \neq q}} \frac{1}{|1-\zeta_q^{-jp}|} \\ &\leq \frac{2^k}{q} \prod_{\substack{p \in B \\ p \neq q}} \frac{1}{|1-\zeta_q|} \\ &\leq \frac{2^k}{q} \left(\frac{q}{b_0}\right)^{r-1} \\ &= \frac{2^k q^{r-2}}{b_0^{r-1}}. \end{aligned}$$

A similar arguments works for the case when $k < 0$. □

Theorem 8.5. *Suppose that $k < r$, and $B \subseteq \mathbb{P}$, satisfies $|B| = r$. Then*

$$|\psi_B^{(k)}(n)| \leq \begin{cases} \frac{2^k}{b_0^{r-1}} \sum_{p \in B} p^{r-1}, & \text{if } k \geq 0; \\ \frac{1}{b_0^{r-k-1}} \sum_{p \in B} p^{r-k-1}, & \text{if } k < 0. \end{cases}$$

Proof. First assume that $k \geq 0$. By Lemmas 8.4 and 8.1,

$$\begin{aligned}
|\psi_B^{(k)}(n)| &= \left| \sum_{p \in B} \sum_{j=1}^{p-1} \beta(\zeta_p^j) \zeta_p^{jn} \right| \\
&\leq \sum_{p \in B} \sum_{j=1}^{p-1} |\beta(\zeta_p^j)| \\
&\leq \sum_{p \in B} \sum_{j=1}^{p-1} \frac{2^k p^{r-2}}{b_0^{r-1}} \\
&= \frac{2^k}{b_0^{r-1}} \sum_{p \in B} (p-1) p^{r-2} \\
&\leq \frac{2^k}{b_0^{r-1}} \sum_{p \in B} p^{r-1}.
\end{aligned}$$

A similar argument works for $k < 0$. □

To obtain bounds for the coefficients α_h, we will use Laurent series.

Lemma 8.6. *Suppose that $k < r$, and $B \subseteq \mathbb{P}$, satisfies $|B| = r$. Denote the largest element of B by Q, and suppose further that $0 < r_0 < |1 - \zeta_Q|$. Let*

$$d_B(r_0) = \prod_{p \in B} \prod_{j=1}^{p-1} (|\zeta_p^j - 1| - r_0).$$

Then

$$|\alpha_h| \leq \frac{1}{r_0^{r-k-h} d_B(r_0)}.$$

Proof. Let γ be the circle $|z - 1| = r_0$. From equations (8.5) and (8.6), and the Laurent Expansion Theorem, we have that

$$\begin{aligned}
|\alpha_h| &= \left| \frac{1}{2\pi i} \int_\gamma f_B^{(k)}(z)(z-1)^{h-1}\, dz \right| \\
&\leq \frac{1}{2\pi} \int_\gamma \frac{1}{|z-1|^{r-k-h+1}} \prod_{p \in B} \prod_{j=1}^{p-1} \frac{1}{|1 - \zeta_p^j z|}\, dz \\
&\leq \frac{1}{r_0^{r-k-h} d_B(r_0)}.
\end{aligned}$$

□

Proposition 8.7. *For $k \leq 0$, and $D_1 \subseteq D_2 \subseteq \mathbb{N}$, we have $p_{D_2}^{(k)}(n) \geq p_{D_1}^{(k)}(n) \geq 0$.*

Proof. This follows immediately from equation 8.3 and the fact that for $k \leq 0$, the power series expansion for $(1-x)^k$ has nonnegative coefficients. □

For the sake of clarity and the comprehensiveness of this exposition we include the following proposition, suitably adapted to our requirements from Bateman and Erdős [4].

Proposition 8.8. *Let $D \subseteq \mathbb{N}$ be an infinite set. For any $t \in \mathbb{N}$, we have that*

$$\frac{p_D(n)}{p_D^{(-1)}(n)} \leq \frac{1}{t+1} + \frac{(t-1)^2}{t+1}\frac{n^{2t-3}}{p_D^{(-1)}(n)}.$$

Proof. Denote by $P_q(n)$, the number of partitions of n into parts from D such that there are exactly q distinct parts. $P_q(n)$ has generating function

$$\sum_{n=0}^{\infty} P_q(n)x^n = \sum_{\{a_1,\dots a_q\}\subseteq D} \frac{x^{a_1}}{1-x^{a_1}} \cdots \frac{x^{a_q}}{1-x^{a_q}}.$$

If $R_q(m)$ is defined by

$$\sum_{m=0}^{\infty} R_q(m)x^m = \sum_{\{a_1,\dots a_q\}\subseteq [n]} \frac{x^{a_1}}{1-x^{a_1}} \cdots \frac{x^{a_q}}{1-x^{a_q}},$$

where $[n] = \{1,\dots,n\}$, then $P_q(n) \leq R_q(n)$.

There are $\binom{n}{q}$ subsets of $[n]$ of size q. Also, the coefficient of x^n in

$$(x^{a_1} + x^{2a_1} + \cdots)\cdots(x^{a_q} + x^{2a_q} + \cdots)$$

is less than or equal to the coefficient of x^n in

$$(x + x^2 + \cdots)^q = \sum_{m=q}^{\infty} \binom{m-1}{q-1} x^m,$$

so

$$P_q(n) \leq \binom{n}{q}\binom{n-1}{q-1} \leq n^{2q-1}.$$

Any partition $n = n_1a_1 + \cdots n_qa_q$, where $a_1, \ldots, a_q \in D$, gives rise to a partition of $n - a_i$ for $i = 1, \ldots, q$, namely

$$\begin{aligned} n - a_1 &= (n_1 - 1)a_1 + n_2a_2 + \cdots + n_qa_q, \\ n - a_2 &= n_1a_1 + (n_2 - 1)a_2 + \cdots + n_qa_q, \\ &\vdots \\ n - a_q &= n_1a_1 + n_2a_2 + \cdots + (n_q - 1)a_q. \end{aligned}$$

Note that no two distinct partitions of n can give rise to the same partition of any $m < n$ in this way, and so

$$\sum_{q=1}^{n} qP_q(n) \leq \sum_{m=0}^{n-1} p_D(m).$$

Now if $t \in \mathbb{N}$, then

$$\begin{aligned} p_D^{(-1)}(n) &= \sum_{m=0}^{n} p_D(m) \\ &\geq p_D(n) + \sum_{q=1}^{n} qP_q(n) \\ &= (t+1)p_D(n) + \sum_{q=1}^{n} (q-t)P_q(n) \\ &\geq (t+1)p_D(n) - (t-1)\sum_{q=1}^{t-1} P_q(n) \\ &\geq (t+1)p_D(n) - (t-1)^2 n^{2t-3}, \end{aligned}$$

and the theorem is proved. □

A consequence of Proposition 8.8 is that for an infinite set D,

$$\lim_{n\to\infty} \frac{p_D(n)}{p_D^{(-1)}(n)} = 0.$$

We are now going to restrict our attention to some more specific subsets of $\mathbb{P}$.

Notation 8.9. *For the remainder of this section, fix* $k \in \mathbb{N}$*, and* $B = \{p_1, p_2, \ldots, p_{k+2}\} \subset \mathbb{P}$*. Furthermore, define positive integers* g, h, t*, and* u *by:*

$$\begin{aligned} g &= \left\lfloor \left(\frac{2}{b_0}\right)^{k+1} \sum_{p \in B} p^{k+1} - 1 \right\rfloor, \\ h &= \left\lceil p_{k+2}\# \left(\frac{2}{b_0}\right)^{k+1} \sum_{p \in B} p^{k+1} \right\rceil, \\ t &= 2h(g+1) - 1, \\ u &= 2t - 2. \end{aligned}$$

Let $C = \mathbb{P} \setminus B$*, and let* $C_1 = \{p_{k+3}, p_{k+4}, \ldots, p_{k+2t}\}$ *be the least* u *elements of* C*,* $Q = \max\{C_1\} = p_{k+2t}$*,* $r_0 = \frac{1}{2}|1 - \zeta_Q|$*, and*

$$d = \prod_{p \in C_1} \Gamma\left(\frac{p+1}{2}\right)^2 \left(\frac{\pi^{2p-1}}{3^{p/2} p^{2p-1}}\right) \frac{\Gamma\left(\left(\frac{\sqrt{3}}{\pi} + \frac{1}{2}\right) p + \frac{1}{2}\right)}{\Gamma\left(\left(\frac{\sqrt{3}}{\pi} - \frac{1}{2}\right) p + \frac{1}{2}\right)},$$

Finally, let

$$h_1 = \left\lceil \frac{Q\#u}{p_{k+2}\#} \left(\frac{2}{d} \left(\frac{2Q}{b_0}\right)^u + 1 + \frac{u^2}{4}(u-1)^{u-1}\right) \right\rceil.$$

Lemma 8.10. *For all* $n \in \mathbb{N}_0$*,* $p_B^{(k)}(n) \geq -g$*, and for all* $n \geq h$*,* $p_B^{(k)}(n) \geq 1$*.*

Proof. From Lemma 8.1,

$$p_B^{(k)}(n) = g_B^{(k)}(n) + \psi_B^{(k)}(n).$$

Then

$$\begin{aligned} p_B^{(k)}(n) &= g_B^{(k)}(n) + \psi_B^{(k)}(n) \\ &= \alpha_1 + \alpha_2(n+1) + \psi_B^{(k)}(n) \\ &\geq \alpha_1 + \alpha_2 + \psi_B^{(k)}(n), \end{aligned}$$

since $\alpha_2 = 1/(p_{k+2}\#) > 0$. But $\alpha_1 + \alpha_2 = 1 - \psi_B^{(k)}(0)$, so

$$\begin{aligned} p_B^{(k)}(n) &\geq 1 + \sum_{p\in B}\sum_{j=1}^{p-1} \beta(\zeta_p^j)(\zeta_p^{jn} - 1) \\ &\geq 1 - 2\sum_{p\in B}\sum_{j=1}^{p-1} |\beta(\zeta_p^j)| \\ &\geq 1 - \left(\frac{2}{b_0}\right)^{k+1} \sum_{p\in B} p^{k+1}, \end{aligned}$$

by Theorem 8.5. The first assertion then follows from the fact that b_0 is transcendental. This establishes the first assertion. For the second, if $n \geq h$, then

$$p_B^{(k)}(n) \geq \alpha_1 + \alpha_2 + \left(\frac{2}{b_0}\right)^{k+1} \sum_{p\in B} p^{k+1} + \psi_B^{(k)}(n) \geq 1,$$

reasoning as before. □

Lemma 8.11. *If $n \geq h_1$ then*

$$\frac{(t-1)^2}{t+1}\frac{n^{2t-3}}{p_{C_1}^{(-1)}(n)} \leq \frac{1}{t+1}, \tag{8.8}$$

and hence

$$\frac{(t-1)^2}{t+1}\frac{n^{2t-3}}{p_{C}^{(-1)}(n)} \leq \frac{1}{t+1}. \tag{8.9}$$

Proof. The second inequality follows from the first via Proposition 8.7. By Lemma 8.1, we may write

$$p_{C_1}^{(-1)}(n) = g_{C_1}^{(-1)}(n) + \psi_{C_1}^{(-1)}(n),$$

where

$$g_{C_1}^{(-1)}(n) = \sum_{h=1}^{u+1} \alpha_h \binom{n+h-1}{h-1}.$$

By Lemma 8.6, we have

$$|\alpha_h| \leq \frac{1}{r_0^u d_{C_1}(r_0)},$$

so

$$g_{C_1}^{(-1)}(n) \geq \frac{p_{k+2}\#}{Q\#}\binom{n+u}{u} - \frac{1}{r_0^u d_{C_1}(r_0)}\sum_{h=1}^{u}\binom{n+h-1}{h-1}$$
$$\geq \frac{p_{k+2}\#}{Q\#}\binom{n+u}{u} - \frac{1}{r_0^u d_{C_1}(r_0)}\binom{n+u}{u-1}.$$

Let

$$\delta_p = \prod_{j=1}^{p-1}\left(|1-\zeta_p^j| - r_0\right),$$

so that

$$d_{C_1}(r_0) = \prod_{p\in C_1}\delta_p.$$

Since all the elements of C_1 are odd, making use of the inequalities (8.7), we have

$$\begin{aligned}
\delta_p &= \prod_{j=1}^{\frac{p-1}{2}}\left(|1-\zeta_p^j| - r_0\right)^2 \\
&= \prod_{j=1}^{\frac{p-1}{2}}|1-\zeta_p^j|^2\left(1-\frac{1}{2}\frac{|1-\zeta_Q|}{|1-\zeta_p^j|}\right)^2 \\
&\geq \prod_{j=1}^{\frac{p-1}{2}}\frac{4\pi^2 j^2}{p^2}\left(1-\frac{\pi^2 j^2}{3p^2}\right)\left(\frac{1}{2}\right)^2 \\
&= \prod_{j=1}^{\frac{p-1}{2}}\frac{\pi^2 j^2}{p^2}\left(1-\frac{\pi^2 j^2}{3p^2}\right) \\
&= \Gamma\left(\frac{p+1}{2}\right)^2\left(\frac{\pi^2}{\sqrt{3}p^2}\right)^{p-1}\prod_{j=1}^{\frac{p-1}{2}}\left(\frac{\sqrt{3}p}{\pi}-j\right)\left(\frac{\sqrt{3}p}{\pi}+j\right) \\
&= \Gamma\left(\frac{p+1}{2}\right)^2\left(\frac{\pi^{2p-1}}{3^{p/2}p^{2p-1}}\right)\frac{\Gamma\left(\left(\frac{\sqrt{3}}{\pi}+\frac{1}{2}\right)p+\frac{1}{2}\right)}{\Gamma\left(\left(\frac{\sqrt{3}}{\pi}-\frac{1}{2}\right)p+\frac{1}{2}\right)}.
\end{aligned} \tag{8.10}$$

Thus we can conclude that $d_{C_1}(r_0) \geq d$. Making use of Theorem 8.5, we have that our Theorem will be proved if for all $n \geq h_1$,

$$\frac{p_{k+2}\#}{Q\#}\binom{n+u}{u} - \frac{1}{r_0^u d}\binom{n+u}{u-1} - \sum_{p\in C_1}\left(\frac{p}{b_0}\right)^u \geq \frac{u^2}{4}n^{u-1}. \tag{8.11}$$

Together with the fact that via Lemma 8.3,

$$\frac{1}{r_0^u} \leq \left(\frac{2Q}{b_0}\right)^u,$$

the inequality (8.11) is implied by

$$n \geq \frac{Q\#u}{p_{k+2}\#\binom{n+u-1}{u-1}} \left[\frac{1}{d}\left(\frac{2Q}{b_0}\right)^u \binom{n+u}{u-1} + \sum_{p\in C_1}\left(\frac{p}{b_0}\right)^u + \frac{u^2}{4}n^{u-1}\right], \tag{8.12}$$

for $n \geq h_1$.

