Unbounded Denominators for Hypergeometric Functions and Modular Forms

by

Tobias Bernstein

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

In [2], Atkin and Swinnerton-Dyer conjectured a simple characterization of those Fuchsian groups whose modular forms have integral Fourier coefficient. It has a natural and far-reaching generalization, which we will call the vASD conjecture, to vector-valued modular forms. We confirm vASD conjecture for all 1-dimensional multipliers of $\Gamma(2)$, and set the stage to test it for higher dimensions for $\Gamma(2)$ and other Fuchsian groups. In order to do so, we investigate the similar question for hypergeometric functions, namely when the denominators of its coefficients are unbounded. We do this using p-adic methods, checking when the coefficients are p-adically unbounded for a given p. We generalize the results of [11] for the standard hypergeometric function ${}_2F_1$ to the generalized hypergeometric function ${}_nF_{n-1}$ with rational parameters. In particular, we provide a necessary and sufficient condition for a given prime p, applicable to all but finitely many primes, which determines when its coefficients are p-adically unbounded; these are equivalent but different to the conditions found earlier by Dwork in [7] and by Christol in [6]. Also, we show that the results from [11] concerning when the density of unbounded primes is 0 or 1 respectively extend to the case of ${}_nF_{n-1}$, and strengthen each slightly. We additionally show that the structure of the set of unbounded primes from the ${}_2F_1$ case extends to the ${}_nF_{n-1}$ case. We end with a discussion of modular forms and a brief overview of how the work on hypergeometric functions will apply to the vASD conjecture.

Acknowledgements

Thanks to my supervisor, Dr Gannon, for all his advice, guidance, understanding and support through my Master's. Thanks as well to Dr Troitsky, with whom I briefly worked on a secondary project and who has also provided a great deal of support and interesting learning throughout my Master's, as well as Dr Franc at the University of Saskatoon, and to all the other wonderful profs who taught me during my time at University of Alberta. Thanks to my wonderful officemates and friends Michelle and Jude for all the times spent working together and supporting each other, as well as my thesis writing partner Louisa without whom I would never have gotten through the writing process. Thanks to all my other friends, family and communities who have been there for me along the way, including but not limited to my mom, Mandy, Drew, Ronnie, Em, Julian, my fellow graduate students, Ilara and the board games group, the Temple Beth Ora community, and all my circus pals at Firefly Circus Academy. Thanks as well to NSERC, my supervisor and the Department of Mathematical and Statistical Sciences for their financial support.

Table of Contents

Chapter 1

Introduction

The theory of modular forms began in the field of number theory, alongside the study of elliptic curves. Since then, it has spread across many fields of mathematics, notably including that of vertex operator algebras. Here they appear as characters of certain 'nice' VOAs; however, as opposed to traditional modular forms, these are multi-dimensional 'vector valued' modular forms, with an additional multiplier or representation attached. We provide a brief sketch here; for a more precise introduction, see chapter 4.

A traditional modular form is, roughly speaking, a function on the upper half plane $\mathbb{H}$ which is invariant, up to a specific factor, under the Möbius action of the modular group, $\Gamma(1) = PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z})/\pm I$, where I is the standard 2×2 identity matrix. We usually refer to matrices in $\Gamma(1)$ by one of their pre-images in $SL_2(\mathbb{Z})$; the action we are about to define remains the same regardless of which representative we choose. The Möbius action is defined, for $\tau \in \mathbb{H}$, as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}.$$

Then our requirement for a function f is that, for a given representation ρ of $\Gamma(1)$, we should have

$$f\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z\right] = \rho\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right](cz + d)^{-k} f(z),$$

where $k \in 2\mathbb{Z}$ is called the *weight* of f. This definition can easily be generalized to discrete subgroups of $PSL_2(\mathbb{R})$, known as Fuchsian groups, by requiring that this hold only for matrices in said subgroup, and can be further generalized by allowing the function f to be vector valued and the representation ρ to be multidimensional. In this thesis we will primarily be interested in (vector valued) modular forms for $\Gamma(1)$ and another group, $\Gamma(2)$, which we will define shortly.

Applying certain minor restrictions to the multiplier ρ, we get that our modular forms will always have nice Fourier expansions in a variable which depends on the group (eg $q = e^{2\pi i\tau}$ for $\Gamma(1)$ or $\tilde{q} = e^{\pi i\tau}$ for $\Gamma(2)$). We restrict here to the cases where these Fourier expansions have rational coefficients. A question then becomes, why is it that in some cases, these Fourier expansions will have rational coefficients with denominators which grow large to an unbounded extent, while others have purely integer coefficients?

For example, take $\frac{\eta(3\tau)}{\eta(\tau)} = q^{1/12}(1 + q + 2q^2 + 2q^3 + 4q^4 + 5q^5 + \ldots)$ and $\sqrt{\frac{\eta(3\tau)}{\eta(\tau)}} = q^{1/24}\left(1 + \frac{1}{2}q + \frac{7}{8}q^2 + \frac{9}{16}q^3 + \frac{171}{128}q^4 + \frac{343}{256}q^5 + \ldots\right)$, where η is the Dedekind eta function. Both are modular forms of weight 0 for $\Gamma(1)$, but for different 1-dimensional representations; is there something about their representations that would let us predict that one has an integer q-expansion, while the other has an expansion whose coefficients are not only rational, but whose denominators grow without bound?

As another example, take one-dimensional vector valued modular forms for $\Gamma(2)$. $\Gamma(2)$ is generated freely by two matrices, $A := \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $B := \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$. One-dimensional representations therefore consist of a choice of two complex numbers, $a := \rho(A)$ and $b := \rho(B)$. Non-trivial modular forms for such a ρ will have rational Fourier coefficients if and only if $a, b \in \mathbb{Q}$. Of these so-called rational representations, almost all will have vector valued forms with Fourier coefficients with unbounded denominators. However, as we show in section 4.6, there is a small class which will have Fourier coefficients with integer coefficients, namely those where $a^{24} = 1$ and $a^8 = b^8$. What is special about these representations? Is anything about the representation

captured by the fact that the Fourier coefficients for its modular forms are integral?

Number theorists have long considered almost exclusively what are known as congruence groups, ie groups that contain some group $\Gamma(N)$, defined as the image in $PSL_2(\mathbb{Z})$ of

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) | a \equiv d \equiv 1 \pmod{N}, b \equiv c \equiv 0 \pmod{N} \right\}.$$

Related to these are congruence representations, which have kernels which contain some $\Gamma(N)$, and congruence modular forms, which are fixed by some $\Gamma(N)$. Indeed, these are the Fuchsian groups which show up almost exclusively in applications of modular forms to geometry, algebra and physics. In spite of this, there are very few of them compared to the total number of Fuchsian groups. Hence the question becomes, is there something special about these groups which makes them interesting, or is their widespread usage merely an accident?

Atkin and Swinnerton-Dyer, in their paper [2], suggested that perhaps congruence modular forms $f(\tau)$ (with rational Fourier coefficients) were precisely those with bounded denominator, that is to say that there exists a non-zero M such that $Mf(\tau)$ has integer Fourier coefficients. We generalize this to vector valued modular forms and ask, is it the case that vector valued modular forms with bounded denominator are precisely those whose components are the congruence modular forms?

We need to make certain restrictions in order for this question to be interesting. For example, it turns out that if ρ has infinite image (or equivalently, its kernel has infinite index) then it necessarily both has unbounded denominators for its modular forms, and is also not congruence. It is also known that if the kernel is a congruence group, then its modular forms' Fourier series have bounded denominators. Hence the interesting case is the case where we have a non-congruence representation that has finite image. We also restrict to vector valued modular forms whose components are linearly independent, as they can otherwise be recast as vector valued modular forms for a subrep-

resentation. We call vector valued modular forms with linearly independent components and rational coefficients *full rational* modular forms.

This leads us to the following question:

Question 1.0.1. Let ρ be a congruence representation with finite image, and let $\mathbb{X}(\tau)$ be a full rational vector valued modular form for ρ. Can $X(\tau)$ have integer Fourier coefficients?

The vASD conjecture is the prediction that they can't.

In the case of $\Gamma(1)$, in 1 and 2 dimensions, every finite image representation is a congruence representation, making the question vacuous in this case. It is for this reason that we look, in section 4.6, at the case of $\Gamma(2)$. There, we confirm the conjecture for 1-dimensional representations of $\Gamma(2)$. However, as explained in section 4.7, to answer the question for higher dimensions and especially in two dimensions, the question comes down to looking at the hypergeometric function.

The hypergeometric functions have a long history, having first been introduced by Euler in 1769. As discussed in more detail in Section 2.4, they are solutions to a certain class of differential equations, and as such show up in many practical applications in physics and elsewhere. Their differential equations are defined over the Riemann sphere, with rational functions as coefficients. These rational functions are permitted to have singularities only at $0, 1$, and ∞.

The connection to vector valued modular forms is that the components of any 2-dimensional vector valued modular form for $\Gamma(2)$ is, up to some change of variables $z = z(\tau)$, some hypergeometric function ${}_2F_1$. Moreover, in each dimension $d > 2$, infinite families of vector valued modular forms will have components which are, up to a change of variables, ${}_dF_{d-1}$. For these vector valued modular forms, verifying the vASD conjecture is intimately related to investigating which ${}_nF_{n-1}$ have unbounded coefficients. Hence, it is to this that we now turn.

The hypergeometric functions have a nice series expansion. For the gener-

alized hypergeometric function ${}_nF_{n-1}$, this takes the form

$${}_nF_{n-1}(\alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots, \beta_{n-1}; z) = \sum_{m=0}^{\infty} \frac{(\alpha_1)_m (\alpha_2)_m \cdots (\alpha_n)_m}{(\beta_1)_m (\beta_2)_m \cdots (\beta_{n-1})_m m!} z^m,$$

where the Pochhammer symbol $(\gamma)_m$ is defined as

$$(\gamma)_m = \begin{cases} 1 & m = 0 \\ \alpha(\alpha+1)\cdots(\alpha+m-1) & m \geqslant 1 \end{cases}.$$

In order to answer Atkin-Swinnerton-Dyer, we must answer the similar question about ${}_nF_{n-1}$: when are the denominators of the coefficients of the generalized hypergeometric function unbounded?

It turns out the question can be answered using p-adic valuations. p-adic valuations come from the study of p-adic numbers, an alternative completion of the rational numbers that can in a sense be thought of as a localization around a given prime p. The valuation measures divisibility by p; hence, by using known facts about p-adic valuations, and especially Kummer's result about the p-adic valuation of binomial coefficients, we are able to examine the divisibility of the denominators of the hypergeometric coefficients by each prime p. We call the series unbounded with respect to p if the series' p-adic valuation is unbounded from below, that is, if arbitrarily high powers of p divide the denominators of the coefficients of the series.

Using these methods we are able to generalize and strengthen the results of [11], which proves various results on this topic for ${}_2F_1$, to arbitrary ${}_nF_{n-1}$. In particular, let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_{n-1}$ be a set of hypergeometric parameters. Let D be the least common multiple of their denominators. Then in Proposition 3.3.2, we show that we can find whether any prime $p > D$ is unbounded simply by checking the first prime $q > D$ such that $p \equiv q \pmod{D}$; that is, whether any prime in a given congruence class of primes modulo D is unbounded is entirely determined by looking at a single sufficiently large prime. This is a strengthening of the original paper's theorem, which claims only that if some prime is unbounded, then all primes larger than it in the same congruence class are unbounded as well. This also ensures that the unbounded primes

have a Dirichlet density of $\frac{N}{\phi(D)}$, where ϕ is the Euler totient function. We then, in Theorem 3.4.2 use this to characterize when a hypergeometric function has unbounded denominators for all but finitely many primes, and in Theorem 3.4.4 when the hypergeometric function has bounded denominators for all but finitely many primes. We of course also generalize the necessary and sufficient conditions for the hypergeometric function ${}_2F_1$ to have unbounded denominators for a given prime, which again apply to all but a finite number of primes. Namely, we introduce the concept of numbers being *semi-interlaced*, and in theorem 3.2.6 show that for sufficiently large p, the question of whether ${}_nF_{n-1}$ has unbounded coefficients is simply the question of whether every p-adic column of the expansion of its parameters with 1 subtracted is semi-interlaced. Since these expansions will be periodic, this is a finite test. These results form the main body of the original work of this thesis.