It is easily deduced that $h_1 \geq Qu^u$ and that

$$\frac{(u-1)^{u-1}}{(Qu^u)^{u-1}} < \frac{b_0^u}{uQ^u}.$$

Hence for $n \geq h_1$, we have

$$\frac{\binom{n+u}{u-1}}{\binom{n+u-1}{u-1}} = \frac{n+u}{n+1} \leq 2,$$

$$\begin{aligned}\frac{1}{\binom{n+u-1}{u-1}} \sum_{p\in C_1}\left(\frac{p}{b_0}\right)^u &\leq \left(\frac{u-1}{n+u-1}\right)^{u-1} \sum_{p\in C_1}\left(\frac{p}{b_0}\right)^u \\ &\leq \frac{(u-1)^{u-1}}{(Qu^u)^{u-1}} \sum_{p\in C_1}\left(\frac{p}{b_0}\right)^u < 1, \text{ and}\end{aligned}$$

$$\frac{n^{u-1}}{\binom{n+u-1}{u-1}} \leq \frac{n^{u-1}(u-1)^{u-1}}{(n+u-1)^{u-1}} \leq (u-1)^{u-1}.$$

The inequality (8.12) is implied by

$$n \geq \frac{Q\#u}{p_{k+2}\#}\left[\frac{2}{d}\left(\frac{2Q}{b_0}\right)^u + 1 + \frac{u^2}{4}(u-1)^{u-1}\right],$$

which is certainly true for all $n \geq h_1$. □

The following Corollary is an immediate consequence of Lemma 8.11 and Proposition 8.8:

Corollary 8.12. *For $n \geq h_1$,*

$$\frac{p_C(n)}{p_C^{(-1)}(n)} \leq \frac{2}{t+1}.$$

Theorem 8.13. *If* $n \geq h + h_1 - 1$*, then*

$$p_{\mathbb{P}}^{(k)}(n) \geq 0.$$

Proof. The following identity is established by considering the generating function for $p_{\mathbb{P}}^{(k)}(n)$, and using the fact that $\mathbb{P}$ is the disjoint union of B and C:

$$p_{\mathbb{P}}^{(k)}(n) = \sum_{m=0}^{n} p_B^{(k)}(n-m)p_C(m).$$

Thus, by Lemma 8.10 and Corollary 8.12, for $n \geq h + h_1 - 1$,

$$\begin{aligned}
p_{\mathbb{P}}^{(k)}(n) &= \sum_{m=0}^{n-h} p_B^{(k)}(n-m)p_C(m) + \sum_{m=n-h+1}^{n} p_B^{(k)}(n-m)p_C(m) \\
&\geq \sum_{m=0}^{n-h} p_C(m) - g \sum_{m=n-h+1}^{n} p_C(m) \\
&= p_C^{(-1)}(n) - (g+1) \sum_{m=n-h+1}^{n} p_C(m) \\
&\geq p_C^{(-1)}(n) - (g+1)p_C^{(-1)}(n) \sum_{m=n-h+1}^{n} \frac{p_C(m)}{p_C^{(-1)}(m)} \\
&\geq p_C^{(-1)}(n) - (g+1)p_C^{(-1)}(n)h\left(\frac{2}{t+1}\right) \\
&= p_C^{(-1)}(n) - p_C^{(-1)}(n) = 0.
\end{aligned}$$

□

Next we will use the information in this section to compute an asymptotic bound in k.

Theorem 8.14. *Let*

$$N(k) = \min\{N \in \mathbb{N} : p_{\mathbb{P}}^{(k)}(n) \geq 0 \text{ for all } n \geq N\}.$$

There is a function $\varepsilon = \varepsilon(k) \ll \frac{\log\log k}{\log k}$ *such that as* $k \to \infty$*,*

$$\log N(k) \ll k^{k(4+\log 16)(1+\varepsilon)}.$$

Proof. Clearly, $\log N(k) \ll \log h + \log h_1$. Now,

$$\begin{aligned}\log h &\ll \log(p_{k+2}\#) + \log\left(\sum_{p\in B} p^{k+1}\right)\\ &\ll \sum_{p\in B}\log p + \log\left(\frac{p_{k+2}^{k+2}}{(k+2)\log p_{k+2}}\right)\\ &\ll k\log p_{k+2} + k\log k \ll k\log k. \qquad (8.13)\end{aligned}$$

From the definition of h_1, it is easily seen that

$$\log h_1 \ll \log\left(\frac{Q\#}{p_{k+2}\#}\right) + \log\left(\frac{1}{d}\right) + u\log Q. \qquad (8.14)$$

But

$$\log\left(\frac{Q\#}{p_{k+2}\#}\right) \ll \sum_{j=1}^{k+2t}\log p_j \ll \sum_{j=1}^{k+2t}\log j \ll t\log t \text{ since } k < t, \text{ and} \qquad (8.15)$$

$$u\log Q \ll t\log t. \qquad (8.16)$$

Also,

$$\log\left(\frac{1}{d}\right) \ll \sum_{p\in C_1}|\log\delta_p|, \qquad (8.17)$$

so we need to look at bounding δ_p and $1/\delta_p$, for $p \in C_1$. By Stirling's Formula (2.5) we have

$$\begin{aligned}\delta_p &\le \prod_{j=1}^{\frac{p-1}{2}}|1-\zeta_p^j|^2\\ &\le \prod_{j=1}^{\frac{p-1}{2}}\frac{4\pi^2 j^2}{p^2}\\ &= \left(\frac{2\pi}{p}\right)^{p-1}\Gamma\left(\frac{p+1}{2}\right)^2\\ &\ll \left(\frac{2\pi}{p}\right)^{p-1}\frac{\pi(p-1)^p}{2^{p-1}e^{p-1}} \ll \left(\frac{\pi}{e}\right)^p p.\end{aligned}$$

On the other hand, if

$$c_1 = \left(\frac{\sqrt{3}}{\pi} - \frac{1}{2}\right)p - \frac{1}{2},$$

$$c_2 = \left(\frac{\sqrt{3}}{\pi} + \frac{1}{2}\right)p - \frac{1}{2} \text{ and}$$

$$c_3 = \frac{\left(\frac{\sqrt{3}}{\pi} - \frac{1}{2}\right)^{\left(\frac{\sqrt{3}}{\pi} - \frac{1}{2}\right)}}{\left(\frac{\sqrt{3}}{\pi} + \frac{1}{2}\right)^{\left(\frac{\sqrt{3}}{\pi} + \frac{1}{2}\right)}},$$

then by (8.10) and Stirling's formula,

$$\begin{aligned}
\frac{1}{\delta_p} &\le \left(\frac{3^{p/2}p^{2p-1}}{\pi^{2p-1}}\right)\frac{\Gamma(c_1+1)}{\Gamma\left(\frac{p+1}{2}\right)^2\Gamma(c_2+1)} \\
&\ll \left(\frac{3^{p/2}p^{2p-1}}{\pi^{2p-1}}\right)\frac{2^{p-1}e^{p-1}}{\pi(p-1)^p}\frac{c_1^{c_1}}{c_2{}^{c_2}}\sqrt{\frac{c_1}{c_2}}e^{c_2-c_1} \\
&\ll \left(\frac{2e^2\sqrt{3}c_3}{\pi^2}\right)^p p^{2p-1}p^{-2p} \ll \left(\frac{2e^2\sqrt{3}c_3}{\pi^2}\right)^p\left(\frac{1}{p}\right).
\end{aligned}$$

We may conclude that $|\log\delta_p| \ll p$, and so by (8.17)

$$\log\left(\frac{1}{d}\right) \ll \sum_{p\in C_1} p \ll \frac{Q^2}{\log Q} \ll t^2\log t. \tag{8.18}$$

In turn, by (8.13) through (8.16), we have that

$$\log N(k) \ll t^2\log t. \tag{8.19}$$

To complete the proof, we must bound $t^2\log t$. Since $t \ll gh$, and $h \ll p_{k+2}\#g$ we have

$$t^2\log t \ll (gh)^2\log h \ll (p_{k+2}\#)^2g^4k\log k \tag{8.20}$$

So we must bound g, but first, from Nathanson [33], p. 269, we invoke that fact that for every positive integer n,

$$\prod_{p\le n} p < 4^n.$$

In particular, we have that

$$\begin{aligned}(p_{k+2}\#)^2 &< 16^{p_{k+2}} \\ &= \exp\left((\log 16)p_{k+2}\right) \\ &= \exp\left[(\log 16)k\log k\left(1+O\left(\frac{\log\log k}{\log k}\right)\right)\right] \\ &= k^{(\log 16)k\left(1+O\left(\frac{\log\log k}{\log k}\right)\right)}. \end{aligned} \tag{8.21}$$

Next,

$$\begin{aligned} g &\ll \left(\frac{2}{b_0}\right)^{k+1}\sum_{p\in B} p^{k+1} \\ &\ll (k+2)\left(\frac{2p_{k+2}}{b_0}\right)^{k+1} \\ &\ll k\left[\left(\frac{2}{b_0}\right)(k+2)\log(k+2)\left(1+O\left(\frac{\log\log k}{\log k}\right)\right)\right]^{k+1} \\ &\ll k^{k+2}(\log k)^{k+1}\left(\frac{2}{b_0}\right)^{k+1}\left[1+O\left(\frac{\log\log k}{\log k}\right)\right]^{k+1} \end{aligned} \tag{8.22}$$

Putting together (8.19) through (8.22) we have

$$\begin{aligned}\log N(k) &\ll k^{\left(4+\log 16+O\left(\frac{\log\log k}{\log k}\right)\right)k+9}(\log k)^{4k+5}\left(\frac{2+O\left(\frac{\log\log k}{\log k}\right)}{b_0}\right)^{4k+4} \\ &\ll \exp\left[k\log k(4+\log 16)+O(k\log\log k)+9\log k+(4k+5)\log\log k+O(k)\right] \\ &\ll k^{k(4+\log 16)\left(1+O\left(\frac{\log\log k}{\log k}\right)\right)},\end{aligned}$$

and the proof is complete.

□

8.3 Asymptotic Formulae

The material found in this section is taken from the paper [48] by the author entitled "On partitions into powers of primes and their difference functions," accepted for publication in the Canadian Journal of Mathematics.

Our purpose in this section is to prove the following asymptotic formula:

$$\log p_{\mathbb{P}(r)}^{(k-1)}(n) = (r+1)\left[\Gamma\left(\frac{1}{r}+2\right)\zeta\left(\frac{1}{r}+1\right)\right]^{r/(r+1)} n^{1/(r+1)}(\log n)^{-r/(r+1)} \times\left(1+O_\epsilon\left(\sqrt{\frac{(\log\log n)^{1+\epsilon}}{\log n}}\right)\right), \text{ as } n\to\infty,$$

for fixed $k, r \geq 1$. This is a generalization of (8.1) which is the $k = 1$ case. The asymptotic (8.1) was first given by Hardy and Ramanujan [19], though without an explicit error term. They did not, however, provide a rigorous proof of this fact. Indeed, they assume that for a given r, $p_{\mathbb{P}(r)}^{(1)}(n) \geq 0$ for all n. This is readily seen to be false for r as low as 2, $n = 5$. Fortunately, we can use the monotonicity results of the previous section and Bateman and Erdős [4] to rectify the dilemma, and still follow their approach to the problem.

The theorem is of the Tauberian type: we shall first prove estimates for the generating functions, and then use them to yield information about the coefficients.

The subsequent analysis will be made easier with the following version of the prime number theorem:

$$\pi(x) = Li(x) + E(x),$$

where, recalling if necessary

$$Li(x) = \int_2^x \frac{1}{\log t}\,dt; \text{ and}$$

$$E(x) = O_\delta\left(\frac{x}{\log^\delta x}\right), \text{ for all } \delta \geq 2.$$

8.3.1 Asymptotic Formula for the Generating Function

In the following argument, s is assumed to be a small positive quantity approaching 0.