The structure of this thesis is as follows.

In Chapter 2, we deal with the hypergeometric function and p-adic numbers. We begin with Section 2.1, an introduction to p-adic numbers and valuations. In Section 2.2 we introduce the reader to p-adic expansions and arithmetic. Section 2.3 provides proofs for the results in the area of p-adics which we will use to prove our main results. Section 2.4 gives a similar background for hypergeometric functions and generalized hypergeometric functions.

Chapter 3 provides the meat of the thesis; this is where the results from [11] are generalized. In s=Section 3.1 we prove some preliminary results and set up some definitions. We then, in Section 3.2, answer the question for the hypergeometric function of when the coefficients have unbounded denominators with respect to given primes. In Section 3.3, we show that the set of unbounded primes has a Dirichlet density, as well as a result about the structure of the set of unbounded primes. In Section 3.4, we show when the set of unbounded primes has density 0 or 1.

In Chapter 4, we give an outline of the theory of modular forms and the Atkin-Swinnerton-Dyer conjecture. We start in Section 4.1 with an introduction to the standard theory of modular forms. In Section 4.2 we then give a brief overview of the generalization to arbitrary Fuchsian groups, especially $\Gamma(2)$. In Section 4.3 we generalize to vector valued modular forms. We then

give a brief overview of the structure of some structure and theory for vector valued modular forms for $\Gamma(1)$ and $\Gamma(2)$ in Section 4.4. Section 4.6 then gives a brief overview of the story of the Atkin-Swinnerton-Dyer conjecture and answers the vASD conjecture in the affirmative in the case of 1-dimensional vector valued modular forms for $\Gamma(2)$. We end in Section 4.7 with a brief overview of how the two parts of this thesis come together for higher dimensional vector valued modular forms.

In the conclusion, Chapter 5, we provide some possible future directions for this work.

Chapter 2

p-adic Numbers and Hypergeometric Functions

2.1 p-adic Numbers and Valuations

We begin with an introduction to the basics of p-adic numbers. For a more thorough introduction, see for example [17] or [21].

Recall that one method of constructing the real numbers $\mathbb{R}$ is as a completion of the rationals, $\mathbb{Q}$, for instance by requiring all Cauchy sequences to converge. In so doing, we implicitly (or explicitly) use the standard metric on the rationals, the Euclidean metric, which measures the distance between two numbers using the standard absolute value. However, this is not the only metric which one can define on the rational numbers, nor is it the only possible absolute value. Another method of measuring size and distance is, roughly speaking, to look at divisibility by some prime p; completing the rationals with respect to this metric gives us the p-adics.

More precisely, and more algebraically, we can approach p-adic numbers via the idea of a *valuation* and the related concept of *absolute values*. For more on valuations and the associated concept of valuation rings, see for example chapter 11 of [8] or chapter XII of [24]. We define only discrete valuations here, using the definition most appropriate to our needs; more general definitions exist, as do definitions written using multiplicative notation.

Definition 2.1.1. *A* discrete valuation *on a field* $\mathbb{K}$ *is a function* $\nu : \mathbb{K} \to \mathbb{Z} \bigcup \{\infty\}$ *such that for any* $a, b \in \mathbb{K}$*,*

- $\nu(ab) = \nu(a) + \nu(b)$*,*
- $\nu(a + b) \geq \min(\nu(a), \nu(b))$*, with equality when* $\nu(a) \neq (\nu(b))$*,*
- $\nu(a) = \infty$ *if and only if* $a = 0$

The last condition can be dropped if we instead use only the multiplicative group of the field, which will also remove the need for the use of infinity, however this notation will be more convenient for us when converting to an absolute value.

Definition 2.1.2. *An* absolute value *on a field* $\mathbb{K}$ *is a function* $|\cdot| : \mathbb{K} \to \mathbb{R}$ *such that for any* $a, b \in \mathbb{K}$*,*

- $|ab| = |a||b|$
- $|a + b| \leqslant |a| + |b|$
- $|a| \geqslant 0$ *and* $|a| = 0$ *if and only if* $a = 0$*.*

We are most interested in a particular class of valuations and absolute values, defined over the rationals.

Definition 2.1.3. *Given a prime* p*, the associated* p-adic valuation *is the function* $\nu_p : \mathbb{Q} \to \mathbb{Z}$ *:*

- *For* $a \in \mathbb{Z}$*, let* $a = p^k b$*, where* $p \nmid b$*. Then* $\nu_p(a) = k$*. That is, the valuation of a whole number is the maximum power of* p *which divides it. Here we define* $\nu_p(0) = \infty$*, both to satisfy the requirements of a valuation and since* p^k *divides* 0 *for all* k*.*
- *For* $\frac{m}{n} \in \mathbb{Q}$*, where without loss of generality* m *and* n *are co-prime,* $\nu_p(\frac{n}{m}) = \nu_p(n) - \nu_p(m)$*.*

It is an easy exercise to check that this does indeed define a discrete valuation. We can now use this valuation to define an absolute value as follows:

Definition 2.1.4. *Given a prime p, we define the* p-adic absolute value $|\cdot|_p : \mathbb{Q} \to \mathbb{R}^+$ *for* $a \in \mathbb{Q}$ *by*

$$|a|_p = \frac{1}{p^{\nu_p(a)}}.$$

Again, seeing that this does in fact define an absolute value is an easy exercise, and follows mostly from the definition of a valuation.

In fact, we can use this absolute value to define a norm as follows.

Definition 2.1.5. *Given a prime p, the associated* p-adic norm *is the function* $||\cdot||_p : \mathbb{Q} \to \mathbb{R}_+$ *defined for* $a \in \mathbb{Q}$ *by*

$$||a||_p = |a|_p$$

That this is a norm in relation to the p-adic absolute value follows immediately from the definition of absolute value.

As a matter of interest, and to show the importance of the p-adic norm, we state the following well-known theorem.

Theorem 2.1.1 (Ostrowski's Theorem)**.** *Up to equivalence, any non-trivial norm on the rationals is either a p-adic norm, or the Euclidean norm.*

For a proof see for example page 3 of [22].

We are now ready to define the p-adic numbers.

Definition 2.1.6. *Given a prime number p, the completion of $\mathbb{Q}$ with respect to the p-adic norm is called the field of* p-adic numbers*, denoted $\mathbb{Q}_p$.*

There are, of course, many extensions of p-adic fields, however the base set of p-adic numbers will be sufficient for our needs.

By Ostrowski's Theorem, the p-adics are, in essence, the only completion of the rationals other than the reals. As such, we can often use the p-adics to find information about the rationals. One common example of this is what is known as the Hasse Principle (see [17, pg 76-77]): the idea that one can study the solutions of Diophantine equations over the rationals by studying their solutions over all the p-adics as well as the reals. This can be thought of as a local-global principle, where the p-adics are thought of as being the local

context around the corresponding prime (and the reals are the local context around the 'infinite prime'). Although it is not in general true that solutions exist for the rationals if and only if they exist in the p-adics for all primes as well as the reals, this does hold for quadratic forms.

2.2 p-adic Expansions and Arithmetic

An important property of p-adic numbers (and, in fact, of p-adic fields in general) is the existence of a unique (for each p) p-adic expansion for any p-adic number, and in particular for the rational numbers. We will make use of this extensively, so we will spend some time introducing the reader to these expansions and how they interact with the p-adic operations, in particular addition. As much of this is well-known, we will focus on examples, and leave most basic theorems without proof.

Theorem 2.2.1. *Given a prime number p, every p-adic number a has a unique expansion, called the* p-adic expansion*, of the form*

$$a = \sum_{k=m}^{\infty} a_k p^k,$$

where $m \in \mathbb{Z}$, $a_m \neq 0$, and for every k, $a_k \in \mathbb{Z}$, $0 \leqslant a_k < p$.

For a proof, see for example [21, pg 22-25].

Definition 2.2.1. *The expansion given in Theorem 2.2.1 is called the* canonical expansion*, or simply the* p-adic expansion*. The numbers a_k are called the* digits *of a, with each a_k being called the* k-th digit *of a.*

In particular, it is well known that a p-adic number has an eventually periodic canonical expansion if and only if it is a rational number; see for example [21, pg 30-32].

This expansion is the source of the intuition that p-adics are local. In number theory, we can take $\mathbb{Q}$ as a global object, which is analogous to global objects in geometry such as the Riemann sphere $\mathbb{CP}^1$. In algebraic geometry these are captured by their function field, in this case the field of rational

functions. Primes in number theory are analogous to points. Evaluating a function at a point is analogous to reducing a rational number modulo p. One can examine the local behaviour of a function at a point by looking at its Laurent expansion; hence this is analogous to the p-adic expansion of a rational number.

The notation above can be somewhat hard to visualise, especially in relation to the commonly used operations of addition and multiplication. As such, it is common to write out p-adic numbers using only their digits without the powers of p, either padded with zeros if $m \geqslant 0$ or with a decimal between a_{-1} and a_0 if $m < 0$. That is, rather than writing

$$\sum_{k=m}^{\infty} a_k p^k,$$

it is common to simply write

$$a_m a_{m+1} \dots a_{-1}.a_0 a_1 a_2 \dots \quad \text{for } m < 0$$

or

$$a_0 a_1 \dots a_m a_{m+1} a_{m+2} \dots \quad \text{for } m \geqslant 0,$$

where in the second case we set $a_i = 0$ for all $i < m$.

In the case where the expansion is eventually periodic, we will use the usual overbar notation, that is, for example,

$$a_0 a_1 \dots a_{k-1} \overline{a_k a_{k+1} a_{k+2}}$$

means that the expansion is continued by repeating $a_k a_{k+1} a_{k+2}$ to infinity. Additionally, in the case where there are only a finite number of non-zero terms, it is traditional to omit the trailing zeros. In this case, we call the number of digits written the *length* of the p-adic expansion. Note that the length always begins its count at or before the zeroth digit.

As it may be hard to distinguish between this and regular, decimal digits, where necessary we will write $[a]_p$ when a is written in p-adic digits and $[a]_{10}$ when it is written in decimal notation. Since 10 is not a prime, no confusion

should arise.

The result of this notation is that positive integers' canonical expansions are precisely their p-nary expansions, only reversed. For example, if we take $p = 5$ and $a = 2482 = 3 \cdot 5^4 + 4 \cdot 5^3 + 4 \cdot 5^2 + 5 + 2$, then its p-nary expansion would be 34412; we write its p-adic expansion as $[21443]_5$.

Negative integers, by contrast, always have infinite series expansions. To see why, let's look at addition using the canonical expansions. Rather than getting bogged down in notation, we will show this using an example. Let's again take $p = 5$, $a = [2482]_10 = [21443]_5$, and $b = [54]_10 = [412]_5$.

Addition in the p-adics is done much as addition in any n-ary system, except with the possibility of infinite carries. In our example above, then, we have

$$\begin{array}{r} 21143 \\ +\ 412 \end{array}$$

Adding $2 + 4 = 6 = 5 + 1$, we get the zeroth digit as a 1, and a carry to the next column; there, we get $1 + 1 = 2$, plus the carry, making the first digit a 3; the second digit is $1 + 2 = 3$, the third is $4 + 0 = 4$ and the fifth is $3 + 0 = 3$. All subsequent digits are zero. That is, the final answer is 13343.

Now let's look at a negative number, for example -1. By definition, we must have $1 + (-1) = 0$. For simplicity, let's keep working with $p = 5$. We need to have a p-adic number $x = x_0x_1x_2 \ldots x_n \ldots$ such that

$$\begin{array}{r} x_0x_1x_2 \ldots x_n \ldots \\ +\ 1 \\ \hline 0\ \ 0\ \ 0\ \ \ldots 0\ \ \ldots \end{array}$$

Then we know that we need $x_0 + 1 = 0 \pmod 5$; since $0 \leqslant x_0 \leqslant 5$, this implies that $x_0 = 4$. The first column now has a carry, meaning that we must again have $x_1 + 1 = 0 \pmod 5$, meaning $x_1 = 4$. Continuing this argument to infinity, we get that $x_n = 4$ for all n, that is, $-1 = [\ \overline{4}\ ]_5$. In fact, the same

argument can be made for any p, giving $-1 = [\ \overline{(p-1)}\ ]_p$ for all primes p.

Similarly, for any positive integer, to get 0 we will be forced to create a stream of infinite carries with a summand of finite length; as such, our p-adic expansion will need infinite non-zero entries.

Multiplication, similarly, works much as it does in n-ary arithmetic, but with the possibility of infinite carries. As it will be less important to this thesis, we refer the reader again to any introductory book on p-adic numbers, such as those mentioned above.

Of particular note is that for rationals $x = \frac{a}{d}$, where $\gcd(a, d) = 1$, if p does not divide d then x has an expansion whose only non-zero digits have non-negative indices. p-adic numbers which have this property are called *p-adic integers* and form a ring denoted $\mathbb{Z}_p$. The multiplicative group $\mathbb{Z}_p^\times$ is made up of precisely those p-adic integers whose zeroth digit is non-zero. In particular, rational numbers $x = \frac{a}{d}$ are units in this ring if and only if p divides neither a nor d.

2.3 Kummer's Theorem and Other Simple Results in p-adics

The following simple results turn out to be quite important for our purposes. The lemmas and proofs in this section follow those of [11].

Lemma 2.3.1. *Let p be a prime, $x = \frac{a}{d} \in \mathbb{Q} \cap \mathbb{Z}_p^\times$, where without loss of generality $\gcd(a, d) = 1$. Then x has a purely periodic p-adic expansion if and only if $x \in [-1, 0)$. Moreover, in this case, the minimal period M of the expansion is the multiplicative order of p in $(\mathbb{Z}/d\mathbb{Z})^\times$, that is, in the multiplicative group of $\mathbb{Z}/d\mathbb{Z}$.*

Proof. Let $x = \frac{a}{d} \in \mathbb{Q} \cap \mathbb{Z}_p^\times$, $\gcd(a, d) = 1$, with purely periodic p-adic expansion $x = \overline{x_0 x_1 x_2 ... x_{M-1}}$, where without loss of generality M is the minimal period. Take $y = [x_0 x_1 x_2 ... x_{M-1}]_p$. We will show that $x = \frac{y}{1-p^M}$, that is, that $x \cdot (1 - p^M) = y$.

First note that p^M has p-adic expansion with Mth digit 1 and all other digits 0. Then similarly to -1, $-p^M$ will have expansion $00 \ldots 0\overline{(p-1)}$, where

the first $p-1$ is at the Mth digit. Then $1-p^M = 1000\ldots0\overline{(p-1)}$, where again the first $p-1$ is at the Mth digit. Thus, multiplying $x \cdot (1-p^M)$, we get the first M digits as the first repeat of the sequence, whereas all following repeats of the sequence and their carries are cancelled by the infinite $p-1$ shifted copies of the sequence. That is, $x \cdot (1-p^M) = y$, or $x = \frac{y}{1-p^M}$. In particular, since by definition of p-adic expansion we have $0 < y \leqslant p^M - 1$, we have that $x \in [-1, 0)$. Moreover, since both y and $1-p^M$ are integers, we have that $d|1-p^M$. Stated differently, $p^M \equiv 1 \pmod{d}$, so that if we let N be the multiplicative order of p in $\mathbb{Z}/d\mathbb{Z}$, N must divide M.

Conversely, suppose that $x = \frac{a}{d} \in \mathbb{Q} \cap \mathbb{Z}_p^\times \cap [-1, 0)$, $\gcd(a, d) = 1$. Let N be the multiplicative order of p in $\mathbb{Z}/d\mathbb{Z}$, as above. Without loss of generality, assume $a > 0, d < 0$. Then we have that $p^N \equiv 1 \pmod{d}$, so that there exists some $u \in \mathbb{Z}_{\geqslant 0}$ such that $1-p^N = du$. Then multiplying both the numerator and denominator by u, we get $x = \frac{nu}{1-p^N}$. Since we know that $-1 \leqslant x < 0$, we must have $0 < nu \leqslant p^N - 1$. In particular, nu must have a finite p-adic expansion with a length of at most N digits.

Doing similar calculations to those above, we find that the p-adic expansion of $\frac{1}{1-p^N}$ is

$$\frac{1}{1-p^N} = \overline{1\underbrace{00\ldots0}_{N-1}},$$

that is, a repeating pattern of a 1 followed by $N-1$ zeroes. Multiplying this by an integer whose p-adic expansion has length at most N, such as nu, will result in a number with a purely periodic expansion, whose minimal period M divides N. By above, we must also have N dividing M, hence the two must be equal. □

In the coming sections we will largely be dealing with rational numbers, for instance x, between 0 and 1, and looking at the p-adic expansions of $x-1$ rather than that of x itself. It is for that reason that we introduce the following potentially confusing notation.

Notation 2.3.1. *Let $x \in \mathbb{Q} \cap [0, 1)$. For any prime p, we define $x^j(p)$ to be the jth coefficient of p^j in the p-adic expansion of $x-1$, that is, $x - 1 = \sum_{j=m}^{\infty} x^j(p)p^j$ in the p-adic norm. When there will be no confusion, we will*

sometimes write simply x^j. For any set A, we will refer to the set $\{x^j(p)|x \in A\}$ as the j-th column *(with respect to p) of A. (Note that this means that the numbering of columns for p-adic integers starts at 0, not 1.) In cases where we want to reference the jth p-adic coeffecient of x itself, we will instead write $x(j,p)$ or simply $x(j)$.*

Additionally, we will sometimes want to look at only part of a p-adic expansion, say the first j digits. Formally, this is a sort of truncation which gives us an integer. For that reason we define the following operators.

Definition 2.3.2. *Let $x \in \mathbb{Z}_p$. For any $j \geqq 0$, there is a unique integer, which we will denote $\tau_j(x)$, such that $0 \leqslant \tau_j(x) < p^j$ and $\tau_j(x) \equiv x \pmod{p^j}$. Then τ_j is a map $\tau_j : \mathbb{Z}_p \to \mathbb{Z}$, which we call the jth* truncation operator*.*

Note that when we take a rational number x between 0 and 1 and subtract 1 from it, we end up, by Lemma 2.3.1, with a rational number with a purely periodic expansion so long as p divides neither the numerator nor the denominator of $x - 1$. In fact, we can say more than this; as the following lemma shows, there is a relatively simple formula for the values of digits of the expansion.

Notation 2.3.3. *For a real number x, we as usual let $\lfloor x \rfloor$ be the* floor *of x, that is the unique integer such that $\lfloor x \rfloor \leqslant x \leqslant \lfloor x \rfloor + 1$. We also define $\{x\}$ to be the* fractional part *of x, that is, $\{x\} = x - \lfloor x \rfloor$.*

Lemma 2.3.2. *Let $x = \frac{a}{d} \in \mathbb{Q} \cap (0,1)$, $\gcd(a,d) = 1$, p a prime such that $x - 1 \in \mathbb{Z}_p^\times$. Let M be any period of $x - 1$. That is, using our above notation, let $x - 1 = \overline{x^0 x^1 x^2 \dots x^{M-1}}$. Then for all $0 \leqslant j < M$, we have*

$$x^j = \lfloor \{-p^{M-1-j} x\} p \rfloor.$$

Proof. Let x, M be as in the hypothesis. Then we have that $-x$ and $x - 1$ are both in $(-1, 0) \cap \mathbb{Q}$, and moreover, since they both have the same denominator, p does not divide the denominator of either. Hence, since both are rational numbers, both have periodic expansions and are p-adic integers. In fact, since we have $-x + (x - 1) = -1 = \overline{(p-1)}$, we know that $-x = \overline{(p - 1 - x^0)(p - 1 - x^1) \dots (p - 1 - x^{M-1})}$.

Let $\{-p^{M-1-j}x\} = \frac{r}{d}$, $n = \lfloor -p^{M-1-j}x \rfloor$. Then multiplying the expansion of $-x$ by p^{M-1-j}, we have that

$$\frac{r}{d} + n = -p^{M-1-j}x = \underbrace{00\ldots0}_{M-1-j \text{ digits}} \overline{(p-1-x^0)(p-1-x^1)\ldots(p-1-x^{M-1})}.$$

To prove that $x^j = \lfloor \{-p^{M-1-j}x\}p \rfloor$, we must show that $0 \leqslant \frac{r}{d}p - x^j < 1$. In order to accomplish this, we first write $\frac{r}{d} + n = (\frac{r}{d} - 1) + (n+1)$. Now by definition, $0 \leqslant \frac{r}{d} < 1$, so that by Lemma 2.3.1 we must have that $\frac{r}{d} - 1$ is purely periodic. Hence, we can find $n+1$ as the unique p-adic number which, when subtracted from $-p^{M-1-j}x = \frac{r}{d} + n$, gives a number with a purely periodic expansion. It is clear, then, from the above expansion of $-p^{M-1-j}x$, that we must have $-(n+1) = [(p-1-x^{j+1})(p-1-x^{j+2})\ldots(p-1-x^{M-1})]_p$, and hence

$$\frac{r}{d} - 1 = \overline{(p-1-x^{j+1})\ldots(p-1-x^{M-1})(p-1-x^1)\ldots(p-1-x^j)}.$$

Truncating this, we get $\tau_M(\frac{r}{d} - 1) = (p-1-x^{j+1})\ldots(p-1-x^{M-1})(p-1-x^1)\ldots(p-1-x^j)$, which allows us to isolate x^j as the final digit. As we showed above, truncation is equivalent to multiplication by $p^M - 1$. If we subtract the final digit of a p-adic number of finite length, we necessarily get a finite p-adic expansion of strictly shorter length; hence we have that

$$0 \leqslant (1 - \frac{r}{d})(p^M - 1) - (p - 1 - x^j)p^{M-1} \leqslant p^{M-1} - 1.$$

Expanding, rearranging and dividing by $-p^{M-1}$, we get that this is equivalent to

$$\frac{r}{d} \cdot \frac{1}{P^{M-1}} \leqslant p\frac{r}{d} - x^j \leqslant 1 - (\frac{1}{p^{M-1}} - \frac{r}{d} \cdot \frac{1}{p^{M-1}}).$$

Since $0 \leqslant \frac{r}{d} < 1$, we have also $0 < (\frac{1}{p^{M-1}} - \frac{r}{d} \cdot \frac{1}{p^{M-1}}) < 1$. Thus, $0 \leqslant p\frac{r}{d} - x^j < 1$, as required.

□

Corollary 2.3.3. *Let $x, y \in \mathbb{Q} \cap (0,1), x \neq y$. Let D be the least common multiple of the denominators of x and y, and let p be a prime such that $p > D$*

and $x-1, y-1 \in \mathbb{Z}_p^\times$. *Then for all* $j \geqslant 0$, $x^j(p) \neq y^j(p)$.