Define

$$\phi(s) = \sum_p e^{-sp^r}.$$

Lemma 8.15. *As $s \to 0^+$,*

$$\phi(s) = \int_2^\infty \frac{e^{-su^r}}{\log u}\,du + O_\delta\left(\frac{s^{-1/r}}{\log^\delta(1/s)}\right),$$

for any $\delta \geq 2$.

Proof. Using Riemann-Stieltjes integration, we have

$$\begin{aligned}\phi(s) &= \int_{2^-}^{\infty} e^{-su^r}\, d\pi(u) \\ &= \int_{2^-}^{\infty} e^{-su^r}\, d(Li(u) + E(u)) \\ &= \int_{2}^{\infty} \frac{e^{-su^r}}{\log u}\, du + \int_{2^-}^{\infty} e^{-su^r}\, dE(u). \qquad (8.23)\end{aligned}$$

Let

$$C = C(s) = \log^{-\delta}(1/s).$$

Note that as $s \to 0^+$, $s = o(C(s))$. Assume that $2^r s < C$. Integration by parts gives

$$\begin{aligned}\int_{2^-}^{\infty} e^{-su^r}\, dE(u) &= rs \int_{2}^{\infty} u^{r-1} e^{-su^r} E(u)\, du + O(1) \\ &\ll_\delta rs \int_{2}^{\infty} \frac{u^r e^{-su^r}}{\log^\delta u}\, du + O(1) \\ &\ll_\delta r^\delta \int_{2^r s}^{\infty} \left(\frac{t}{s}\right)^{1/r} \frac{e^{-t}}{\log^\delta (t/s)} dt + O(1), \text{ via the substitution } t = su^r \\ &= Cs^{-1/r} \int_{2^r s}^{\infty} \frac{t^{1/r} e^{-t}}{\left(\frac{\log t}{\log(1/s)} + 1\right)^\delta}\, dt + O(1) \\ &= Cs^{-1/r} \left[\int_{2^r s}^{C} \frac{t^{1/r} e^{-t}}{\left(\frac{\log t}{\log(1/s)} + 1\right)^\delta}\, dt + \int_{C}^{\infty} \frac{t^{1/r} e^{-t}}{\left(\frac{\log t}{\log(1/s)} + 1\right)^\delta}\, dt\right] + O(1). \qquad (8.24)\end{aligned}$$

Now,

$$\begin{aligned}
\int_{2^r s}^{C} \frac{t^{1/r}e^{-t}}{\left(\frac{\log t}{\log(1/s)}+1\right)^{\delta}}\,dt &\ll \int_{2^r s}^{C} \frac{t^{1/r}e^{-t}}{\left(\frac{\log(2^r s)}{\log(1/s)}+1\right)^{\delta}}\,dt \\
&= \int_{2^r s}^{C} \frac{t^{1/r}e^{-t}}{\left(\frac{r\log 2}{\log(1/s)}\right)^{\delta}}\,dt \\
&\ll_\delta \log^{\delta}(1/s)\int_{2^r s}^{C} t^{1/r}e^{-t}\,dt \\
&\ll_\delta C\log^{\delta}(1/s) \\
&= 1.
\end{aligned}$$

On the other hand for $C \le t < \infty$, we have that $\log t/\log(1/s)+1$ is minimized when $t = C$, so that

$$\begin{aligned}
\int_{C}^{\infty} \frac{t^{1/r}e^{-t}}{\left(\frac{\log t}{\log(1/s)}+1\right)^{\delta}}\,dt &\ll \int_{C}^{\infty} \frac{t^{1/r}e^{-t}}{\left(\frac{-\delta\log\log(1/s)}{\log(1/s)}+1\right)^{\delta}}\,dt \\
&\ll_\delta \int_{0}^{\infty} t^{1/r}e^{-t}\,dt \\
&\ll_\delta 1.
\end{aligned}$$

Hence by (8.24),

$$\int_{2^-}^{\infty} e^{-su^r}\,dE(u) \ll_\delta Cs^{-1/r},$$

which together with (8.23) completes the proof.

□

Lemma 8.16. *As* $s \to 0^+$,

$$\int_{2}^{\infty} \frac{e^{-su^r}}{\log u}\,du = r\Gamma\left(\frac{1}{r}+1\right)s^{-1/r}(\log(1/s))^{-1} + O\left(\frac{s^{-1/r}\log\log(1/s)}{\log^2(1/s)}\right).$$

Proof. Making the substitution $t = su^r$ into the integral gives

$$\begin{aligned}
\int_{2}^{\infty} \frac{e^{-su^r}}{\log u}\,du &= s^{-1/r}\int_{2^r s}^{\infty} \frac{t^{1/r-1}e^{-t}}{\log(1/s)}\,dt \\
&= s^{-1/r}(\log(1/s))^{-1}(I_1+I_2+I_3), \qquad (8.25)
\end{aligned}$$

where

$$I_1 = \int_{2^r s}^{1/\log^{2r}(1/s)} \frac{t^{1/r-1}e^{-t}}{1+\frac{\log t}{\log(1/s)}}\,dt,$$

$$I_2 = \int_{1/\log^{2r}(1/s)}^{\log^2(1/s)} \frac{t^{1/r-1}e^{-t}}{1+\frac{\log t}{\log(1/s)}}\,dt, \text{ and}$$

$$I_3 = \int_{\log^2(1/s)}^{\infty} \frac{t^{1/r-1}e^{-t}}{1+\frac{\log t}{\log(1/s)}}\,dt.$$

We will consider each of these integrals individually.

For $t \in [2^r s, 1/\log^{2r}(1/s)]$, $\log t/\log(1/s)$ is closest to -1 when $t = 2^r s$. Hence

$$\begin{aligned} I_1 &\ll \int_{2^r s}^{1/\log^{2r}(1/s)} \frac{t^{1/r-1}e^{-t}}{1+\frac{\log 2^r s}{\log(1/s)}}\,dt, \\ &\ll \log(1/s)\int_{2^r s}^{1/\log^{2r}(1/s)} t^{1/r-1}e^{-t}\,dt \\ &\ll \log(1/s)\int_{0}^{1/\log^{2r}(1/s)} t^{1/r-1}\,dt \\ &\ll \frac{1}{\log(1/s)}. \end{aligned} \tag{8.26}$$

Now we consider I_2. For $t \in [1/\log^{2r}(1/s), \log^2(1/s)]$, we have

$$\frac{1}{1+\frac{\log t}{\log(1/s)}} = 1 + O\left(\frac{\log\log(1/s)}{\log(1/s)}\right),$$

and so using integration by parts:

$$\begin{aligned} I_2 =& \int_{1/\log^{2r}(1/s)}^{\log^2(1/s)} t^{1/r-1}e^{-t}\,dt + O\left(\frac{\log\log(1/s)}{\log(1/s)}\right) \\ =& r\left[t^{1/r}e^{-t}\right]_{1/\log^{2r}(1/s)}^{\log^2(1/s)} + r\int_0^\infty t^{1/r}e^{-t}\,dt + O\left(\int_0^{1/\log^{2r}(1/s)} t^{1/r}e^{-t}\,dt\right) \\ &+ O\left(\int_{\log^2(1/s)}^\infty t^{1/r}e^{-t}\,dt\right) + O\left(\frac{\log\log(1/s)}{\log(1/s)}\right). \end{aligned}$$

But

$$\int_0^\infty t^{1/r}e^{-t}\,dt = \Gamma\left(\frac{1}{r}+1\right),$$

and all the remaining terms are $O(\log\log(1/s)/\log(1/s))$, so

$$I_2 = r\Gamma\left(\frac{1}{r}+1\right) + O\left(\frac{\log\log(1/s)}{\log(1/s)}\right). \tag{8.27}$$

Finally,

$$\begin{aligned} I_3 &\ll \frac{1}{\log(1/s)} \int_{\log^2(1/s)}^{\infty} t^{1/r-1} e^{-t}\, dt \\ &\ll \frac{1}{\log(1/s)}. \end{aligned} \tag{8.28}$$

The proof is completed by combining (8.25), (8.26), (8.27), and (8.28). □

The previous two lemmas yield:

Corollary 8.17. *As $s \to 0^+$,*

$$\phi(s) = r\Gamma\left(\frac{1}{r}+1\right) s^{-1/r} (\log(1/s))^{-1} + O\left(\frac{s^{-1/r}\log\log(1/s)}{\log^2(1/s)}\right).$$

Let $k \in \mathbb{N}$, and define

$$f(s) = \sum_{n=0}^{\infty} p^{(k)}_{\mathbb{P}(r)}(n) e^{-ns} = (1-e^{-s})^k \prod_p (1-e^{-sp^r})^{-1}.$$

That is, $f(s)$ is the generating function in e^{-s} of the kth difference function of $p_{\mathbb{P}(r)}(n)$. Taking logarithms we have

$$\begin{aligned} \log f(s) &= k\log(1-e^{-s}) - \sum_p \log(1-e^{-sp^r}) \\ &= k\log(1-e^{-s}) + \sum_p \sum_{j=1}^{\infty} \frac{e^{-jsp^r}}{j} \\ &= k\log(1-e^{-s}) + \sum_{j=1}^{\infty} \frac{1}{j} \sum_p e^{-jsp^r} \\ &= k\log(1-e^{-s}) + \sum_{j=1}^{\infty} \frac{\phi(js)}{j}. \end{aligned} \tag{8.29}$$

We wish to use our approximations for $\phi(s)$ to evaluate the sum $\sum_{j=1}^{\infty} \frac{\phi(js)}{j}$. To do this we break up the sum into two parts. Let

$$N = (1/s)/(\log(1/s)).$$

Then by Corollary 8.17:

$$\begin{aligned}
&\sum_{j\le N}\frac{\phi(js)}{j}\\
&=r\Gamma\left(\frac{1}{r}+1\right)s^{-1/r}\\
&\quad\times\left[\sum_{j\le N}\frac{1}{j^{1+1/r}\log(1/js)}+O\left(\sum_{j\le N}\frac{\log\log(1/js)}{j^{1+1/r}\log^2(1/js)}\right)\right]. \qquad (8.30)
\end{aligned}$$

Now,

$$\begin{aligned}
\sum_{j\le N}\frac{1}{j^{1+1/r}\log(1/js)}&=\frac{1}{\log(1/s)}\sum_{j\le N}\frac{1}{j^{1+1/r}\left(1-\frac{\log j}{\log(1/s)}\right)}\\
&=\frac{1}{\log(1/s)}\left[\sum_{j\le N}\frac{1}{j^{1+1/r}}+\frac{1}{\log(1/s)}\sum_{j\le N}\frac{\log j}{j^{1+1/r}\left(1-\frac{\log j}{\log(1/s)}\right)}\right]\\
&=(\log(1/s))^{-1}\left(\zeta\left(\frac{1}{r}+1\right)+O\left(\frac{1}{N^{1/r}}\right)\right)\\
&\quad+\frac{1}{\log^2(1/s)}O\left(\sum_{j\le N}\frac{1}{j^{1+1/2r}\left(1-\frac{\log j}{\log(1/s)}\right)}\right). \qquad (8.31)
\end{aligned}$$

We have,

$$\frac{(\log(1/s))^{-1}}{N^{1/r}}=O(s^{1/r}), \qquad (8.32)$$

and

$$\sum_{j\le N}\frac{1}{j^{1+1/2r}\left(1-\frac{\log j}{\log(1/s)}\right)}=\Sigma_1+\Sigma_2,$$

where

$$\begin{aligned}
\Sigma_1&=\sum_{j\le 1/\sqrt{s}}\frac{1}{j^{1+1/2r}\left(1-\frac{\log j}{\log(1/s)}\right)},\ \text{and}\\
\Sigma_2&=\sum_{1/\sqrt{s}<j\le N}\frac{1}{j^{1+1/2r}\left(1-\frac{\log j}{\log(1/s)}\right)}.
\end{aligned}$$

But

$$\Sigma_1\ll\sum_{j\le 1/\sqrt{s}}\frac{1}{j^{1+1/2r}}\ll 1,$$

and

$$\begin{aligned}\Sigma_2 &\ll \sum_{1/\sqrt{s}<j\le N} \frac{1}{j^{1+1/2r}\left(1-\frac{\log N}{\log(1/s)}\right)}\\ &\ll \sum_{1/\sqrt{s}<j\le N} \frac{\log(1/s)}{j^{1+1/2r}\log\log(1/s)}\\ &\ll \frac{s^{1/4r}\log(1/s)}{\log\log(1/s)},\end{aligned}$$

Hence

$$\sum_{j\le N} \frac{1}{j^{1+1/2r}\left(1-\frac{\log j}{\log(1/s)}\right)} \ll 1. \tag{8.33}$$

We use a similar technique to bound the error term in (8.30). Write

$$\sum_{j\le N} \frac{\log\log(1/js)}{j^{1+1/r}\log^2(1/js)} = \frac{1}{\log^2(1/s)}(\Sigma_1' + \Sigma_2'),$$

where

$$\begin{aligned}\Sigma_1' &= \sum_{j\le 1/\sqrt{s}} \frac{\log\log(1/js)}{j^{1+1/r}\left(1-\frac{\log j}{\log(1/s)}\right)^2}, \text{ and}\\ \Sigma_2' &= \sum_{1/\sqrt{s}<j\le N} \frac{\log\log(1/js)}{j^{1+1/r}\left(1-\frac{\log j}{\log(1/s)}\right)^2}.\end{aligned}$$