Proof. Suppose, by way of contradiction, that $x^j(p) = y^j(p)$. By Lemma 2.3.2, this is equivalent to

$$\lfloor \{-p^{M-1-j}x\}p \rfloor = \lfloor \{-p^{M-1-j}y\}p \rfloor.$$

Let $\{-p^{M-1-j}x\} = \frac{a}{D}$, $\{-p^{M-1-j}y\} = \frac{b}{D}$. Then we have

$$\begin{aligned} \left\lfloor \frac{a}{D}p \right\rfloor &= \left\lfloor \frac{b}{D}p \right\rfloor \\ \implies 0 \leqslant |\frac{a}{D}p - \frac{b}{D}p| &\leqslant 1 \\ \implies 0 \leqslant |a-b| &\leqslant \frac{D}{p}. \end{aligned}$$

Since $p > D$ and a, b are integers, this implies that $a = b$, that is, that $\{-p^{M-1-j}x\} = \{-p^{M-1-j}y\}$. But then this means that since $-p^{M-1-j}x, -p^{M-1-j}y$ have the same fractional part, we must have that $-p^{M-1-j}(x-y) \in \mathbb{Z}$. However, since $0 < |x-y| < 1$, and since $p > D$ implies $\gcd(p, D) = 1$, this is impossible. Hence, we must have $x^j(p) \neq y^j(p)$. □

Corollary 2.3.4. *Let* $x = \frac{a}{d} \in \mathbb{Q}$ *such that* $0 < x < 1$. *Let* p *be a prime such* $p > D$ *(in particular,* $\gcd(d, p) = 1$*). Then for all* j, $x^j(p) \neq p-1$.

Proof. Let M be the order of $p \pmod{d}$. By Lemma 2.3.2, $x^j = \lfloor \{-p^{M-1-j}x\}p \rfloor$. Since $x = \frac{a}{d}$ and p is, in particular, an integer, we have that $\{-p^{M-1-j}x\} \leqslant \frac{d-1}{d}$. Then $\{-p^{M-1-j}x\}p \leqslant p - \frac{p}{d} < p-1$, so that $\lfloor \{-p^{M-1-j}\alpha\}p \rfloor < p-1$. □

For our purposes we will need to know about the valuations of binomial coefficients. It turns out that this information is encoded within the number of carries performed during addition. This important, yet fairly elementary, result will be the basis of our upcoming analysis. To prove it, we will need the following well-known formula, attributed to Legendre ([25]) and now often left as an exercise, eg in [17][pg 113]. We provide a proof here for completeness.

Lemma 2.3.5 (Legendre)**.** *Let n be a positive integer, and let $s_p(n)$ be the sum of the p-adic digits of n. Then*

$$\nu_p(n!) = \frac{n - s_p(n)}{p-1}.$$

Proof. First note that we can count the valuation of $n!$ by counting first the number of multiples of p less than n, then the multiples of p^2, and so on, since each of these and only these will contribute to the valuation of $n!$. The number of factors in each of these cases is $\left\lfloor \frac{n}{p^i} \right\rfloor$. That is,

$$\nu_p(n!) = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor.$$

Now, let's let $n = n(0) + n(1)p + n(2)p^2 + \cdots + n(k)p^k$ be the p-adic expansion of n. We write it out in full since we will be using it as a sum. We can then rewrite the above formula as

$$\begin{aligned}
\sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor &= \sum_{i=1}^{\infty} \left\lfloor \sum_{j=0}^{k} (\frac{n(j)p^j}{p^i}) \right\rfloor & &= \sum_{i=1}^{k} \sum_{j=i}^{k} (n(j)p^{j-i}) \\
&= \sum_{j=0}^{k} \sum_{i=1}^{j} (n(j)p^{j-i}) & &= \sum_{j=0}^{k} n(j) \sum_{i=0}^{j-1} p^i \\
&= \sum_{j=0}^{k} n(j) \frac{1-p^j}{1-p} & &= \frac{1}{p-1} \sum_{j=0}^{k} n(j)(p^j - 1) \\
&= \frac{1}{p-1} \sum_{j=0}^{k} n(j)p^j - n(j) & &= \frac{1}{p-1}(n - s_p(n)).
\end{aligned}$$

□

We can now prove Kummer's result. We will first show it for positive integers, then extend the result to p-adic integers. We will first introduce some notation which we will use throughout the rest of this thesis.

Notation 2.3.4. *Let $\gamma \in \mathbb{Z}_p$, $m \in \mathbb{Z}_{\geqslant 0}$. Define*

$$c_p^j(\gamma, m) = \begin{cases} 1 & \text{if adding } m+\gamma \text{ } p\text{-adically causes a carry from the } j\text{th digits} \\ 0 & \text{otherwise.} \end{cases}$$

We denote the total number of carries by $c_p(\gamma, m) := \sum_{j=0}^{\infty} c_p^j(\gamma, m)$.

Theorem 2.3.6 (Kummer)**.** *Let $\gamma \in \mathbb{Z}_p$, $m \in \mathbb{Z}_{\geqslant 0}$. Then the valuation of the binomial coefficient $\binom{\gamma+n}{n}$ is exactly the number of carries when adding $\gamma + n$ in the p-adics, that is,*

$$\nu_p\left(\binom{\gamma+n}{n}\right) = c_p(\gamma, n).$$

Proof. Suppose first that $\gamma \in \mathbb{Z}_{\geqslant 0}$. Then we have that

$$\binom{\gamma+n}{n} = \frac{(\gamma+n)!}{\gamma! n!}.$$

Hence by Lemma 2.3.5 and the properties of valuations, we have that

$$\begin{aligned} \nu_p\left(\binom{\gamma+n}{n}\right) &= \nu_p((\gamma+n)!) - \nu_p(\gamma!) - \nu_p(n!) \\ &= \frac{(\gamma+n) - s_p(\gamma+n)}{p-1} - \frac{\gamma - s_p(\gamma)}{p-1} - \frac{n - s_p(n)}{p-1} \\ &= \frac{s_p(\gamma) + s_p(n) - s_p(\gamma+n)}{p-1}. \end{aligned}$$

For notational simplicity, let's call $x := \gamma + n$. Then we have, for each $j \geqslant 0$, $x(j) = \gamma(j) + n(j) + c_p^{j-1}(\gamma, n) - p \cdot c_p^j(\gamma, n)$, where we define $c_p^j(\gamma, n) = 0$. Since all these are integers, we can find some index such that all following digits are zero in all three of these numbers' p-adic expansions. Call this index k. Then, since both $c_p^k(\gamma, n)$ and $c_p^{-1}(\gamma, n)$ must equal 0, we have

$$\begin{aligned} &s_p(\gamma) + s_p(n) - s_p(\gamma+n) \\ =&\sum_{j=0}^{k} \gamma(j) + \sum_{j=0}^{k} n(j) - \sum_{j=0}^{k} (\gamma(j) + n(j) + c_p^{j-1}(\gamma, n) - p \cdot c_p^j(\gamma, n)) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=0}^{k} p \cdot c_p^j(\gamma, n) - c_p^{j-1}(\gamma, n) = \sum_{j=0}^{k} p \cdot c_p^j(\gamma, n) - \sum_{j=-1}^{k-1} c_p^j(\gamma, n) \\
&= \sum_{j=0}^{k} p \cdot c_p^j(\gamma, n) - c_p^j(\gamma, n) = (p-1) \sum_{j=0}^{k} c_p^j(\gamma, n) \\
&= (p-1) c_p(\gamma, n).
\end{aligned}
$$

Hence, combining these two, we get

$$\nu_p\left(\binom{\gamma+n}{n}\right) = c_p(\gamma, n)$$

for $\gamma, n \in \mathbb{Z}_{\geqslant 0}$, as required.

Now suppose that $\gamma \in \mathbb{Z}_p$. Then either there are infinitely many carries when adding $\gamma + n$, or else there is some finite index l_0 such that no carries occur beyond this index. Then for any index $l \geqslant l_0$, $c_p(\gamma, n) = c_p(\tau_l(\gamma), n)$.

Recall that

$$\binom{\gamma+n}{n} = \frac{(\gamma+1)(\gamma+2)\dots(\gamma+n)}{n!}.$$

The denominator will always have finite valuation; the numerator will have finite valuation so long as none of the factors is zero, that is, so long as $\gamma \notin \{-1, -2, \dots, -n\}$. Then supposing we have a finite valuation, there is some index l_1 such that for all $l \geqslant l_1$, for $1 \leqslant k \leqslant n$,

$$\nu_p(\gamma + k) = \nu_p(\tau_l(\gamma + k)) = \nu_p(\tau_l(\gamma) + k).$$

Then

$$\nu_p\left(\binom{\gamma+n}{n}\right) = \nu_p\left(\binom{\tau_l(\gamma)+n}{n}\right).$$

Let $L = \max(l_0, l_1)$. Then supposing $\gamma \notin \{-1, -2, \dots, -n\}$ and $c_p(\gamma, n) < \infty$, we have $c_p(\gamma, n) = c_p(\tau_L(\gamma), n)$ and $\nu_p\left(\binom{\gamma+n}{n}\right) = \nu_p\left(\binom{\tau_L(\gamma)+n}{n}\right)$. Since $\tau_L(\gamma)$ is an integer, by the first part of our proof we have $\nu_p\left(\binom{\tau_L(\gamma)+n}{n}\right) = c_p(\tau_L(\gamma), n)$. Hence, we have $\nu_p\left(\binom{\gamma+n}{n}\right) = c_p(\gamma, n)$.

It remains to show that $c_p(\gamma, n) = \infty$ if and only if $\nu_p\left(\binom{\gamma+n}{n}\right) = \infty$. We have already shown that $\nu_p\left(\binom{\gamma+n}{n}\right) = \infty$ if and only if $\gamma \in \{-1, -2, \dots, -n\}$.

Note that, as strictly negative integers, each of $-1, \ldots, -n$ have infinitely many non-zero digits. However, for $\gamma \in \{-1, -2, \ldots, -n\}$, we have that $\gamma + n \in \mathbb{Z}_{\geqslant 0}$, hence $\gamma + n$ has only finitely many non-zero digits. Since n has only finitely many non-zero digits, this can only occur if there are infinitely many carries. Hence for $\gamma \in \{-1, -2, \ldots, -n\}$, $c_p(\gamma, n) = \infty$.

Conversely, suppose that $c_p(\gamma, n) = \infty$. Since n has finitely many non-zero digits, γ must have all digits past some index K be $p-1$. This can only occur if $\gamma = m - p^N$ for some $m \in \mathbb{Z}_{\geqslant 0}$, $0 \leqslant m < p^N$, $N \geqslant 1$. In particular, γ is a strictly negative integer. Since infinite carries from a finite summand will cause all but finitely many of the the $p-1$s to become 0s, this forces $\gamma + n$ to be a finite integer. That is, $m - p^N + n \geqslant 0$, or equivalently, $m - p^N \geqslant -n$. Hence $-n \leqslant m - p^N < 0$, so that $\gamma \in \{-1, -2, \ldots, -n\}$, as required.

□

The p-adics are a vast and interesting area which have been applied in many areas of mathematics, far too vast for the scope of this thesis. These results, however, are sufficient for our needs, which largely consist of applying p-adic methods to rational contexts. As such, we will now move on from this area, and onto the focus of our application.

2.4 Hypergeometric Functions and Generalized Hypergeometric Functions

We begin with a very brief history of the hypergeometric function; for a more in-depth history, see [18], from which most of these historical notes have been taken. The hypergeometric function and its generalizations are solutions to a certain class of differential equations, which show up in many areas of mathematics and physics. They were first introduced by Euler in 1769. As we will only be dealing with rational parameters, we will give definitions in these terms, although more general definitions exist.