Then

$$\Sigma_1' \ll \sum_{j\le 1/\sqrt{s}} \frac{\log\log(1/s)}{j^{1+1/r}} \ll \log\log(1/s),$$

and

$$\begin{aligned}\Sigma_2' &\ll \sum_{1/\sqrt{s}<j\le N} \frac{\log\log(1/s)}{j^{1+1/r}\left(1-\frac{\log N}{\log(1/s)}\right)^2}\\ &\ll \sum_{1/\sqrt{s}<j\le N} \frac{\log^2(1/s)}{j^{1+1/r}\log\log(1/s)}\\ &\ll \frac{s^{1/2r}\log^2(1/s)}{\log\log(1/s)}.\end{aligned}$$

Hence,

$$\sum_{j\le N}\frac{\log\log(1/js)}{j^{1+1/r}\log^2(1/js)}\ll\frac{\log\log(1/s)}{\log^2(1/s)}. \tag{8.34}$$

Next we must consider the tail of the sum $\sum\phi(js)/j$:

$$\begin{aligned}
\sum_{j>N}\frac{\phi(js)}{j}&\ll\sum_{n=2}^{\infty}\sum_{j>N}\frac{e^{-jsn}}{j}\\
&\ll\frac{1}{N}\sum_{n=2}^{\infty}\sum_{j>N}e^{-jsn}\\
&\ll\frac{1}{N}\sum_{n=2}^{\infty}\frac{e^{-Nsn}}{1-e^{-sn}}\\
&=\frac{1}{N}\sum_{2\le n\le 1/s}\frac{e^{-Nsn}}{1-e^{-sn}}+\frac{1}{N}\sum_{n>1/s}\frac{e^{-Nsn}}{1-e^{-sn}}\\
&\ll\frac{1}{N}\sum_{2\le n\le 1/s}\frac{e^{-Nsn}}{sn}+\frac{1}{N}\sum_{n>1/s}e^{-Nsn}\\
&\ll\log(1/s)\sum_{n=2}^{\infty}e^{-Nsn}+\frac{1}{N}\frac{e^{-N}}{1-e^{-Ns}}\\
&\ll\log(1/s)\frac{e^{-2Ns}}{1-e^{-Ns}}+\frac{s\log(1/s)e^{-1/(s\log(1/s))}}{1-e^{-Ns}}\\
&\ll\log^2(1/s)e^{-2/\log(1/s)}+s\log^2(1/s)e^{-1/(s\log(1/s))}\\
&\ll\log^2(1/s).
\end{aligned} \tag{8.35}$$

Combining (8.29) through (8.35), and the fact that

$$\log(1-e^{-s})\ll\log(1/s)\ll\frac{s^{-1/r}\log\log(1/s)}{\log^2(1/s)},$$

as $s\to 0^+$, we have the following theorem:

Theorem 8.18. *As $s\to 0^+$,*

$$\log f(s)=r\Gamma\left(\frac{1}{r}+1\right)\zeta\left(\frac{1}{r}+1\right)s^{-1/r}(\log(1/s))^{-1}+O\left(\frac{s^{-1/r}\log\log(1/s)}{\log^2(1/s)}\right).$$

8.3.2 Bounding from Above

Now we are in a position to prove our main theorem, which we do in two parts, the first being the simplest. First let us introduce some new notation.

Let $k, r \geq 1$, $a_n = p^{(k)}_{\mathbb{P}(r)}(n)$, $A_n = \sum_{i=0}^{n} a_i = p^{(k-1)}_{\mathbb{P}(r)}(n)$, and denote the following constants:

$$A = r\Gamma\left(\frac{1}{r}+1\right)\zeta\left(\frac{1}{r}+1\right),$$
$$B = (r+1)\left[\Gamma\left(\frac{1}{r}+2\right)\zeta\left(\frac{1}{r}+1\right)\right]^{r/(r+1)}.$$

Furthermore, choose $C_1 > 0$ such that if

$$\delta(s) = C_1 \frac{\log\log(1/s)}{\log(1/s)},$$

then

$$|1-(1/A)s^{1/r}\log(1/s)\log f(s)| < C_1 \frac{\log\log(1/s)}{\log(1/s)}.$$

We begin by bounding $\log A_n$ from above.

Lemma 8.19. *There exists a function $\beta \ll \log\log n/\log n$ such that for all n sufficiently large,*

$$\log A_n < \frac{Bn^{1/(r+1)}}{(\log n)^{r/(r+1)}}(1+\beta).$$

Proof. We have that

$$(1-\delta(s))As^{-1/r}(\log(1/s))^{-1} < \log f(s) < (1+\delta(s))As^{-1/r}(\log(1/s))^{-1}. \tag{8.36}$$

Since $\lim_{n\to\infty} a_n = \infty$, there exists an $N \in \mathbb{N}$ depending on k and r such that $n > N$ implies that $a_n \geq 0$. We define a constant C_2 by

$$C_2 = \sum_{j=0}^{N} |a_j|.$$

Thus if $n > N$, then

$$\begin{aligned}
A_n e^{-ns} &= \sum_{j=0}^{N} a_j e^{-ns} + \sum_{j=N+1}^{n} a_j e^{-ns} \\
&< \sum_{j=0}^{N} a_j e^{-ns} + \sum_{j=N+1}^{n} a_j e^{-js} \\
&= \sum_{j=0}^{N} a_j (e^{-ns} - e^{-js}) + \sum_{j=0}^{n} a_j e^{-js} \\
&< f(s) + C_2,
\end{aligned}$$

and so

$$\begin{aligned}
\log A_n <& ns + (1+\delta(s))As^{-1/r}(\log(1/s))^{-1} \\
&+ \log(1 + C_2 e^{-(1-\delta(s))As^{-1/r}(\log(1/s))^{-1}}) \\
<& ns + (1+\delta(s))As^{-1/r}(\log(1/s))^{-1} \\
&+ O(e^{-(1-\delta(s))As^{-1/r}(\log(1/s))^{-1}}).
\end{aligned} \tag{8.37}$$

For a large value of n, we can, by continuity, choose a corresponding $s > 0$ such that

$$\frac{1-\delta(s)}{r} As^{-(r+1)/r}(\log(1/s))^{-1} < n < \frac{1+\delta(s)}{r} As^{-(r+1)/r}(\log(1/s))^{-1}. \tag{8.38}$$

For these values of s and n, we deduce from (8.38) that

$$\frac{1}{s} = \left[\left(\frac{rn\log(1/s)}{A}\right)\left(1 + O\left(\frac{\log\log(1/s)}{\log(1/s)}\right)\right)\right]^{r/(r+1)}, \tag{8.39}$$

and hence

$$\log(1/s) = \frac{r\log n}{r+1}\left(1 + O\left(\frac{\log\log(1/s)}{\log n}\right)\right). \tag{8.40}$$

Note that this implies that $\log(1/s) \ll \log n \ll \log(1/s)$ as $s \to 0$, or equivalently, as $n \to \infty$, so we may use $\log n$, and $\log(1/s)$ interchangeably in various error terms. This fact, together with equations (8.39) and (8.40) implies that

$$\begin{aligned} s &= \left[\left(\frac{A}{rn\log(1/s)}\right)\left(1+O\left(\frac{\log\log(1/s)}{\log(1/s)}\right)\right)\right]^{r/(r+1)} \\ &= \left[\left(\frac{A(r+1)}{r^2 n\log n}\right)\left(1+O\left(\frac{\log\log(1/s)}{\log(1/s)}\right)\right)\right]^{r/(r+1)} \\ &= \frac{B}{(r+1)(n\log n)^{r/(r+1)}}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right). \end{aligned} \tag{8.41}$$

From (8.40) and (8.41), we infer that

$$\begin{aligned} &ns+As^{-1/r}(\log(1/s))^{-1} \\ &\quad=\frac{Bn^{1/(r+1)}}{(r+1)(\log n)^{r/(r+1)}}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right) \\ &\qquad+\frac{A(r+1)^{(r+1)/r}n^{1/(r+1)}}{rB^{1/r}(\log n)^{r/(r+1)}}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right) \\ &\quad=\frac{Bn^{1/(r+1)}}{(\log n)^{r/(r+1)}}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right). \end{aligned}$$

Therefore, by (8.37),

$$\log A_n < \frac{Bn^{1/(r+1)}}{(\log n)^{r/(r+1)}}\left(1+O\left(\frac{\log\log n}{\log n}\right)\right). \tag{8.42}$$

This completes the proof of the lemma. □

8.3.3 Bounding from Below

Lemma 8.19 is one half of what we require. We use it to prove the other half.

Lemma 8.20. *Let $\epsilon > 0$ be given. Then there is a function*

$$\beta \ll_\epsilon \sqrt{\frac{(\log\log n)^{1+\epsilon}}{\log n}}$$

such that for all n sufficiently large,

$$\log A_n > \frac{Bn^{1/(r+1)}}{(\log n)^{r/(r+1)}}(1-\beta).$$

First let us introduce a convenient bit of notation. At times throughout the following argument, we are guaranteed the existence of certain positive functions which are $O(\log\log(1/s)/\log(1/s))$ in magnitude, as $s \to 0^+$. Rather than rename each such function, we may simply write η. Thus the precise η may vary, depending on the context, even within the same equation, but will always be used to denote such a positive function whose existence is guaranteed.

Proof. Let $\mathbf{A}(x) = A_n$, for $n \le x < n+1$. Hence by (8.42), there is a constant $C_3 > 0$, such that if

$$\eta_1(x) = C_3 \frac{\log\log x}{\log x},$$

then

$$\log \mathbf{A}(x) < \frac{Bx^{1/(r+1)}}{(\log x)^{r/(r+1)}} (1 + \eta_1(x)). \tag{8.43}$$

Now

$$\begin{aligned} f(s) &= \sum_{n=0}^{\infty} a_n e^{-ns} \\ &= \sum_{n=0}^{\infty} A_n (e^{-ns} - e^{-(n+1)s}) \\ &= s \sum_{n=0}^{\infty} A_n \int_n^{n+1} e^{-sx}\, dx \\ &= s \int_0^{\infty} \mathbf{A}(x) e^{-sx}\, dx. \end{aligned} \tag{8.44}$$

The inequalities (8.36) together with equation (8.44) imply that

$$\begin{aligned} \exp\Big((1-\delta(s))As^{-1/r}(\log(1/s))^{-1}\Big) &< s\int_0^{\infty} \mathbf{A}(x)e^{-sx}\, dx \\ &< \exp\Big((1+\delta(s))As^{-1/r}(\log(1/s))^{-1}\Big). \end{aligned} \tag{8.45}$$

Given a small value of $s > 0$, we can, by continuity, choose a corresponding $m > 0$ such that

$$\frac{1}{s} = \frac{r+1}{B}(m \log m)^{r/(r+1)}. \tag{8.46}$$

Now, denote

$$\begin{aligned} f(s) &= s\int_0^\infty \mathbf{A}(x)e^{-sx}\,dx \\ &= s\left(\int_0^{m/H} + \int_{m/H}^{(1-\zeta)m} + \int_{(1-\zeta)m}^{(1+\zeta)m} + \int_{(1+\zeta)m}^{Hm} + \int_{Hm}^{\infty}\right) \\ &= J_1 + J_2 + J_3 + J_4 + J_5, \end{aligned} \tag{8.47}$$

where

$$\zeta = \sqrt{\frac{(\log\log m)^{1+\epsilon}}{\log m}},$$

and $H > 1$ is a constant yet to be determined. We will see that the dominant term here is J_3, but first we shall prove that the terms J_1, J_2, J_4, and J_5 are negligible in comparison to the exponentials on either side of (8.45).