Definition 2.4.1. *Let $\alpha, \beta, \gamma \in \mathbb{Q}$. The associated* (standard) hypergeometric

equation *is*

$$x(1-x)\frac{d^2y}{dx^2} + [\gamma - (\alpha+\beta+1)x]\frac{dy}{dx} - \alpha\beta y = 0.$$

The associated (standard) hypergeometric function ${}_2F_1$ *is*

$${}_2F_1(\alpha,\beta;\gamma;z) := 1 + \sum_{m=1}^{\infty} \frac{\alpha(\alpha+1)\cdots(\alpha+m-1)\beta(\beta+1)\cdots(\beta+m-1)}{\gamma(\gamma+1)\cdots(\gamma+m-1)m!} z^m$$

As the nomenclature implies, the hypergeometric function is a solution to the hypergeometric equation. For ease of notation, following [11], we introduce the Pochhammer symbol:

Notation 2.4.2. *The* Pochhammer symbol $(\gamma)_m$ *is defined as*

$$(\gamma)_m = \begin{cases} 1 & m = 0 \\ \alpha(\alpha+1)\cdots(\alpha+m-1) & m \geqslant 1 \end{cases}$$

Then using this notation, the hypergeometric function is

$${}_2F_1(\alpha,\beta;\gamma;z) = \sum_{m=0}^{\infty} \frac{(\alpha)_m(\beta)_m}{(\gamma)_m m!} z^m.$$

Gauss later studied these and made several observations, including that the series is a polynomial if either $\alpha - 1$ or $\beta - 1$ is a negative integer, since then after some point all of the Pochhammer symbols will evaluate to zero; and for the same reason, it is undefined if $\gamma - 1$ is a negative integer. In all other cases, he showed that it is convergent for $z \in \mathbb{C}$ where $|z| < 1$. In fact it analytically continues to a multivalued function on the Riemann sphere, with singularities at $0, 1$ and ∞. Moreover, he introduced the notion of contiguous functions:

Definition 2.4.3. *Two hypergeometric functions are* contiguous *if at least one of their parameters differs by exactly one and the others remain the same. For example,* ${}_2F_1(\alpha,\beta;\gamma;z)$ *is* contiguous *to each of* ${}_2F_1(\alpha\pm1,\beta\pm1;\gamma\pm1;z)$.

Recall that the *monodromy* of a function is a representation consisting

of transformations encoding what happens to the function as it goes around its singularities. It turns out, by Corollary 2.6 and Proposition 2.7 of [5], that suppose two contiguous functions have irreducible monodromy; then their monodromy representations are equivalent.

Fuchs later began to study more general homogeneous linear ordinary differential equations. Specifically, he studied those of the form

$$\frac{d^n y}{dx^n} + p_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + p_{n-1} \frac{dy}{dx} + p_n y = 0,$$

where the p_i are single-valued meromorphic functions of x on the complex plane $\mathbb{C}$, or some simply connected region thereof. The p_i were moreover required to have finitely many singular points. Recall that a *regular singular point* is a singularity where the growth of solutions is bounded by some algebraic function. We restrict our interest here to equations of this type which have come to be known as (generalized) hypergeometric equations, which we will define shortly.

The naming thereof comes from the fact that its solutions are (related to) a clear generalization of the standard hypergeometric function, the generalized hypergeometric function introduced by Thomae. There are, of course, other generalizations of the hypergeometric function; see for example [9].

Definition 2.4.4. *The generalized hypergeometric function* ${}_nF_{n-1}$ *is*

$${}_nF_{n-1}(\alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots, \beta_{n-1}; z) = \sum_{m=0}^{\infty} \frac{(\alpha_1)_m (\alpha_2)_m \cdots (\alpha_n)_m}{(\beta_1)_m (\beta_2)_m \cdots (\beta_{n-1})_m m!} z^m.$$

Notation 2.4.5. *For notational convenience we will often shorten the notation to* ${}_nF_{n-1}(\alpha_i; \beta_k; z)$*. Similarly we will often denote two sets of hypergeometric parameters* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}, \{\beta_1, \beta_2, \ldots, \beta_{n-1}\}$ *simply as* $\{\alpha_i\}, \{\beta_k\}$*, where it is implicit that we have the indices* $i = 1, \ldots, n$ *and* $k = 1, \ldots, n-1$*. We denote the least common multiple of the denominators of* $\{\alpha_i\} \cup \{\beta_k\}$ *as* D*.*

Thomae discovered that these extend many of the properties of the standard hypergeometric function, ${}_2F_1$; for instance, they converge for $|z| < 1$, and contiguous functions, which are defined using the obvious extension of the

original definition, satisfy the condition that any $n+1$ contiguous functions have linear relations with rational coefficients. Similarly to the hypergeometric equation and for the same reasons, ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ is a polynomial when any $\alpha_i \in \mathbb{Z}_{<0}$, and is undefined if some $\beta_i \in \mathbb{Z}_{<0}$. Additionally, they satisfy an nth order linear ODE, the aforementioned hypergeometric equation:

Definition 2.4.6. *Let $\theta = z\frac{d}{dz}$, $n \geqslant 2$, and $p_1, \ldots, p_n \in \mathbb{C}(z)$, such that each $p_j = p_{j0} + p_{j1}(z-1)^{-1}$ for $p_{jk} \in \mathbb{C}$. Let P be the differential operator*

$$P := \theta^n + p_1\theta^{n-1} + \cdots + p_{n-1}\theta + p_n.$$

The equation $Px = 0$ is called a hypergeometric equation.

Multiplying both sides of a hypergeometric equation by $1-z$, we get an equation which we can write in the form $Du = 0$, where

$$\begin{aligned} D &= D(\alpha_1, \ldots, \alpha_n; \beta_1, \ldots, \beta_n) \\ &= (\theta + \beta_1 - 1)\cdots(\theta + \beta_n - 1) - z(\theta + \alpha_1)\cdots(\theta + \alpha_n) \end{aligned}$$

By abuse of notation, we will also call this a hypergeometric equation. The generalized hypergeometric function ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ satisfies the hypergeometric equation with β_n set to 1. (In fact when β_n does not equal 1 the solutions are hypergeometric equations shifted by powers of z; see [5].)

These generalized hypergeometric functions also show up in many places, for instance, as we will explain later, in the study of vector valued modular forms. They will be our main topic of study, although we will often use the standard hypergeometric function as a more manageable example.

Chapter 3

Unbounded Coefficients for Hypergeometric Functions

3.1 Preliminary Results

The results we prove were originally shown by Franc, Gannon and Mason for ${}_2F_1$ in [11]; the main work of this chapter is to generalize them to ${}_nF_{n-1}$. There is some overlap with earlier work by Dwork ([7]) and Christol ([6]).

As mentioned previously, we will restrict our attention to generalized hypergeometric functions with rational parameters. From here on we will call such things simply hypergeometric functions. Furthermore, we will assume that no $\alpha_k - \beta_j \in \mathbb{Z}$, since if this is the case then the monodromy representation is reducible by Proposition 2.7 of [5].

Notation 3.1.1. *Let ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ be a (generalized) hypergeometric function. We denote the coefficient of z^m by A_m. That is,*

$$A_m := \frac{(\alpha_1)_m(\alpha_2)_m\cdots(\alpha_n)_m}{(\beta_1)_m(\beta_2)_m\cdots(\beta_{n-1})_m m!}$$

Definition 3.1.2. *Let ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ be a hypergeometric function. ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ is said to have p-*adically unbounded coefficients *if $\inf_m(\nu_p(A_m)) = -\infty$. In this case, p is said to be an* unbounded prime for ${}_nF_{n-1}(\alpha_i;\beta_k;z)$.

The main question here is, when is it the case that ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ has

p-adically unbounded coefficients? This question has applications in vector valued modular forms, as will be explained later. We care only about when $\inf_m(\nu_p) = -\infty$ since the question of when $\sup_m \nu_p(A_m) = \infty$ is much easier, as will be shown (following Proposition 3.8 of [11]) in Proposition 3.1.4. First, however, we need some additional definitions.

Our main tool in analyzing the hypergeometric equation is the following theorem, which is in fact an easy corollary of Theorem 2.3.6.

Theorem 3.1.1. *Let $\{\alpha_i\}, \{\beta_k\}$ be rational hypergeometric parameters, and p a prime such that all of $\alpha_i - 1, \beta_j - 1$ are p-adic integers. Then in ${}_nF_{n-1}(\alpha_i; \beta_k; z)$, for the mth coefficient A_m, we have*

$$\nu_p(A_m) = \sum_{i=1}^{n} c_p(\alpha_i - 1, m) - \sum_{j=1}^{n-1} c_p(\beta_j - 1, m).$$

Proof. We have

$$\begin{aligned} A_m &= \frac{(\alpha_1)_m(\alpha_2)_m \cdots (\alpha_n)_m}{(\beta_1)_m(\beta_2)_m \cdots (\beta_{n-1})_m m!} \\ &= \left(\prod_{i=1}^{n} \frac{(\alpha_i)_m}{m!}\right)\left(\prod_{j=1}^{n-1} \frac{m!}{(\beta_j)_m}\right) \\ &= \frac{\prod_{i=1}^{n} \binom{\alpha_i - 1 + m}{m}}{\prod_{j=1}^{n-1} \binom{\beta_j - 1 + m}{m}}. \end{aligned}$$

Hence, by Theorem 2.3.6 and the properties of valuations, we have

$$\begin{aligned} \nu_p(A_m) &= \nu_p\left(\frac{\prod_{i=1}^{n} \binom{\alpha_i - 1 + m}{m}}{\prod_{j=1}^{n-1} \binom{\beta_j - 1 + m}{m}}\right) \\ &= \sum_{i=1}^{n} \nu_p\left[\binom{\alpha_i - 1 + m}{m}\right] - \sum_{j=1}^{n-1} \nu_p\left[\binom{\beta_j - 1 + m}{m}\right] \\ &= \sum_{i=1}^{n} c_p(\alpha_i - 1, m) - \sum_{j=1}^{n-1} c_p(\beta_j - 1, m), \end{aligned}$$

as required. □

As shown in Corollary 2.6 of [5], subject to our existing conditions, contiguous hypergeometric functions will have the same monodromy. In Lemma 4.1 of [11], it is shown that in the case of ${}_2F_1$ they also share the same unbounded primes. Specifically, Lemma 4.1 of [11] states the following:

Lemma 3.1.2. *Suppose we have two sets of rational hypergeometric parameters for ${}_2F_1$, $(a, b; c)$ and $(r, s; t)$, and suppose furthermore that*

1. *none of $a, b, c, a - c$, or $b - c$ is an integer;*

2. *$a - r, b - s$, and $c - t$ are all integers.*

Then a prime p is an unbounded prime for ${}_2F_1(a, b; c; z)$ if and only if it is an unbounded prime for ${}_2F_1(r, s; t; z)$.

Here we will generalize this, showing that it holds for all ${}_nF_{n-1}$, of which ${}_2F_1$ is a special case.

Theorem 3.1.3. *Let $\alpha_i, \beta_j \in \mathbb{Q} \setminus \mathbb{Z}, i = 1, \ldots, n, j = 1, \ldots, n-1$ be such that for all i, j, $\alpha_i - \beta_j \notin \mathbb{Z}$. Take any set of $k_i, l_j \in \mathbb{Z}$. Then p is an unbounded prime for ${}_nF_{n-1}(\alpha_i; \beta_j; z)$ iff the same is true for ${}_nF_{n-1}(\alpha_i + k_i; \beta_j + l_j; z)$.*

Proof. We'll show that if any of the following series have p-adically unbounded coefficients for some prime p, then so does $F := {}_nF_{n-1}(\alpha_i; \beta_k; z)$:

1. $G^1 := {}_nF_{n-1}(\alpha_1 + 1, \alpha_2 + 1, \ldots, \alpha_n + 1; \beta_1 + 1, \beta_2 + 1, \ldots, \beta_n + 1; z)$,

2. $G^2 := {}_nF_{n-1}(\alpha_1 + 1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots, \beta_n; z)$,

3. $G^3 := {}_nF_{n-1}(\alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1 - 1, \beta_2, \ldots, \beta_n; z)$, and

4. $G^4 := {}_nF_{n-1}(\alpha_1 - 1, \alpha_2, \ldots, \alpha_n; \beta_1, \beta_2, \ldots, \beta_n; z)$.

A more natural way to view the above results is that, for example, (1) says that if ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically unbounded coefficients, then so does $F^1 := {}_nF_{n-1}(\alpha_1 - 1, \alpha_2 - 1, \ldots, \alpha_n - 1; \beta_1 - 1, \beta_2 - 1, \ldots, \beta_n - 1; z)$. We use the equivalent statement above as it allows more convenient notation for the proof. It is also clear from the definition that we have ${}_nF_{n-1}(\alpha_i; \beta_k; z) = {}_nF_{n-1}(\alpha_{\sigma(i)}; \beta_{\tau(k)}; z)$ for any $\sigma \in S_n, \tau \in S_{n-1}$. Using this alternate statement

and combining (1), (3) and (4), along with the symmetry, we can show that if $G^5 := {}_nF_{n-1}(\alpha_1, \alpha_2, \ldots, \alpha_n; \beta_1 + 1, \beta_2, \ldots, \beta_n; z)$ has p-adically unbounded coefficients, so does ${}_nF_{n-1}(\alpha_i; \beta_k; z)$. Iterating these statements completes the proof of the original lemma.