We first despatch J_1, and J_5. From Lemma 8.19, we have

$$\begin{aligned} J_1 &= s\int_0^{m/H} \mathbf{A}(x)e^{-sx}\,dx \\ &< \exp\left[(1+\eta_1(m/H))B(m/H)^{1/(r+1)}(\log{(m/H)})^{-r/(r+1)}\right]. \end{aligned} \tag{8.48}$$

Taking logarithms of (8.46), we see that

$$\frac{r+1}{r}(1-\eta) < \frac{\log m}{\log{(1/s)}}.$$

We can in light of this fact, select a positive function $\eta_3(s) \ll \log\log{(1/s)}/\log{(1/s)}$ such that for m sufficiently large relative to H (i.e. s sufficiently small),

$$\left(\frac{r+1}{r}\right)\frac{1+\eta_1(m/H)}{\left(1-\frac{\log H}{\log m}\right)^{r/(r+1)}} < \frac{(1+\eta_3(s))\log m}{\log{(1/s)}}.$$

This leads to the following string of inequalities:

$$\frac{(1+\eta_1(m/H))A(r+1)^{1+(r+1)/r}}{r^2\log m\left(1-\frac{\log H}{\log m}\right)^{r/(r+1)}} < \frac{(1+\eta_3(s))A(r+1)^{(r+1)/r}}{r\log(1/s)},$$

$$\frac{(1+\eta_1(m/H))B^{(r+1)/r}}{\log m\left(1-\frac{\log H}{\log m}\right)^{r/(r+1)}} < \frac{(1+\eta_3(s))A(r+1)^{(r+1)/r}}{r\log(1/s)},$$

$$\frac{(1+\eta_1(m/H))Bm^{1/(r+1)}}{(\log m)^{r/(r+1)}\left(1-\frac{\log H}{\log m}\right)^{r/(r+1)}} < \frac{(1+\eta_3(s))A(r+1)^{(r+1)/r}}{rB^{1/r}\log(1/s)} \times m^{1/(r+1)}(\log m)^{1/(r+1)}$$

$$\frac{(1+\eta_1(m/H))Bm^{1/(r+1)}}{(\log(m/H))^{r/(r+1)}H^{1/(r+1)}} < \frac{(1+\eta_3(s))A(r+1)}{rH^{1/(r+1)}s^{1/r}\log(1/s)}.$$

Comparing the final inequality with (8.48) yields:

$$J_1 < \exp\left[(1+\eta_3(s))AH^{-1/(r+1)}s^{-1/r}(\log(1/s))^{-1}\right],$$

for s sufficiently small. Choose H large enough such that for all s in the range in question,

$$\frac{1+\eta_3(s)}{H^{1/(r+1)}} \le \frac{1+\delta(s)}{2}.$$

Then

$$J_1 < \exp\left[((1+\delta(s))/2)As^{-1/r}(\log(1/s))^{-1}\right].$$

We now consider J_5. Note that $\max\{\eta_1(x) : x > 1\} = C_3/e$. We may choose H sufficiently large such that

$$\frac{1}{r+1} > \frac{2(1+C_3/e)}{H^{r/(r+1)}}.$$

Then

$$\begin{aligned} s &= \frac{B}{(r+1)(m\log m)^{r/(r+1)}} \\ &> \frac{2(1+C_3/e)B}{(Hm\log(Hm))^{r/(r+1)}} \\ &\ge \frac{2(1+\eta_1(x))B}{(x\log x)^{r/(r+1)}} \text{ for all } x \ge Hm, \end{aligned}$$

and so

$$\frac{(1+\eta_1(x))Bx^{1/(r+1)}}{(\log x)^{r/(r+1)}} < \frac{sx}{2},$$

for all $x \geq Hm$. Thus

$$\begin{aligned} J_5 &= s\int_{Hm}^{\infty} \mathbf{A}(x)e^{-sx}\,dx \\ &< s\int_{Hm}^{\infty} \exp\left[\frac{Bx^{1/(r+1)}(1+\eta_1(x))}{(\log x)^{r/(r+1)}} - sx\right] dx \\ &< s\int_0^{\infty} e^{-sx/2}\,dx \\ &= 2, \end{aligned}$$

where the first inequality follows from (8.43).

Now we take a look at the integrals J_2, and J_4, beginning with the latter. By (8.43),

$$J_4(s) = s\int_{(1+\zeta)m}^{Hm} \mathbf{A}(x)e^{-sx}\,dx < s\int_{(1+\zeta)m}^{Hm} e^{\psi(x)}\,dx,$$

where

$$\psi(x) = (1+\eta_1(x))Bx^{1/(r+1)}(\log x)^{-r/(r+1)} - sx. \tag{8.49}$$

If the maximum for $\psi(x)$ occurs at x_0, then, via a straightforward differentiation, it transpires that

$$\frac{1}{s} = \left(1 + O\left(\frac{\log\log x_0}{\log x_0}\right)\right)\frac{r+1}{B}x_0^{r/(r+1)}(\log x_0)^{r/(r+1)}. \tag{8.50}$$

Comparing this with (8.46), we conclude that $\log m \asymp \log x_0$, and that

$$x_0 = \left(1 + O\left(\frac{\log\log x_0}{\log x_0}\right)\right)m, \tag{8.51}$$

and therefore, for s sufficiently small,

$$(1-\zeta)m < x_0 < (1+\zeta)m. \tag{8.52}$$

Writing $x = x_0 + \xi$, Taylor's formula gives us

$$\psi(x) = \psi(x_0) + \frac{B}{2}\xi^2\frac{d^2}{dx^2}\left[(1+\eta_1(x))x^{1/(r+1)}(\log x)^{-r/(r+1)}\right]\Bigg|_{x=x_1},$$

where $x_0 < x_1 < x$, and hence

$$(1-\zeta)m < x_1 < Hm. \tag{8.53}$$

From this, it is easily seen that there exist positive constants C_4, C_5 such that

$$\begin{aligned}\frac{d^2}{dx_1^2}(1+\eta_1(x_1))\left[x_1^{1/(r+1)}(\log x_1)^{-r/(r+1)}\right] &< -C_4 x_1^{1/(r+1)-2}(\log x_1)^{-r/(r+1)} \\ &< -C_5 m^{1/(r+1)-2}(\log m)^{-r/(r+1)}.\end{aligned} \tag{8.54}$$

Equations (8.49), and (8.50) yield that

$$\psi(x_0) = As^{-1/r}(\log(1/s))^{-1}\left(1+O\left(\frac{\log\log(1/s)}{\log(1/s)}\right)\right). \tag{8.55}$$

Combining the information on $\psi(x)$, we see that there is a constant $C_6 > 0$ such that

$$\begin{aligned}J_4(s) <& s\exp\left[(1+\eta)As^{-1/r}(\log(1/s))^{-1}\right] \\ &\times \int_{(\zeta-\eta)m}^{\infty} \exp\left[-C_6\xi^2 m^{1/(r+1)-2}(\log m)^{-r/(r+1)}\right] d\xi.\end{aligned}$$

The integral on the right-hand side of this inequality is simplified by observing that it is of the form

$$\int_D^\infty e^{-Cx^2}\,dx,$$

for $C, D > 0$. Substituting $u^2 = Cx^2 - CD^2$, we have that

$$\int_D^\infty e^{-Cx^2}\,dx = \frac{1}{\sqrt{C}}\int_0^\infty \frac{ue^{-CD^2-u^2}}{\sqrt{u^2+CD^2}}\,du < \frac{e^{-CD^2}}{\sqrt{C}}\int_0^\infty e^{-u^2}\,du = \frac{e^{-CD^2}}{2}\sqrt{\frac{\pi}{C}}.$$

Hence with $D = (\zeta-\eta)m$, and $C = C_6 m^{1/(r+1)-2}(\log m)^{-r/(r+1)}$, there is a $C_7 > 0$ such that

$$J_4(s) \ll \frac{s\exp\left[(1+\eta)As^{-1/r}(\log(1/s))^{-1} - C_7\zeta^2 m^{1/(r+1)}(\log m)^{-r/(r+1)}\right]}{\sqrt{m^{1/(r+1)-2}(\log m)^{-r/(r+1)}}}. \tag{8.56}$$

Now, by definition of m,

$$\frac{s}{\sqrt{m^{1/(r+1)-2}(\log m)^{-r/(r+1)}}} = s\sqrt{m(m\log m)^{r/(r+1)}}$$
$$\ll \sqrt{sm} \ll \frac{1}{\sqrt{s^{1/r}\log(1/s)}}.$$

As we similarly have

$$s^{-1/r}(\log(1/s))^{-1} \asymp m^{1/(r+1)}(\log m)^{-r/(r+1)},$$

there is a constant $C_8 > 0$ such that

$$J_4(s) \ll \frac{\exp\left[(1+\eta-C_8\zeta^2)As^{-1/r}(\log(1/s))^{-1}\right]}{\sqrt{s^{1/r}\log(1/s)}}$$
$$\ll \exp\left[(1-C_8\zeta^2/2)As^{-1/r}(\log(1/s))^{-1}\right]. \tag{8.57}$$

Virtually the same analysis works for $J_2(s)$ giving a bound of a similar form. The results thus far have guaranteed us the existence of a constant $C_9 > 0$, such that

$$J_1, J_2, J_4, J_5 \ll \exp\left((1-C_9\zeta^2)As^{-1/r}(\log(1/s))^{-1}\right). \tag{8.58}$$

Hence by (8.45), we may select a new function $\delta_1(s)$ of the form $C\log\log(1/s)/\log(1/s)$ ($2\delta(s)$ works), such that for s sufficiently small,

$$\exp\left[(1-\delta_1(s))As^{-1/r}(\log(1/s))^{-1}\right] < s\int_{(1-\zeta)m}^{(1+\zeta)m} \mathbf{A}(x)e^{-sx}\,dx \tag{8.59}$$

Since $\mathbf{A}(x)$ is eventually increasing, we have

$$\exp\left[(1-\delta_1(s))As^{-1/r}(\log(1/s))^{-1}\right] < s\mathbf{A}((1+\zeta)m)\int_{(1-\zeta)m}^{(1+\zeta)m} e^{-sx}\,dx.$$

Evaluating the integral leads to

$$(e^{\zeta sm}-e^{-\zeta sm})\mathbf{A}((1+\zeta)m) > \exp\left[(1-\delta_1(s))As^{-1/r}(\log(1/s))^{-1}+ms\right]. \tag{8.60}$$

Substituting s in terms of m into the right-hand side of (8.60), we obtain an expression of the form

$$\exp\left[Bm^{1/(r+1)}(\log m)^{-r/(r+1)}\left(1+O\left(\frac{\log\log m}{\log m}\right)\right)\right].$$

Now, equation (8.46) yields

$$e^{\zeta sm} - e^{-\zeta sm} = e^{\frac{\zeta B}{r+1} m^{1/(r+1)} (\log m)^{-r/(r+1)}} \left(1 - e^{-\frac{2\zeta B}{r+1} m^{1/(r+1)} (\log m)^{-r/(r+1)}}\right),$$

and so by (8.60),

$$\begin{aligned}&\mathbf{A}((1+\zeta)m) \\ &\quad > \exp\left[B m^{1/(r+1)} (\log m)^{-r/(r+1)} \left(1 - \frac{\zeta}{r+1} + O\left(\frac{\log\log m}{\log m}\right)\right)\right].\end{aligned} \tag{8.61}$$

But $(1+\zeta)m$ is a continuous function of m, which is ultimately increasing. Thus for all n sufficiently large, we may choose a unique value of s, and hence of m such that $(1+\zeta)m = n$. Substituting

$$m = \frac{n}{1+\zeta}$$

into (8.61), and observing that $\log m \asymp \log n$, we have the lemma. □

Together, Lemmas 8.19 and 8.20 yield our main theorem:

Theorem 8.21. *For a fixed $k \geq 1$,*

$$\begin{aligned}\log p_{\mathbb{P}(r)}^{(k-1)}(n) =& (r+1)\left[\Gamma\left(\frac{1}{r}+2\right)\zeta\left(\frac{1}{r}+1\right)\right]^{r/(r+1)} n^{1/(r+1)} (\log n)^{-r/(r+1)} \\ &\times \left(1 + O_\epsilon\left(\sqrt{\frac{(\log\log n)^{1+\epsilon}}{\log n}}\right)\right), \quad \text{as } n \to \infty.\end{aligned}$$

Chapter 9

Sequences of Iterates

The simplest classes of prime symmetric functions we have looked at are those of the form e_k and s_k. The s_1-sequences and s_2-sequences of n are ultimately periodic (in fact of period 1) for any n. It seems highly probable, but is not yet proven whether the sequences must terminate for any other functions of the form $s_{k,\ell}$. For the most part in this section we will narrow our focus to s_1, which for ease of notation we denote by s and study its corresponding s-sequences.

9.1 Density of Sets with Given Terminal Value

Lal [26] investigates the function s, which he writes as J. We summarize his observations which are based on accumulated empirical data via computer searches. We will also expand on his data via our own computer searches.

Corresponding to each integer m, are two sets which we define as follows:

$$B_m = \{n \in \mathbb{N} : s(n) = m, n \neq m\}, \text{ and} \tag{9.1}$$

$$\mathfrak{B}_m = \{n \in \mathbb{N} : s^{(i)}(n) = m \text{ for some } i \geq 0\}. \tag{9.2}$$

That is, B_m is the set of all $n > m$ such that $s(n) = m$, whereas $\mathfrak{B}_m$ consists of all those integers n which include m in their s-sequence. The letter "B" is chosen in both instances to correspond to the word "branching." We can think of the set $\mathfrak{B}_m$ as a partially ordered set where $n \geq \ell$ if $s^{(i)}(n) = \ell$ for some $i \geq 0$.

With the exceptions of $m = 0, 1, 2, 3$, and 4, the sets $\mathfrak{B}_m$ are infinite. The sets B_m are always finite with cardinality $p_{\mathbb{P}}(m)$ if m is not prime, and $p_{\mathbb{P}}(m) - 1$ otherwise. For $n \geq 5$, the s-sequence of n always terminates in a prime p. Thus we have

$$\mathbb{N} = \{1, 2, 3, 4\} \cup \left(\dot{\bigcup_{p \geq 5}} \mathfrak{B}_{\mathfrak{p}} \right), \tag{9.3}$$

where all the unions are disjoint, and the second union is of infinite sets taken over all primes $p \geq 5$. We also have that

$$\mathfrak{B}_{\mathfrak{m}} = \{m\} \cup \left(\dot{\bigcup_{n \in \mathfrak{B}_{\mathfrak{m}}}} B_n \right), \tag{9.4}$$

where again, all unions are disjoint.

Lal observes that the set $\mathfrak{B}_p$ appears to define a positive asymptotic density for each prime $p \geq 5$. That is, if $\mathfrak{B}_p(x) = \#\{n \in \mathfrak{B}_p : n \leq x\}$ is the counting function for $\mathfrak{B}_p$, then $\mathfrak{B}_p(x)/x$ has a positive limit as $x \to \infty$. Let us consider some empirical data that might motivate such a conjecture. The tables below are extensions of those provided by Lal. In the first two, the entries correspond to values of $\mathfrak{B}_p(x)$ for different primes p, with x increasing in increments of 5000. In the second two, we evaluate the ratio $\mathfrak{B}_p(x)/x$. The algorithm we use to compute $\mathfrak{B}_p(x)$ is found in the appendix on Maple algorithms.

Table 4. Selected Values of $\mathfrak{B}_p(x)$

$x\backslash p$	5	7	11	13	17	19	23	29	31	37
5000	1426	810	327	374	152	184	121	62	86	56
10000	2830	1605	649	714	306	377	213	134	172	104
15000	4188	2397	941	1049	497	573	306	183	241	156
20000	5534	3202	1241	1400	677	780	426	235	316	200
25000	6846	3996	1556	1737	856	976	538	285	392	235
30000	8197	4812	1860	2066	1027	1175	664	341	472	289
35000	9547	5597	2156	2421	1203	1358	788	404	552	332
40000	10879	6380	2477	2758	1360	1566	900	462	622	374
45000	12217	7176	2785	3086	1506	1755	1012	520	703	423
50000	13585	7964	3081	3440	1667	1948	1118	585	776	463
55000	14879	8742	3367	3767	1809	2157	1251	642	850	522
60000	16226	9509	3662	4111	1964	2351	1356	698	927	572
65000	17599	10319	3938	4453	2088	2542	1473	765	998	620
70000	18967	11049	4237	4812	2243	2734	1597	827	1068	672
75000	20288	11858	4512	5138	2383	2927	1707	882	1128	706
80000	21643	12677	4820	5458	2535	3123	1832	939	1198	764
85000	22976	13450	5126	5819	2680	3297	1937	998	1279	801
90000	24340	14243	5417	6128	2826	3472	2059	1063	1353	849
95000	25655	15017	5736	6483	2993	3669	2175	1125	1413	898
100000	27023	15753	6032	6821	3146	3837	2289	1194	1497	937

$x\backslash p$	41	43	47	53	59	61	67	71	73	79
5000	51	70	40	37	28	45	22	27	33	22
10000	88	113	88	76	50	85	58	52	74	43
15000	139	170	139	97	77	124	84	76	110	65
20000	180	221	189	128	101	158	102	103	139	82
25000	221	289	223	156	121	193	124	130	171	110
30000	260	346	263	187	145	231	143	153	191	126
35000	287	398	303	221	171	266	168	171	212	146
40000	331	453	340	253	199	303	188	195	243	165
45000	368	520	375	275	225	341	213	220	272	183
50000	396	588	409	312	246	367	235	239	305	206
55000	441	643	449	345	275	399	258	259	341	226
60000	469	705	492	378	294	437	289	271	374	244
65000	503	757	523	403	318	469	304	293	416	268
70000	537	804	565	434	335	503	325	323	455	284
75000	577	869	597	474	362	536	350	351	493	306
80000	609	921	640	498	390	567	365	371	523	323
85000	647	983	678	527	406	613	385	398	557	341
90000	682	1041	704	567	435	638	397	413	592	366
95000	708	1092	752	593	454	672	417	430	625	384
100000	747	1161	792	626	477	719	444	444	658	401

Table 5. Selected Values of $\mathfrak{B}_p(x)/x$

$x\backslash p$	5	7	11	13	17	19	23	29	31	37
5000	.2852	.1620	.0654	.0748	.0304	.0368	.0242	.0124	.0172	.0112
10000	.2830	.1605	.0649	.0714	.0306	.0377	.0213	.0134	.0172	.0104
15000	.2792	.1598	.0627	.0699	.0331	.0382	.0204	.0122	.0161	.0104
20000	.2767	.1601	.0621	.0700	.0339	.0390	.0213	.0118	.0158	.0100
25000	.2738	.1598	.0622	.0695	.0342	.0390	.0215	.0114	.0157	.0094
30000	.2732	.1604	.0620	.0689	.0342	.0392	.0221	.0114	.0157	.0096
35000	.2728	.1599	.0616	.0692	.0344	.0388	.0225	.0115	.0158	.0095
40000	.2720	.1595	.0619	.0690	.0340	.0392	.0225	.0116	.0156	.0094
45000	.2715	.1595	.0619	.0686	.0335	.0390	.0225	.0116	.0156	.0094
50000	.2717	.1593	.0616	.0688	.0333	.0390	.0224	.0117	.0155	.0093
55000	.2705	.1590	.0612	.0685	.0329	.0392	.0227	.0117	.0155	.0095
60000	.2704	.1585	.0610	.0685	.0327	.0392	.0226	.0116	.0155	.0095
65000	.2708	.1588	.0606	.0685	.0321	.0391	.0227	.0118	.0154	.0095
70000	.2710	.1578	.0605	.0687	.0320	.0391	.0228	.0118	.0153	.0096
75000	.2705	.1581	.0602	.0685	.0318	.0390	.0228	.0118	.0150	.0094
80000	.2705	.1585	.0603	.0682	.0317	.0390	.0229	.0117	.0150	.0096
85000	.2703	.1582	.0603	.0685	.0315	.0388	.0228	.0117	.0150	.0094
90000	.2704	.1583	.0602	.0681	.0314	.0386	.0229	.0118	.0150	.0094
95000	.2701	.1581	.0604	.0682	.0315	.0386	.0229	.0118	.0149	.0095
100000	.2702	.1575	.0603	.0682	.0315	.0384	.0229	.0119	.0150	.0094

$x\backslash p$	41	43	47	53	59	61	67	71	73	79
5000	.0102	.0140	.0080	.0074	.0056	.0090	.0044	.0054	.0066	.0044
10000	.0088	.0113	.0088	.0076	.0050	.0085	.0058	.0052	.0074	.0043
15000	.0093	.0113	.0093	.0065	.0051	.0083	.0056	.0051	.0073	.0043
20000	.0090	.0111	.0095	.0064	.0051	.0079	.0051	.0052	.0070	.0041
25000	.0088	.0116	.0089	.0062	.0048	.0077	.0050	.0052	.0068	.0044
30000	.0087	.0115	.0088	.0062	.0048	.0077	.0048	.0051	.0064	.0042
35000	.0082	.0114	.0087	.0063	.0049	.0076	.0048	.0049	.0061	.0042
40000	.0083	.0113	.0085	.0063	.0050	.0076	.0047	.0049	.0061	.0041
45000	.0082	.0116	.0083	.0061	.0050	.0076	.0047	.0049	.0060	.0041
50000	.0079	.0118	.0082	.0062	.0049	.0073	.0047	.0048	.0061	.0041
55000	.0080	.0117	.0082	.0063	.0050	.0073	.0047	.0047	.0062	.0041
60000	.0078	.0118	.0082	.0063	.0049	.0073	.0048	.0045	.0062	.0041
65000	.0077	.0116	.0080	.0062	.0049	.0072	.0047	.0045	.0064	.0041
70000	.0077	.0115	.0081	.0062	.0049	.0072	.0046	.0047	.0065	.0041
75000	.0077	.0116	.0080	.0063	.0048	.0071	.0046	.0047	.0066	.0041
80000	.0076	.0115	.0080	.0062	.0049	.0071	.0046	.0046	.0065	.0040
85000	.0076	.0116	.0080	.0062	.0048	.0071	.0045	.0047	.0066	.0040
90000	.0076	.0116	.0078	.0063	.0048	.0071	.0044	.0046	.0066	.0041
95000	.0075	.0115	.0079	.0062	.0048	.0071	.0044	.0045	.0066	.0040
100000	.0075	.0116	.0079	.0063	.0048	.0072	.0044	.0044	.0066	.0040

Observe that for each value of p, the ratio $\mathfrak{B}_p(x)/x$ appears to converge to a fixed positive value, or else generally decreases at an increasingly slow rate. Presently, the positivity, or even existence of the limit remains unresolved. To investigate this conjecture, it would, in the least, be necessary to determine if generally, the composite functions $s^{(i)}(n)$ behave "nicely" for i fixed, or at least suitably small relative to n. That is, given that $s(n)$ is often approximately $\frac{\pi^2 n}{12\log n}$, is the average value of $s^{(i)}(n)$ asymptotic to the i-th composition of $\frac{\pi^2 n}{12\log n}$ with itself?

Empirical evidence leads to corresponding conjectures for each s_k:

Conjecture 9.1. *Let $k, n \in \mathbb{N}$, and suppose that $s_k(n) = n$. If the set*

$$\{m \in \mathbb{N} : s_k^{(i)}(m) = n, \text{ for some } i \geq 0\}$$

is infinite, then it has a positive asymptotic density.

Example 9.2. *The number 39 satisfies* $s_2(39) = 39$*, and the set of numbers which terminate in 39 under iteration by* s_2 *is infinite. If we denote this set by* $\mathfrak{B}$*, then we have the following table of values for* $\mathfrak{B}(x)$ *and* $\mathfrak{B}(x)/x$ *with* x *ranging in increments of 5000 up to 100000:*

Table 6. Iterations of s_2 Terminating in 39

x	$\mathfrak{B}(x)$	$\mathfrak{B}(x)/x$
5000	40	.00800
10000	54	.00540
15000	68	.00453
20000	85	.00425
25000	96	.00384
30000	107	.00357
35000	117	.00334
40000	132	.00330
45000	144	.00320
50000	153	.00306
55000	165	.00300
60000	188	.00313
65000	199	.00306
70000	210	.00300
75000	222	.00296
80000	232	.00290
85000	241	.00284
90000	253	.00281
95000	263	.00277
100000	276	.00276

It should be noted however, that it is still well within the range of possibility that the sets under discussion do not have an asymptotic density. Any illuminating information on their distribution would be of value. The corresponding sets for the functions $e_k(n)$ would be much less easily understood for $k \geq 2$, because e_k-sequences tend to fluctuate unpredictably. This makes sense since the average value of $e_k(n)$

$$\frac{1}{x}\sum_{n\leq x} e_k(n) \sim \frac{\zeta(k+1)x^k}{(k+1)\log x},$$

which grows significantly faster than x for $k \geq 2$.

9.2 Distribution of Products in Partitions into Primes

The problem of the asymptotic density of $\mathfrak{B}_n$ remains tantalizingly open. The value of $\mathfrak{B}_n(x)$ can be determined from the values of $B_m(x)$ using the formula

$$\mathfrak{B}_n(x) = 1 + \sum_{m \in \mathfrak{B}_n} B_m(x). \tag{9.5}$$

Hence, a good understanding of the distribution of the elements of B_m would lead to information regarding the density of $\mathfrak{B}_m$. That is, we require an understanding of the distribution of the products of parts in partitions into primes. We studied this problem in Section 5.2, when we looked at $b_k(x, y)$, however, the value of y was dependent on x. We wish to understand $b_1(x, y)$ for a fixed y. We drop the subscript and write $b(x, y)$ for $b_1(x, y)$. We have the relation

$$B_m(x) = \begin{cases} b(x, m) - b(x, m-1), & \text{if } m \text{ is not prime or } 4, \\ b(x, m) - b(x, m-1) - 1, & \text{if } m \text{ is prime or } 4. \end{cases}$$

Hansraj Gupta [18] shows that the largest element of $s^{(-1)}[\{m\}]$ for $m \geq 2$ is given by $t \cdot 3^{\lfloor m/3 \rfloor}$, where $t = 1, 4/3$, or 2 according as $m \equiv 0, 1$ or 2 (mod 3) respectively. He goes on to show that corresponding to any prime $p \geq 5$, the largest integer $h_r = h_r(p)$ such that $s^{(r)}(h_r) = p$, but $s^{(r-1)}(h_r) \neq p$ is given by

$$h_r = 3^{h_{r-1}/3}; \text{ with } h_1 = 4 \cdot 3^{(p-4)/3} \text{ or } 2 \cdot 3^{(p-2)/3},$$

according as $p \equiv 1$ or 2 (mod 3). The similar question as to the least element of B_m could at best be answered approximately, and depends on the Goldbach conjecture.

Thus the set B_m has a very large range relative to m. Empirical evidence would suggest however that x needn't be as high as $O(3^{m/3})$ for $B_m(x)$ to capture most of B_m. A possible line of attack on the function $B_m(x)$ could involve exponential sums and the circle method. With the customary notation $e(\alpha) = e^{2\pi i \alpha}$, for $0 \leq \alpha \leq 1$, let

$$f(\alpha) = \sum_{n \leq x} e(\alpha s(n)).$$

Then

$$\#\{n \leq x : s(n) = m\} = \int_0^1 f(\alpha) e(-\alpha m)\, d\alpha.$$

Ideally, we would proceed by estimating $f(\alpha)$ when $\alpha = a/q$, a rational number in lowest terms. Clearly $f(0) = f(1) = \lfloor x \rfloor$. In Chapter 7 we bounded the absolute value of sums of the form $f(a/q)$ using Perron's formula. In fact, we did so to far more general sums. We were able to do this because the additive properties of the divisor functions in question enabled us to express corresponding Dirichlet series in terms of L-functions. The problem here however arises from the fact that we are trying not simply to bound an integral from above, but approximate its actual value. This poses difficulties even in the case when $\alpha = 1/2$. By Perron's formula, we have the heuristic

$$f(1/2) \approx \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{1+2^{-s}}{1-2^{-s}} \frac{\zeta(2s)}{\zeta(s)} \frac{x^s}{s}\, ds,$$

for suitably chosen $T > 0$ and $c > 1$. Of course the function in the integrand is analytic in a region that extends beyond the line $\sigma = 1$, but even assuming the Riemann Hypothesis, it is not clear what the value of the integral should be. In case $\alpha = 1/3$, after some work, we obtain the heuristic

$$f(1/3) \approx$$

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{1}{(1-3^{-s})^{3/2}} \exp\left[\sum_p \sum_{j=2}^{\infty} (\zeta^{pj} - \zeta^{p^j}) p^{-js}/j\right] L(s,\chi)^{i\sqrt{3}/2} \zeta(s)^{-1/2} \frac{x^s}{s}\, ds,$$

where $\zeta = e(1/3)$, and χ is the non-principal character modulo 3. Assuming the Generalized Riemann Hypothesis, the function in the integrand can be analytically continued to the region in the plane $\{s : \sigma > 1/2\} \backslash \{s : \sigma \leq 1, t = 0\}$. Trying to integrate around the branch cut proves problematic. Furthermore, in general, it is not clear what size the minor arcs should be in application of the circle method. This will in turn affect the values of c and T chosen for a given $\alpha = a/q$. The problems and complexity increase with the value of the denominator q. To conclude, if the circle method is to be used to estimate $B_m(x)$, then it may not be practicable to do so in conjunction with Perron's formula.