From here on for any power series $\tilde{F}$ we denote its mth coefficient by $\tilde{F}_m$; that is, for instance, we define F_m by $F(z) = \sum_{i=0}^{\infty} F_m z^m$, and similarly each G^i_m by $G^k(z) = \sum_{i=1}^{\infty} G^i_m z^m$.

The first observation is that if $\frac{d^k}{dz^k}F$ has p-adically unbounded coefficients, then so does F. This is due to the fact that taking the derivative simply multiplies each coefficient by an integer and shifts its index, that is to say, $(\frac{d}{dz}(F))_{m-1} = mF_m$. Thus while it is possible that $\frac{d}{dz}F$ could have p-adically unbounded coefficients for *fewer* primes p, it could not possibly have p-adically unbounded coefficients for any additional primes.

The technique in general is to prove that the given function G^k is a polynomial in z and derivatives of F. This implies that if G^k has p-adically unbounded coefficients, then so does F or one of its derivatives, and thus by the previous argument F must have p-adically unbounded coefficients. (1), (2) and (3) are relatively straightforward and the relation can be given by general formulas; (4) requires a bit more work, and we provide only a proof of the existence of a formula for each n.

Fix a set of parameters $\{\alpha_i\}, \{\beta_k\}$.

For (1), note that

$$\begin{aligned}
\frac{d}{dz}F &= \sum_{m \geqslant 0} m \cdot \frac{\prod_{i=1}^{n}(\alpha_i)_m}{(m)! \prod_{j=1}^{n-1}(\beta_j)_m} z^{m-1} \\
&= \frac{\prod_{i=1}^{n} \alpha_i}{\prod_{j=1}^{n-1} \beta_j} \sum_{m \geqslant 1} \frac{\prod_{i=1}^{n}(\alpha_i + 1)_{m-1}}{(m-1)! \prod_{j=1}^{n-1}(\beta_j + 1)_{m-1}} z^{m-1} \\
&= \frac{\prod_{i=1}^{n} \alpha_i}{\prod_{j=1}^{n-1} \beta_j} \cdot G^1.
\end{aligned}$$

Thus any unbounded primes for G^1 must also be unbounded primes for $\frac{d}{dz}F$ and thus also for F.

For (2), we have

$$\begin{aligned} G^2 &= \sum_{m\geqslant 0} \frac{(\alpha+1)_m \prod_{i=2}^{n}(\alpha_i)_m}{(m)!\prod_{j=1}^{n-1}(\beta_j)_m} z^m \\ &= \frac{1}{\alpha_1}\sum_{m\geqslant 0} \frac{\prod_{i=1}^{n}(\alpha_i)_m}{(m)!\prod_{j=1}^{n-1}(\beta_j)_m} z^m \cdot (\alpha_1+m) \\ &= F + \alpha_1^{-1} z\frac{d}{dz}F. \end{aligned}$$

(3) is similar:

$$\begin{aligned} G^3 &= \sum_{m\geqslant 0} \frac{\prod_{i=1}^{n}(\alpha_i)_m}{(m)!(\beta-1)_m\prod_{j=2}^{n-1}(\beta_j)_m} z^m \\ &= \frac{1}{\beta_1-1}\sum_{m\geqslant 0} \frac{\prod_{i=1}^{n}(\alpha_i)_m}{(m)!\prod_{j=1}^{n-1}(\beta_j)_m} z^m \cdot (\beta_1-1+m) \\ &= F + (\beta_1-1)^{-1} z\frac{d}{dz}F. \end{aligned}$$

For case (4), we will involve the following series:

$$H := \sum_{m\geqslant 0} \frac{\prod_{i=1}^{n}(\alpha_i)_{m-1}}{(m)!\prod_{j=1}^{n-1}(\beta_j)_m} z^m = \sum_{m\geqslant 0} H_m z^m.$$

The motivation for this is that, coefficientwise, using this series gives the following equations:

$$(G^4)_m = (\alpha_1-1)\prod_{i=2}^{n}(m+\alpha_i-1)\cdot H_m \tag{3.1}$$

$$(z^k\frac{d^k}{dz^k}F)_m = m^{\underline{k}}\prod_{i=1}^{n}(m+\alpha_i-1)\cdot H_m \tag{3.2}$$

$$(z^{k+1}\frac{d^k}{dz^k}F)_m = m^{\underline{k+1}}\prod_{j=1}^{n-1}(m+\beta_j-1)\cdot H_m \tag{3.3}$$

where $k \in \mathbb{Z}_{\geqslant 0}$, F_m is the coefficient of z^m in F and

$$m^{\underline{k}} := \begin{cases} 1 & k = 0 \\ m(m-1)(m-2)\dots(m-k+1) & k > 0. \end{cases}$$

Thus, we can use H to relate the coefficients of these series to each other by polynomials in m, allowing us to construct systems of linear equations. Note that the polynomials above are all monic of degree $n-1$, $n+k$ and $n+k$ respectively, and that the only ones with a non-zero constant term are the one for G^4 and F (ie $k=0$). Allowing k to range from 0 to $n-1$ gives us $2n+1$ polynomials in m of degree at most $n+n-1 = 2n-1$. Comparing coefficients of these polynomials then gives us a system of $2n$ homogeneous linear equations in $2n+1$ variables, which guarantees infinite nontrivial solutions by basic linear algebra arguments. Each solution will be a linear relation between $G^4, F, z\frac{d}{dz}F, \dots z^k\frac{d^k}{dz^k}F, zF, \dots, z^{k+1}\frac{d^k}{dz^k}F$, with coefficients which are rational functions in the α_i, β_j.

It remains to show that these solutions do not all require the coefficient of $(G^4)_m$ to be zero. Since only the polynomials of $(G_4)_m$ and F_m have constant terms, this would also require that F_m have a coefficient of zero. Thus, if such a solution existed it would imply that zF could be written as a linear combination of $z^k\frac{d^k}{dz^k}F$ and $z^{k+1}\frac{d^k}{dz^k}F$ for $k > 0$.

Suppose this were true. This is equivalent to saying that F can be written as a linear combination of $z^{k-1}\frac{d^k}{dz^k}F$ and $z^k\frac{d^k}{dz^k}F$ for $1 \leqslant k \leqslant n-1$. Note that

$$(z^k\frac{d^k}{dz^k}F)_m = m^{\underline{k}}F_m$$

and

$$\begin{aligned}(z^{k-1}\frac{d^k}{dz^k}F)_m &= (m+1)^{\underline{k}}F_{m+1} \\ &= (m+1)^{\underline{k}}\frac{\prod_{i=1}^{n}(\alpha_i)_{m+1}}{(m+1)!\prod_{j=1}^{n-1}(\beta_j)_{m+1}} \\ &= (m)^{\underline{k-1}}\frac{\prod_{i=1}^{n}(\alpha_i+m)}{\prod_{j=1}^{n-1}(\beta_j+m)}F_m.\end{aligned}$$

But then this implies that there are some constants c_k, d_k such that for all m,

$$(1 - m\sum_{k=1}^{n-1} c_k(m-1)^{\underline{k-1}})F_m = \frac{\prod_{i=1}^{n}(\alpha_i + m)}{\prod_{j=1}^{n-1}(\beta_j + m)}(\sum_{k=1}^{n-1} d_k(m)^{\underline{k-1}})F_m.$$

Since the left hand side is a polynomial, the right hand side must be too. Then since the α_i are distinct from the β_j, the denominator must be cancelled by the polynomial $\sum_{k=1}^{n-1} d_k(m)^{\underline{k-1}}$. However $\sum_{k=1}^{n-1} d_k(m)^{\underline{k-1}}$ is a polynomial of degree $n-2$ in m and thus cannot possibly cancel the $n-1$ factors in the denominator, leading to a contradiction.

Thus, any solution of the system of linear equations will give us a formula for G^4 in terms of a finite linear combination of powers of z times derivatives of F. Thus, if G^4 has p-adically unbounded coefficients then so does F or one of its derivatives, and thus so does F. □

Example 1. In the case of ${}_2F_1$, we can use the following equations, originally proven by Gauss. We once again let $\theta = z\frac{d}{dz}$.

$$\begin{aligned}
{}_2F_1(\alpha+1,\beta;\gamma;z) &= \left(1 + \frac{1}{\alpha}\theta\right){}_2F_1(\alpha,\beta;\gamma;z),\\
{}_2F_1(\alpha-1,\beta;\gamma;z) &= \left((1-z) - \frac{(\alpha+\beta-\gamma)z}{\gamma-\alpha} + \frac{1-z}{\gamma-\alpha}\theta\right){}_2F_1(\alpha,\beta;\gamma;z),\\
{}_2F_1(\alpha,\beta;\gamma+1;z) &= \left(\frac{(\alpha+\beta-\gamma)\gamma}{(\gamma-\alpha)(\gamma-\beta)} + \frac{(1-z)\gamma}{(\gamma-\alpha)(\gamma-\beta)}\frac{d}{dz}\right){}_2F_1(\alpha,\beta;\gamma;z),\\
{}_2F_1(\alpha,\beta;\gamma-1;z) &= \left(1 + \frac{1}{\gamma-1}\theta\right){}_2F_1(\alpha,\beta;\gamma;z).
\end{aligned}$$

This lemma allows us to shift any parameters by any integer without affecting the results. Between this and the reducibility condition above, we can look simply at the following case.

Definition 3.1.3. *Let $\{\alpha_i\}, \{\beta_k\}$ be rational hypergeometric parameters. We call these parameters* admissible *if both of the following conditions are satisfied:*

1. *$0 < \alpha_i, \beta_j < 1$ for all i, j,*
2. *$\alpha_i \neq \beta_j$ for all i, j.*

Similarly, for many of our results we will require our primes to be good in the following sense. All but finitely many primes will be good for a given set of hypergeometric parameters.

Definition 3.1.4. *Given rational hypergeometric parameters* $\{\alpha_i\}, \{\beta_k\}$*, a prime* p *is called a* good *prime if* $\nu_p(\alpha_i - 1) = \nu_p(\beta_j - 1) = 0$*, that is, each* $\alpha_i - 1, \beta_j - 1$ *is in* $\mathbb{Z}_p^\times$*.*

Note that, by Lemma 2.3.1, for good primes, admissible rational hypergeometric parameters will have purely periodic expansions for each of $\alpha_i - 1, \beta_j - 1$, with period the multiplicative order of p modulo each of their denominators.

Definition 3.1.5. *By the* period *of the hypergeometric parameters, we mean the least common multiple of the multiplicative order of* p *modulo each of the denominators, that is to say the least common multiple of the periods of the different* $\alpha_i - 1, \beta_j - 1$*.*

We are now ready to show, as promised, that the question of when $\sup_m \nu_p(A_m) = \infty$ is fairly straightforward. We follow Proposition 3.8 of [11].

Proposition 3.1.4. *Let* $\{\alpha_i\}, \{\beta_k\}$ *be admissible hypergeometric parameters. Then* $\sup_m \nu_p(A_m) = \infty$ *for all good primes such that* $p > D$*. In particular, this is the case for all but finitely many primes.*

Proof. Given a good prime, we will construct a sequence of coefficients A_{m_k} such that $\nu_p(A_{m_k}) \xrightarrow{k\to\infty} \infty$.