9.3 The Number of Iterations of s until termination

Definition 9.3. *Suppose $f : \mathbb{N}_0 \to \mathbb{N}_0$ is an arithmetic function. Then we define the function $t_f : \mathbb{N}_0 \to \mathbb{N}_\infty$ by letting $t_f(n)$ be the number of distinct elements in the f-sequence of n. The value of t_f is infinity if the f-sequence never becomes periodic.*

If $f = s$, we let $t = t_f$.

For $k > 1$, the e_k-sequences tend to be rather erratic. This is due to the fact that when n has few divisors relative to k, then $e_k(n)$ tends to be very large relative to n. This lack of stability makes it difficult to predict the behaviour of the sequence of iterates, and hence of the function t_f.

None the less, the case $k = 1$ proves to provide a wealth of fascinating material for study, and at least empirically seems to behave with a high degree of regularity. Clearly $t(n) > 0$ for all n, and $t(n) = 1$ if and only if n is 0, a prime, or 4. A trivial upper-bound for $t(n)$ is found as follows.

Proposition 9.4. *For all $n \geq 5$,*

$$t(n) \leq \log(n)/\log(2) + 1.$$

Proof. It is straightforward to see that if n is not a prime, then $s(n) \leq n/2 + 2$. Thus

$$\begin{aligned} s^{(k)}(n) &\leq \frac{n}{2^k} + \frac{2}{2^{k-1}} + \frac{2}{2^{k-2}} + \cdots + 2, \\ &< \frac{n}{2^k} + 4. \end{aligned}$$

providing $s^{(i)}(n)$ is never prime for $0 \leq i \leq k-1$. This sequence must terminate at a prime $p \geq 5$. Thus the sequence can be iterated at most $k = \log_2(n)$ times before terminating at a prime. That is, $t(n) \leq \log_2(n) + 1$ as claimed. □

Observe that the bound $s(n) \leq n/2 + 2$ is only attained when $n = 2p$, for p prime, apart from the case when n itself is prime. Apart from these cases, we have that $s(n) \leq n/3 + 3$.

If $n = 2p$, then $s(n) = p + 2$, and so for $p > 2$, we can not have $s(n) = 2p_1$ for some other prime p_1. This observation will allow us to effect a slight improvement in the constant $1/\log 2$.

Lemma 9.5. *Let $k \geq 0$, $n \geq 2$. If $s^{(i)}(n)$ is not prime for any i, $0 \leq i < k$, then*

$$s^{(k)}(n) < \frac{n}{2^{\lceil k/2 \rceil} \cdot 3^{\lfloor k/2 \rfloor}} + \frac{22}{5}.$$

Proof. This can be verified numerically for $n = 2, 3, 4, 5$. For $n > 5$, we have seen that

$$s(n) \leq \frac{n}{2} + 2,$$

when n is not prime, so by the preceding remarks,

$$\begin{aligned} s^{(2)}(n) &\leq \frac{n}{2\cdot 3} + \frac{2}{3} + 3 \\ s^{(3)}(n) &\leq \frac{n}{2^2\cdot 3} + \frac{2}{2\cdot 3} + \frac{3}{2} + 2 \\ s^{(4)}(n) &\leq \frac{n}{2^2\cdot 3^2} + \frac{2}{2\cdot 3^2} + \frac{3}{2\cdot 3} + \frac{2}{3} + 3, \end{aligned}$$

and in general,

$$\begin{aligned} s^{(2k-1)}(n) &\leq \frac{n}{2^k\cdot 3^{k-1}} + 2\sum_{i=0}^{k-1}\frac{1}{2^i\cdot 3^i} + \frac{3}{2}\sum_{i=0}^{k-2}\frac{1}{2^i\cdot 3^i} < \frac{n}{2^k\cdot 3^{k-1}} + \frac{7}{2}\sum_{i=0}^{\infty}\frac{1}{6^i} \\ s^{(2k)}(n) &\leq \frac{n}{2^k\cdot 3^k} + \frac{2}{3}\sum_{i=0}^{k-1}\frac{1}{2^i\cdot 3^i} + 3\sum_{i=0}^{k-1}\frac{1}{2^i\cdot 3^i} < \frac{n}{2^k\cdot 3^k} + \frac{11}{3}\sum_{i=0}^{\infty}\frac{1}{6^i}. \end{aligned}$$

Combining these we have that

$$s^{(k)}(n) < \frac{n}{2^{\lceil k/2\rceil}\cdot 3^{\lfloor k/2\rfloor}} + \frac{22}{5}.$$

□

Proposition 9.6. *If $n \geq 2$, then*

$$t(n) < \frac{2\log n}{\log 6} + \frac{2\log 5}{\log 6}.$$

Proof. This can be verified numerically for $n = 2, 3, 4, 5$. For $n > 5$, write $t(n) = k + 1$. This means $s^{(k)}(n) = p \in \mathbb{P}$. By the lemma, we have

$$5 \leq p = s^{(k)}(n) < \frac{n}{2^{\lceil k/2\rceil}\cdot 3^{\lfloor k/2\rfloor}} + \frac{22}{5},$$

and so

$$2^{\lceil k/2\rceil}\cdot 3^{\lfloor k/2\rfloor} < \frac{5}{3}n.$$

But

$$2^{\lceil k/2\rceil}\cdot 3^{\lfloor k/2\rfloor} \geq 6^{k/2}\sqrt{\frac{2}{3}},$$

so

$$6^{k/2} < \frac{5}{\sqrt{6}}n.$$

This implies that

$$k+1 < \frac{2\log n}{\log 6} + \left(\frac{2\log(5/\sqrt{6})}{\log 6} + 1\right) = \frac{2\log n}{\log 6} + \frac{2\log 5}{\log 6},$$

which is the statement of the proposition. □

The improvement of the multiplicative constant is somewhat slight, from $1/\log 2 \approx 1.44270$, to $2/\log 6 \approx 1.11622$.

9.3.1 The Average order of $t(n)$

The following formulae are useful:

$$\mathfrak{B}_n(x) = 1 + \sum_{m \in \mathfrak{B}_n} B_m(x) = 1 + \sum_{m \in B_n} \mathfrak{B}_m(x). \tag{9.6}$$

Proposition 9.7. *The average order of $t(n)$ has the following expressions:*

$$\sum_{n \le x} t(n) = 2 + \sum_{2 \le n \le x} \mathfrak{B}_n(x) = 1 + \lfloor x \rfloor + \sum_{2 \le n \le x} \sum_{m \in \mathfrak{B}_n} B_m(x)$$

Proof. The second equation follows from the first, and the first equation in (9.6). To prove the first, first note that $t(1) = 2$. The sum $\sum t(n)$ counts each n, $2 \le n \le x$ once for each $m \le x$ satisfying $s^{(i)}(m) = n$ for some $i \ge 0$. Simply combine these two facts. □

Though we have shown that $t(n) \ll \log x$, it is no surprise that this does not seem to be optimal on average. Empirical results appear to indicate that

$$\sum_{n \le x} t(n) \asymp x \log\log x.$$

Indeed, consider the following table of values:

Table 7. Average Order of $t(n)$

x	$\sum_{n\leq x} t(n)$	$\sum_{n\leq x} t(n)\frac{1}{x\log x}$	$\sum_{n\leq x} t(n)\frac{1}{x\log\log x}$
5000	19571	.4595645436	1.827283521
10000	40471	.4394082995	1.822749691
15000	61620	.4272133008	1.814962097
20000	83104	.4195693666	1.812213183
25000	104673	.4134563567	1.808472967
30000	126593	.4093302983	1.808719966
35000	148545	.4056294504	1.807668137
40000	170568	.4024107756	1.806453766
45000	192657	.3995799640	1.805230307
50000	215000	.3974203434	1.805678899
55000	237041	.3948512538	1.803168591
60000	259322	.3928371307	1.802284432

There is a heuristic argument that lends credence to the possibility that

$$\sum_{n\leq x} t(n) \sim Cx\log\log x$$

for some constant $C > 0$. In Section 5.2, we saw that $s(n) \leq x^{\alpha}$ a positive proportion of the time on the interval $[1, x]$. Iterating the function x^{α} j times gives us x^{α^j}. Setting $x^{\alpha^j} \leq c$ for some constant c, we see that the minimum required value of j is roughly $\log\log x$. At present, unfortunately, we can do no better than argue heuristically and provide supporting empirical evidence.

9.4 Iterating $s + b$

The main goal of this section is to study the sequences formed by iterating the function $s+b$ for positive integral values of b. We begin with a definition.

Definition 9.8. *Let $p_1, \ldots, p_r$ be primes such that $p_1 \leq \ldots \leq p_r$, and let $n = p_1 \cdots p_r$. We say that n is first s_k-defective of order r if n is s_k-defective and if whenever primes $q_1, \ldots, q_r$ satisfying $q_1 \leq \ldots \leq q_r$, are such that $q_1 \cdots q_r$ is s_k-defective and $q_i \leq p_i$ for each i, then $q_i = p_i$ for each i.*

Example 9.9. $42 = 2 \cdot 3 \cdot 7$ *is first s_2-defective of order 3.*

For a given k, and fixed r, there are only finitely many first s_k-defective numbers of order r. Let $r(k)$ be defined as in Lemma 2.14. If $r < k$, or if $r \geq r(k)$ then 2^r is the only first s_k-defective number of order r. If $r = k$, there are no such numbers, since no such number is s_k-defective.

If $n = p_1 \cdots p_r$ is s_k-defective, then there exists an $n' = p'_1 \cdots p'_r$ which is first s_k-defective of order r, and such that $p'_i \leq p_i$ for each i.

We make use of the fact that there are only finitely many such numbers for each r in the following lemma.

Lemma 9.10. *Let $k \in \mathbb{N}$, and let $M > 0$. Then there exists an $N \in \mathbb{N}$ such that if $n \geq N$ and n is s_k-defective, then $n - s_k(n) > M$.*

Proof. Let $n = p_1 \cdots p_r$. The inequality

$$n > s_k(n) + M$$

holds if and only if

$$1 > \sum_{1 \leq i_1 < \ldots < i_{r-k} \leq r} \frac{1}{p_{i_1} \cdots p_{i_{r-k}}} + \frac{M}{n}.$$

This is implied by

$$1 > \binom{r}{k} \frac{1}{2^{r-k}} + \frac{M}{2^r}. \tag{9.7}$$

There is an R such that $r \geq R$ implies that the inequality (9.7) holds. On the other hand, if $r < k$, then we need only ensure $N > M$. Thus we may assume $k < r < R$, for which there are only finitely many possible values of r. Fix r in this range.

For the given r, there is a finite nonempty set of first s_k-defective numbers of order r. Let $n' = p'_1 \cdots p'_r$, where $p'_1 \leq \ldots \leq p'_r$, be one such number. Then since $s_k(n') < n'$, we have that

$$1 > \sum_{1 \leq i_1 < \ldots < i_{r-k} \leq r} \frac{1}{p'_{i_1} \cdots p'_{i_{r-k}}} = \alpha.$$

Now if $n = p_1 \cdots p_r$, where $p_1 \leq \ldots \leq p_r$ and $p_i \geq p'_i$ for $i = 1, \ldots, r$, then

$$\begin{aligned} 1 &> \alpha \\ &\geq \sum_{1 \leq i_1 < \ldots < i_{r-k} \leq r} \frac{1}{p_{i_1} \cdots p_{i_{r-k}}}. \end{aligned}$$

This implies that

$$\sum_{1\leq i_1<\ldots<i_{r-k}\leq r} \frac{1}{p_{i_1}\cdots p_{i_{r-k}}} + \frac{M}{n} \leq \alpha + \frac{M}{n}.$$

If we choose n large enough such that

$$\frac{M}{n} < \frac{1-\alpha}{2},$$

then we have that $n - s_k(n) < M$ as desired. There are only finitely many r for which we must apply this analysis. For each such r, there are only finitely many first s_k-defective numbers of order r, and as already stated, for any s_k-defective number $n = p_1 \cdots p_r$, there is a first s_k-defective number $n' = p'_1 \cdots p'_r$ of order r such that $p'_i \leq p_i$. This proves the lemma. □

The following corollary is a companion to Corollary 4.3, and is an immediate consequence of the preceding lemma.