Recall that for a good prime, $\nu_p(\alpha_i) = \nu_p(\beta_j) = 0$ for all i, j. Hence in particular, $\alpha_i^0, \beta_j^0 \neq 0$ for all i, j. Recall that $p^N - 1$ has the expansion

$$p^N - 1 = \overbrace{(p-1)(p-1)\dots(p-1)}^{\text{N times}}.$$

Hence when adding $(p^N - 1) + a$ for any p-adic integer a with non-zero zeroth digit, we get at least N carries. We get more than N carries if and only if the Nth digit of a is $p - 1$. Since for $p > D$, each of the $\{\alpha_i\}, \{\beta_k\}$ have no $p - 1$ digits, they will each have exactly N carries for $p^N - 1$.

Let $m_k = p^k - 1$. Then by Theorem 3.1.1,

$$\begin{aligned}\nu_p(A_{m_k}) &= \sum_{i=1}^{n} c_p(\alpha_i - 1, m_k) - \sum_{j=1}^{n-1} c_p(\beta_j - 1, m_k) \\ &= \sum_{i=1}^{n} k - \sum_{j=1}^{n-1} k \\ &= k \quad \xrightarrow{k\to\infty} \infty,\end{aligned}$$

as required.

□

3.2 Unbounded Coefficients

We can now proceed to proving generalizations of the theorems which appear in Section 4 of [11]. First we will introduce some definitions and notations which will simplify our discussion.

Due to Theorem 3.1.1, we will often want to talk about all the carries in a single column at once. That is, we will want to discuss the number of α_i and β_k for which, when adding m to $\alpha_i - 1$ or $\beta_k - 1$ respectively, there is a p-adic carry from the jth digit to the $j+1$st digit. For that reason we introduce the following terminology.

Definition 3.2.1. *Let p be a prime, $\{\alpha_i\}, \{\beta_k\}$ admissible hypergeometric parameters, m a positive integer. For any $\gamma \in \mathbb{Z}$, define*

$$c_p^j(\gamma, m) = \begin{cases} 1 & \text{if when adding } m + (\gamma) \text{ there is a carry from the } j\text{th column} \\ 0 & \text{otherwise.} \end{cases}$$

The (p-adic) net carries *(with respect to m and p) for the jth column of these parameters is defined to be $\sum_{i=1}^{n} c_p^j(\alpha_i - 1, m) - \sum_{i=1}^{n-1} c_p^j(\beta_k - 1, m)$. Similarly, we will sometimes refer to carries from $(\alpha_i - 1) + m$ as* positive carries *and carries from $(\beta_k - 1) + m$ as* negative carries.

Note that by Theorem 3.1.1, $v_p(A_m)$ is exactly the sum of the net carries

in each column with respect to m and p.

Definition 3.2.2. *Let $A = \{a_1, a_2, \ldots, a_n\}, B = \{b_1, b_2, \ldots, b_{n-1}\}$ be two sets of (not necessarily distinct) numbers. Suppose, rearranging if necessary, that $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n$ and $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_{n-1}$. If for each $i = 1, \ldots, n-1$, we have $a_i \geqslant b_i$, we say that A and B are* semi-interlaced downwards*; if $b_i \geqslant a_{i+1}$ we say that A and B are* semi-interlaced upwards*. If the inequalities are strict, we say that the sets are* strictly interlaced upwards *or* downwards *respectively. If the sets are not semi-interlaced, call any i for which the relevant inequality does not hold a* flip point *or* flip index*. (Whether this refers to an upwards or downwards semi-interlacing will usually be clear from context.)*

If A and B are strictly semi-interlaced both upwards and downwards, that is if

$$a_1 > b_1 > a_2 > b_2 > \cdots > a_{n-1} > b_{n-1} > a_n,$$

we say that the sets are interlaced*.*

Example 2. $\left\{\frac{1}{6}, \frac{2}{3}, \frac{5}{6}\right\}, \left\{\frac{1}{4}, \frac{3}{4}\right\}$ are interlaced, since

$$\frac{1}{6} < \frac{1}{4} < \frac{2}{3} < \frac{3}{4} < \frac{5}{6}.$$

$\left\{\frac{1}{4}, \frac{3}{4}, \frac{5}{6}\right\}, \left\{\frac{1}{6}, \frac{2}{3}\right\}$ are semi-interlaced downwards, but not upwards, since

$$\frac{1}{4} \not< \frac{1}{6} < \frac{3}{4} \not< \frac{2}{3} < \frac{5}{6}.$$

$\left\{\frac{1}{4}, \frac{1}{3}, \frac{3}{4}\right\}, \left\{\frac{1}{2}, \frac{5}{6}\right\}$ are semi-interlaced upwards, but not downwards, since

$$\frac{1}{4} < \frac{1}{2} \not< \frac{1}{3} < \frac{5}{6} \not< \frac{3}{4}.$$

This notion of being interlaced is essentially identical to the notion of being *interlaced on the unit circle* in [5], and the terminology was chosen to reflect this. Interlaced sets are used in Theorem 2.8 of [5], which we later use as part of our Theorem 3.4.4, as a finiteness condition for the so-called hypergeometric group, ie the monodromy group of a hypergeometric equation. In particular, a set of rational hypergeometric parameters are interlaced if and only if the

parameters of the corresponding hypergeometric group as defined in [5] are interlaced on the unit circle.

We also add the following lemma which will simplify some proofs.

Proposition 3.2.1. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters, p a good prime. Then ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically bounded coefficients iff $v_p(A_m) \geqslant 0$ for all m, that is to say all of its coefficients are p-adic integers.*

Proof. The converse is obvious. For the forward implication, we'll prove that if there is some m such that $v_p(A_m) < 0$ we can use this to construct a sequence of coefficients $(A_{m_i})_{i=0}^{\infty}$ such that $v_p(A_{m_i}) \xrightarrow{i\to\infty} -\infty$.

Let M be the period of the data. Since the p-adic expansion of any $m' \in \mathbb{Z}_{\geqslant 0}$ is finite, we can write the digits of the given m as $m = m(0)m(1)\dots m(s)$. Let $N = \left\lceil \frac{s}{M} \right\rceil + 1$. Since none of the parameters can be integers, in particular none of the $\alpha_i - 1$ or $\beta_k - 1$ is $-1 = \overline{(p-1)}$. Let $\gamma \in \{\alpha_i\} \cup \{\beta_k\}$. Then since each is purely periodic of period dividing M, in every collection of M consecutive digits $\gamma - 1$, there is at least one digit which is not $p-1$. When adding two p-adic numbers x and y, if there is a carry from the jth column to the $j+1$st column and $y(j+1, p) = 0$, then there will be a p-adic carry from the $j+1$st column to the $j+2$nd column if and only if $x(j+1, p) = p-1$. Thus any single digit of m can, on its own, cause at most M carries when m is added to any $\gamma - 1$. In particular, for any $m' \in \mathbb{Z}_{\geqslant 0}$, for any integer $t > s + M, c_p(\gamma - 1, p^t m' + m) = c_p(\gamma - 1, p^t m') + c_p(\gamma - 1, m)$. Since $\gamma - 1$ is purely periodic with period dividing M, we also have that for any $t' \in \mathbb{Z}_{\geqslant 0}$, $c_p(\gamma - 1, p^{t'M} m) = c_p(\gamma - 1, m)$. Combining these and repeating inductively, we get that for any $r \in Z_{\geqslant 0}$, $c_p(\gamma - 1, \sum_{j=0}^{r} p^{jNM} m) = (r+1)c_p(\gamma - 1, m)$.

Let $m^r := \sum_{j=0}^{r} p^{jNM} m$. Then by Theorem 3.1.1 and the above argument, we have

$$\begin{aligned} v_p(A_{m_r}) &= \sum_{i=1}^{n} c_p(\alpha_i - 1, m_r) - \sum_{k=1}^{n-1} c_p(\alpha_i - 1, m_r) \\ &= \sum_{i=1}^{n} (r+1) c_p(\alpha_i - 1, m) - \sum_{k=1}^{n-1} (r+1) c_p(\alpha_i - 1, m) \\ &= (r+1) v_p(A_m) \xrightarrow{r\to\infty} -\infty \end{aligned}$$

□

Remark 3.2.2. As in Remark 4.3 of [11], this proof implies that if there is some $m_0 < M$ such that $v_p(A_{m_0}) < 0$, then we can find some subsequence of $v_p(A_m)$ which diverges to $-\infty$ at least as fast as $\frac{1}{M}\log_p(m)$. In fact we will usually be using this proposition in this context.

In Theorem 4.2 of [11], we are given necessary and sufficient conditions for ${}_2F_1$ to have p-adically unbounded coefficients. Specifically, the following is proven:

Theorem 3.2.3. *Let (a,b;c) be admissible hypergeometric parameters, and let p be a good prime. The following are equivalent:*

1. *there exists an index j such that $\tau_j(c-1) > \tau_j(a-1)$ and $\tau_j(c-1) > \tau_j(b-1)$;*

2. *p is an unbounded prime for ${}_2F_1(a,b;c;z)$.*

Here, we generalize this to ${}_nF_{n-1}$. Unfortunately, the possibility of some complications are introduced by the possible interactions within columns when we are dealing with more parameters, specifically in the case where some digits may be $p-1$. Luckily, as shown earlier, we can solve this by excluding finitely many primes. We begin by proving the necessary and sufficient conditions separately before combining them into an equivalence theorem. Moreover, it is no longer sufficient to simply have a single parameter as the maximum; it is here that we begin to use the concept of being semi-interlaced.

We quickly recall that we have defined the jth column of the parameters to be the jth digits after subtracting one from each of the parameters.

Lemma 3.2.4. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters, p a good prime. If for every j the sets of j-truncations of each of these parameters minus one are semi-interlaced downwards, then ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ has p-adically bounded coefficients. If $p > D$ this is equivalent to every column of these parameters being semi-interlaced.*

Proof. Following the proof of Theorem 4.2 in [11], we will prove the following result: Choose any $m \in \mathbb{Z}_{\geqslant 0}$ and any $\alpha \in \{\alpha_i\}, \beta \in \{\beta_k\}$ with $\tau_{j+1}(\alpha - 1) \geqslant \tau_{j+1}(\beta - 1)$. If there is a p-adic carry from the jth column when evaluating $(\beta - 1) + m$, then there is also a p-adic carry from the jth column when evaluating $(\alpha - 1) + m$; that is, using our earlier notation, $c_p^j(\beta, m) \leqslant c_p^j(\alpha, m)$. This would imply that, for any j where the sets of j-truncations are semi-interlaced downwards, we have that the net number of carries in that column is non-negative for every m. If this is the case for every j, then by Theorem 3.1.1 we have that for every m, $v_p(A_m) \geqslant 0$, proving the lemma. The equivalence for $p > D$ follows directly from the fact that by Lemma 2.3.3, in this case $\beta_k^j \neq \alpha_i^j$ for all i, k.

We use induction on j. If $j = 0$, $\tau_1(\alpha - 1) = \alpha^0, \tau_1(\beta - 1) = \beta^0$. Then if $c_p^0(\beta - 1, m) = 1$, necessarily $m(0) \geqslant p - \beta^0 \geqslant p - \alpha^0$. Then $m(0) + \alpha^p \geqslant p$, so that there must be a p-adic carry, that is, $c_p^0(\alpha - 1, m) = 1$.