Corollary 9.11. *Let $k \in \mathbb{N}$, and let $d > 0$. The equation $n - s_k(n) = d$ has at most finitely many solutions.*

Now let us look at some new prime symmetric functions. For $b > 0$, let $f = s + b$. In [11], the case when $b = 1$ was studied It was shown that repeated iteration of f eventually terminates in the cycle $7, 8, 7, 8, ...$, except when $n < 7$, in which case the sequence terminates in 1 or 6. We will show that for any $b > 0$, the function f will ultimately terminate under repeated iteration in an f-cycle, or an f-perfect number. Furthermore, we will show that there are only finitely many such f-cycles. For convenience, we will consider an f-perfect number to be an f-cycle of length 1.

Theorem 9.12. *Let $b \in \mathbb{N}$ and let $f = s + b$. If $n \in \mathbb{N}$, then the f-sequence of n terminates in an f-cycle. Furthermore, there are only finitely many f-cycles.*

Proof. We first need to show that the f-sequence of n must terminate in an f-cycle. If not, then $\lim_{i\to\infty} f^{(i)}(n) = \infty$. Let p be the least prime not dividing b.

If $r > p$, then elements of the sequence $r, r+b, r+2b, \ldots, r+(p-1)b$ can not all be prime since one must be divisible by p but strictly greater than p.

By Lemma 9.10, there is an N such that if $n > N$ is not prime, then $n > f(n) + 2pb$.

When we iterate by f, all strictly increasing subsequences of consecutive terms beginning with $q \geq N$ are of the form $q, q+b, \ldots, q+ib$, where $q, q+b, \ldots, q+(i-1)b$ are prime, $q+ib$ is not prime, and $i \leq p-1$. In this case, we have that $f^{(i+1)}(q) = f(q+ib) < q+ib-2pb < q-pb$.

If $q > \max\{N+pb, f(1)+pb, f(2)+pb, \ldots, f(N)+pb\}$, then $f^{(j)}(q)$ can never ascend above $q+pb$ for any j. Thus the f-sequence of n, must terminate in an f-cycle, since in order for $\lim_{i\to\infty} f^{(i)}(n) = \infty$, it must contain such primes q.

Showing there are only finitely many such cycles is done similarly to the first part of the proof. If there were infinitely many, we could choose arbitrarily large least elements of such cycles. Clearly, we can choose N large enough so that (1), if $n > N$ is the least element of an f-cycle, then n is prime, and (2), if $m > N$ is not prime, then $m > f(m)+pb$. Let $n > N$ be the least element of an f-cycle. Then n is prime, and there is a least $i < p$ such that $n+ib$ is not prime. Then $f^{(i+1)}(n) = f(n+ib) < n+ib-pb < n$. This contradicts that n is the least element of an f-cycle, and the Theorem is proved. □

We will illustrate this Theorem with the case when $b = 2$.

Example 9.13. *Let $f = s + 2$. Then for $n \in \mathbb{N}$, the f-sequence of n terminates in the f-cycle $\{8\}$. Indeed it is easily verified that this is so for $n = 1, 2, \ldots, 35$. It is also readily seen that for n larger, if n is not prime, then $f(n) + 4 < n$. If there is another f-cycle, it must have a least element $q > 35$ which is prime. However, one of $q+2$, and $q+4$ can not be prime. In either case, if we apply f, we get a contradiction to q being the least element.*

Below is a table of complete lists of f-cycles, where $f = s + b$, for $b = 1, \ldots, 10$.

Table 8. $(s+b)$-Cycles

b	$(s+b)$-cycles
1	{1},{6},{7,8}
2	{8}
3	{9},{10}
4	{11,15,12}, {13,17,21,14}
5	{12},{13,18},{14}
6	{14,15}
7	{15}
8	{16}
9	{17,26,24,18}, {22}
10	{18}, {19,29,39,26,25,20}

Bibliography

[1] Alladi, K., and Erdos, P., *On an additive arithmetic function,* Pacific J. Math., **71** (2), (1977), 275-294.

[2] Apostol, T., *Introduction to analytic number theory,* Springer, New York-Berlin-Heidelberg-Hong Kong-London-Milan-Paris-Tokyo, 1976.

[3] Bateman, P. T. and Erdos, P., *Partitions into primes,* Publ. Math. (Debrecen), **4** (1956), 198–200.

[4] Bateman, P. T. and Erdos, P., *Monotonicity of partition functions,* Mathematika, **3** (5), (1956), 1-14.

[5] Bateman, P., and Diamond, H. G., *Analytic Number Theory - An Introductory Course,* World Scientific Publishing Co. Pte. Ltd., 2004.

[6] Butson, A. T., *Generalized Hadamard Matrices,* Proc. Am. Math. Soc., **13**, (1962), 894-898.

[7] Butson, A. T., *Relations among generalized Hadamard matrices, relative difference sets, and maximal length linear recurring sequences,* Canad. J. Math., **15**, (1963), 42-48.

[8] Davenport, H., *Multiplicative Number Theory,* Markham Publishing Company, Chicago, 1966.

[9] de Bruijn, N. G., *On the number of positive integers $\leq x$ and free of prime factors $> y$,* Nederl. Akad. Wetensch. Proc. Ser. A **54** (1951), 50-60.

[10] Burton, D. M., *Elementary Number Theory,* Allan and Bacon, Inc. Boston-London-Sydney, 1976.

[11] Cadogan, C. C. and Callendar, B. A., *A problem on positive integers* New Zealand Math. Mag. 11 (1974), 87-91, 94.

[12] Chandran, V. R., *On generalized unitary perfect numbers,* Math. Student, **61** (1992), 54-56.

[13] Cohen, G. L. and te Riele, H. J. J., *Iterating the sum of divisors function,* Experiment. Math., **5** (1996), 91-100.

[14] Davenport, H., *Multiplicative Number Theory,* Markham Publishing Company, Chicago, 1966.

[15] Dunham, W., *Euler The Master of Us All,* The Mathematical Association of America, (1999).

[16] Erdős, P., and Pomerance, C., *On the largest prime factors of n and $n+1$,* Aequationnes Math., **17** (1978), 311-321.

[17] Grosswald, E., *Partitions into prime powers,* Michigan Math. J., **7** (1960), 97122.

[18] Gupta, H., *Products of parts in partitions into primes,* Res. Bull. Panjab. Univ. (N.S.), **21** (1970), 251-253.

[19] Hardy, G. H. and Ramanujan, S., *Asymptotic formulae for the distribution of integers of various types,* Proc. London Math. Soc., Ser. 2, **16** (1916), 112-132.

[20] Hardy, B. E. and Subbarao, M. V., *On hyperperfect numbers,* Congr. Numer., **42** (1984), 183-198.

[21] Hildebrand, A., *On the number of positive integers $\leq x$ and free of prime factors $> y$,* J. Number Theory **22** (1986), 289-307.

[22] Hildebrand, A. and Tenenbaum, G., *Integers without large prime factors,* J. Theorie des Nombres de Bordeaux, **5** (1993), 411-484.

[23] Hunsucker, J. L. and Pomerance, C., *There are no odd super perfect numbers less than 7×10^{24},* Indian J. Math. **17** (1975), 107-120.

[24] Kerawala, S. M., *A note on the orders of two arithmetic functions $F(n,k)$ and $F^*(n,k)$,* J. Natur. Sci. and Math. **9** (1969), 105-107.

[25] Kerawala, S. M., *On the asymptotic values of $\ln p_A(n)$ and $\ln p_A^{(d)}(n)$ with A as the set of primes,* J. Natur. Sci. and Math. **9** (1969), 209-216.

[26] Lal, M., *Iterates of a Number-Theoretic Function,* Mathematics of Computation **23** (1968), 181-183.

[27] LeVan, Marijo O., *On the order of $F(x, r)$,* J. Natur. Sci. and Math. **10** (1970), 163–166.

[28] Loweke, G. P., *The Lore of Prime Numbers,* Vantage Press, New York-Washington-Atlanta-Los Angeles-Chicago, 1982.

[29] McCranie, J. S., *A study of hyperperfect numbers,* J. of Integer Seq., **3** (2000), 153-157.

[30] Minoli, D., *Issues in nonlinear hyperperfect numbers,* Mathematics of Computation, **34** (1980), 639-645.

[31] Mitsui, T., *On the partitions of a number into the powers of prime numbers,* J. Math. Soc. Jpn., **9** (1957), 428-447.

[32] Montgomery, H. L., and Vaughn, R. C., *Multiplicative Number Theory 1: Classical Theory,* Cambridge University Press, 2006.

[33] Nathanson, M., *Elementary Methods in Number Theory,* Springer, New York-Berlin-Heidelberg-Hong Kong-London-Milan-Paris-Tokyo, 2000.

[34] Niven, I., Zuckerman, H., Montgomery, H., *An Introduction to the Theory of Numbers,* John Wiley and Sons, Inc., New York-Chichester-Brisbane-Toronto-Singapore, 1991.

[35] Pomerance, C., *Multiply perfect numbers, Mersenne primes and effective computability,* Math. Ann. **226** (1977), 195-206.

[36] Richard, L. B., *Asymptotic relations for partitions,* J. Number Theory, **7** (1975), 389-405.

[37] Richard, L. B., *Asymptotic results for partitions (I) and the distribution of certain integers,* J. Number Theory, **8** (1976), 372-389.

[38] Richard, L. B., *Asymptotic results for partitions (II) and a conjecture of Bateman and Erdős,* J. Number Theory, **8** (1976), 390-396.

[39] Richard, L. B., *Asymptotic relations for partitions,* Trans. Amer. Math. Soc., **219** (1976), 379-385.

[40] Suryanarayana, D., *Superperfect Numbers,* Elem. Math., **24** (1967), 16-17.

[41] te Riele, H. J. J., *Hyperperfect numbers with three different prime factors,* Mathematics of Computation, **36** (1981), 297-298.

[42] Tattersall, J. J., *Elementary Number Theory in Nine Chapters,* Cambridge University Press, 1999.

[43] Turán, D., *On a theorem of Hardy and Ramanujan,* J. London Math. Soc., **9** (1934), 274-276.

[44] Eric W. Weisstein. "Newton-Girard Formulas." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Newton-GirardFormulas.html

[45] Eric W. Weisstein et al. "Symmetric Polynomial." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/SymmetricPolynomial.html

[46] Woodford, R., *A Variation on Perfect Numbers,* Integers: Electronic Journal of Combinatorial Number Theory, **4** (2004), A11.

[47] Woodford, R., *Prime Symmetric Divisor Functions,* Masters Thesis: University of British Columbia, Department of Mathematics (2005).

[48] Woodford, R., *On partitions into powers of primes and their difference functions,* Accepted for publication in the Canadian Journal of Mathematics in 2006.

[49] Woodford, R., *Bounds for the Eventual Positivity of Difference Functions of Partitions into Prime Powers,* Journal of Integer Sequences, **10** (2007), 07.1.3.

[50] Woodford, R., *On the Average Orders of a Class of Divisor Functions,* Integers: Electronic Journal of Combinatorial Number Theory, **7** (2007), A13.

Appendix A

Maple Algorithms

Computer searches and much of the experimentation for this exposition was done using Maple. In this section I include some of the algorithms for computing some of the prime symmetric functions. First note that the `with(linalg):`, `with(numtheory):`, `with(combinat):` and `with(LinearAlgebra):` commands should be invoked at the beginning of the worksheet.

The `omega` procedure with input n returns the value of $\omega(n)$, that is, the number of distinct prime factors of n.

```
omega:=proc(n)
om:=Dimension(Vector(ifactors(n)[2]))/2:
return(om):
end proc:
```

The `pvect` procedure with input n returns a vector of length $\Omega(n)$ whose entries are the prime factors of n with repetition.

```
pvect:=proc(n)
c:=0:
v:=array(1..1,1..bigomega(n)):
k:=omega(n):
for i from 1 to k do
for j from 1 to ifactors(n)[2][i][2] do
c:=c+1:
v[1,c]:=ifactors(n)[2][i][1]:
od:od:
return(v):
end proc:
```

To calculate $e_k(n)$, use the following:

```
e:=proc(k,n)
v:=ifactors(n)[2]:
r:=omega(n):
```

```
s:=sum(v[j][2]*v[j][1]^k,j=1..r);
return(s);
end proc:
```

So, for example, if we wished to calculate $e_3(1444)$, we would type

`e(3,1444);`

Maple returns the correct value of `13734`. To compute the function s_k at n, use the following:

```
s:=proc(k,n):
b:=bigomega(n):
v:=pvect(n):
a:=choose(b,k);
sk:=sum(product(v[1,a[i][j]],j=1..k),i=1..binomial(b,k));
return(sk);
end proc:
```

For example, to compute $s_5(174960)$, we would type

`s(5,174960);`

Hitting return yields the correct output of `134942`.

The algorithm we use to compute the final value of the s_k-sequence of n is as follows:

```
term:=proc(k,n):
d:=n:
while d<>s(k,d) do
d:=s(k,d):
od:
return(d);
end proc:
```

For instance, typing

`term(1,16)`

will yield an output of `5`.

We use the `term` procedure to compute $\mathfrak{B}_p(x)$ as follows. You must first input values for `p` and `x`. Use only integer values for `x`.

```
y:=0:
for m from 1 to x do
tt:=term(1,m):
if tt=p then y:= y+1:
fi:
od;
y;
```

To compute the number $t_{s_k}(n)$ of elements in the s_k-sequence of n, use the following variant of the `term` procedure:

```
t:=proc(k,n):
d:=n:
te:=1:
while d<>s(k,d) do
d:=s(k,d):
te:=te+1:
od:
return(te);
end proc:
```

Thus, `t(1,16)`
will yield an output of `4`.