Now assuming we have proven the result for some $j - 1$, suppose that we have $\tau_j(\alpha - 1) \geqslant \tau_j(\beta - 1)$, and that there is a p-adic carry at the jth digit when evaluating $(\beta - 1) + m$. If $\beta^j < \alpha^j$, regardless of whether or not the carry depends on a carry from a previous carry, we must have

$$m(j) \geqslant p - (\beta^j + 1) \geqslant p - \alpha^j$$

so that

$$m(j) + \alpha^j \geqslant p,$$

that is, $c_p^j(\alpha - 1, m) = 1$. If, on the other hand, $\alpha^j = \beta^j$, then we can only have $c_p^j(\beta - 1, j) > c_p^j(\alpha - 1, j)$ if there is a carry from the $j - 1$st column for β but not for α. But $\tau_{j+1}(\alpha - 1) \geqslant \tau_{j+1}(\beta - 1)$ and $\alpha^j = \beta^j$ implies that $\tau_j(\alpha - 1) \geqslant \tau_j(\beta - 1)$, so by our induction hypothesis this is impossible. □

For convenience we introduce the following notation:

Notation 3.2.3. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters. For each j, we can find permutations $\sigma_j \in S_n$, $\tau_j \in S_{n-1}$ such that*

$$\alpha_{\sigma_j(1)}^j \geqslant \alpha_{\sigma_j(2)}^j \geqslant \cdots \geqslant \alpha_{\sigma_j(n)}^j$$

and

$$\beta^j_{\tau_j(1)} \geqslant \beta^j_{\tau_j(2)} \geqslant \cdots \geqslant \beta^j_{\tau_j(n-1)}.$$

Note that these permutations are not necessarily unique when $p \leqslant D$. We will call any such permutations orderings of the jth column *or* j-orderings.

Note that if k is a flip point for the jth column of one pair of orderings, it will be for any other pair of orderings as well, since different orderings can only interchange parameters with equal jth digits.

Lemma 3.2.5. *Let p be a good prime for admissible hypergeometric parameters $\{\alpha_i\}, \{\beta_k\}$. Suppose that there exists some column j of these parameters which is not semi-interlaced downwards, and furthermore, that for some j-orderings σ_j of $\{\alpha_i\}$ and τ_j of $\{\beta_k\}$ and for some flip point K, we have $\alpha^{j+1}_{\sigma_j(i)} < p-1$ for all $i < K$. Then ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically unbounded coefficients.*

Proof. For notational simplicity we assume without loss of generality that the j-ordering in the assumption coincides with the usual ordering, that is, that σ_j and τ_j are both the identity.

Since K is a flip point, by definition we have $\beta^j_K > \alpha^j_K$. Then in particular, $\beta^j_K > 0$. Then we can define $m = p^j(p - \beta^j_K)$, and this will be its p-adic expansion. Now since $\beta^j_k \geqslant \beta^j_K$ for all $k \leqslant K$, we have that for all such k, $c^j_p(m, \beta_k - 1) = 1$, so that $c_p(m, \beta_k - 1) \geqslant 1$ for all such k. For all $k < K$ we also have that, by assumption, $\alpha^{j+1}_k + 1 < p$; that is, a carry from the jth column cannot by itself cause a carry in the $(j+1)$th column. Then for these k, $c_p(m, \alpha_k - 1) \leqslant 1$. Meanwhile, for all $k \geqslant K$, $\beta^j_K > \alpha^j_K \geqslant \alpha^j_k$, so that $c_p(m, \alpha_k - 1) = 0$.

Combining these three facts and Theorem 3.1.1, we get that

$$\begin{aligned} v_p(A_m) &= \sum_{i=0}^{n} c_p(m, \alpha_i - 1) - \sum_{k=0}^{n-1} c_p(m, \beta_k - 1) \\ &\leqslant (K-1) - K \\ &< 0. \end{aligned}$$

Thus, by Lemma 3.2.1, ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically unbounded coefficients. □

This is clearly very close to being a true converse; the only impediment is the possibility of α_is with $p-1$ digits in inconvenient places. Luckily, by Lemma 2.3.4 for all but finitely many primes, this is not a possibility. Hence, we have the following theorem.

Theorem 3.2.6 (Generalization of Theorem 3.2.3)**.** *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters, D the least common multiple of their denominators, and let $p > D$. Then the following are equivalent:*

1. *For some index j, the j-truncations $\tau_j(\alpha_i - 1)$, $\tau_j(\beta_k - 1)$ are not semi-interlaced downwards.*

2. *For some index j, the jth column of the parameters is not semi-interlaced downwards.*

3. *For some index m, the mth coefficient A_m of ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ is not a p-adic integer.*

4. *${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically unbounded coefficients.*

Proof. (1) and (2) are equivalent by Lemma 2.3.3. We have already proven the equivalence of (3) and (4), and that (4) implies (2). By Lemma 2.3.4, satisfying (2) automatically satisfies Lemma 3.2.5, showing (2) implies (4). □

Example 3. We will examine whether 13 is an unbounded prime for ${}_3F_2$ with parameters $\left\{\frac{1}{6}, \frac{2}{3}, \frac{5}{6}\right\}, \left\{\frac{1}{4}, \frac{3}{4}\right\}$ and $\left\{\frac{1}{6}, \frac{3}{4}, \frac{5}{6}\right\}, \left\{\frac{1}{4}, \frac{2}{3}\right\}$.

Note that we have $D = 12$. Since $13 \equiv 1 \pmod{12}$, the period of the parameters will be 1. We can use lemma 2.3.2 to calculate the expansions of these parameters (with 1 subtracted) for the 13-adics. For example,

$$\begin{aligned} \frac{1}{6} - 1 \qquad &= -\frac{5}{6} \qquad &&= \overline{\left\lfloor \left\{ -13^0 \frac{1}{6} \right\} 13 \right\rfloor} \\ &= \overline{\left\lfloor \frac{5}{6} \cdot 13 \right\rfloor} \qquad &&= \overline{\left\lfloor \frac{65}{5} \right\rfloor} \\ &= \overline{(10)}. \end{aligned}$$

Similarly, $\frac{2}{3} - 1 = \overline{4}$; $\frac{5}{6} - 1 = \overline{2}$; $\frac{1}{4} - 1 = \overline{9}$; $\frac{3}{4} - 1 = \overline{3}$.

We have only a single repeated column, so it is sufficient to check whether this column is semi-interlaced downwards for each of the sets of parameters. In the first case, they are, since $10 > 9$ (corresponding to $\frac{1}{6}$ and $\frac{1}{4}$) and $4 > 3$ (corresponding to $\frac{2}{3}$ and $\frac{3}{4}$). Hence, we must have that ${}_3F_2(\frac{1}{6}, \frac{2}{3}, \frac{5}{6}; \frac{1}{4}, \frac{3}{4}; z)$ has 13-adically bounded coefficients. Indeed, it is easy to see that any digit added to this column will result in a net non-negative number of carries.

In the second case, the column is *not* semi-interlaced, since although $10 > 9$ (corresponding to $\frac{1}{6}$ and $\frac{1}{4}$), it is not the case that $3 > 4$ (corresponding to $\frac{3}{4}$ and $\frac{2}{3}$). Hence, ${}_3F_2(\frac{1}{6}, \frac{3}{4}, \frac{5}{6}; \frac{1}{4}, \frac{2}{3}; z)$ will have 13-adically unbounded coefficients. Indeed, by adding 9 to each of the parameters, we get a net of -1 carries, since it will cause a carry for $\frac{1}{6}, \frac{1}{4}$, and $\frac{2}{3}$ but not $\frac{3}{4}$ or $\frac{5}{6}$. Either repeating this for every column, or using Proposition 3.2.1, we see that we must have 13-adically unbounded coefficients.

3.3 Densities and Structure

The previous theorem gives a finite algorithm, for admissible hypergeometric parameters $\{\alpha_i\}, \{\beta_k\}$, to check whether ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ will have p-adically unbounded coefficients for any prime $p > D$; by periodicity, we can simply check the first M columns, where M is the period of the data. However checking each prime individually would still be an impossible task. In [11], this task is made somewhat more manageable by their Proposition 4.10, which states the following for the case of ${}_2F_1$.

Definition 3.3.1. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters. We define $S(\alpha_i; \beta_k)$ to be the set of primes for which ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically unbounded coefficients.*

Proposition 3.3.1. *Suppose $\{a, b\}, \{c\}$ are admissible hypergeometric parameters, and let D as usual be the lowest common multiple of their denominators. Let $p > D$ be a good prime such that $p \in S(a, b; c)$. Then for all good primes $q \geqslant p$ which are in the same congruence class modulo D, that is, such that $p \equiv q \pmod{D}$, we must also have $q \in S(a, b; c)$. In particular,*

$S(a,b;c)$ has a Dirichlet density of the form $\frac{\alpha}{\phi(D)}$, where α is an integer such that $0 \leqslant \alpha \leqslant \phi(D)$, and $\phi(D)$ is the Euler totient function.

Recall that the Euler totient function $\phi(D)$ is defined as the number of integers less than D which are relatively prime to D, ie the number of integers $1 \leqslant k \leqslant D$ such that $\gcd(k, D) = 1$.

We strengthen this here to show that it is in fact sufficient to check a single prime in each congruence class modulo D to find the answer for all primes larger than D. This, of course, also generalizes the result about the Dirichlet density of the set of unbounded primes.

Proposition 3.3.2. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters, $p > D$ a good prime. If $p \in S(\alpha_i; \beta_k)$, then all primes $q > D$ such that $p \equiv q \pmod{D}$ are also in $S(\alpha_i; \beta_k)$. That is, either all primes greater than D in a given congruence class modulo D are elements of $S(\alpha_i; \beta_k)$, or none of them are. In particular, $S(\alpha_i; \beta_k)$ has a Dirichlet density of the form $\frac{N}{\phi(D)}$ for an integer N satsifying $0 \leqslant N \leqslant \phi(D)$.*

Proof. Following the proof of Proposition 4.9 in [11], we prove that for any $a, b \in \{\alpha_i\} \cup \{\beta_k\}$ and any primes $p, q > D$ such that $p \equiv q \pmod{D}$ and $q > p$, if $a^j(p) < b^j(p)$ then $a^j(q) < b^j(q)$, that is, strict inequalities amongst digits in the same column are preserved. It is sufficient to prove this for $0 \leqslant j < M$, where M is the period of the data.

By Lemma 2.3.2, $a_j(p) = \lfloor \{-p^{M-1-j}a\}p \rfloor$ and $b_j(p) = \lfloor \{-p^{M-1-j}b\}p \rfloor$.

Since $q > p$, then we have $q = p + tD$ for some positive integer t. Thus we have

$$\begin{aligned}
a_j(p) < b_j(p) &\iff \lfloor \{-p^{M-1-j}a\}p \rfloor < \lfloor \{-p^{M-1-j}b\}p \rfloor \\
&\implies \{-p^{M-1-j}a\} < \{-p^{M-1-j}b\} \\
&\implies \lfloor \{-p^{M-1-j}a\}p \rfloor + \{-p^{M-1-j}a\}tD \\
&\qquad < \lfloor \{-p^{M-1-j}ab\}p \rfloor + \{-p^{M-1-j}b\}tD \\
&\iff \lfloor \{-q^{M-1-j}a\}q \rfloor < \lfloor \{-q^{M-1-j}b\}q \rfloor \\
&\iff a_j(q) < b_j(q).
\end{aligned}$$

Suppose $p \in S(\alpha_i; \beta_k)$. By Lemma 3.2.4, there is some column j which is not semi-interlaced for p. Since $p > D$, this can be stated purely in terms of strict inequalities by Lemma 2.3.3. By the above result, strict inequalities will be the same for q, so that the jth column will also not be semi-interlaced for q. By Theorem 3.2.6, this implies that $q \in S(\alpha_i; \beta_k)$.

Now suppose that $p \notin S(\alpha_i; \beta_k)$. By Theorem 3.2.6, all of the columns of the parameters must be semi-interlaced with respect to p. Since $p > D$, this can be stated entirely in term of strict inequalities by Lemma 2.3.3, which as shown above are preserved for $q > p$, $p \equiv q \pmod{D}$. Thus all the columns of the parameters must also be semi-interlaced for q, which again by Theorem 3.2.6 implies that $q \notin S(\alpha_i; \beta_k)$. Thus, $p \in S(\alpha_i; \beta_k)$ iff $q \in S(\alpha_i; \beta_k)$. □

Corollary 3.3.3. *Suppose $p \notin S(\alpha_i; \beta_k)$ and $p > D$. Then for all $q > D$ such that $q \equiv p^m \pmod{D}$, $q \notin S(\alpha_i; \beta_k)$. In particular, the congruence classes of primes in $(\mathbb{Z}/D\mathbb{Z})^\times$ for which ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ has p-adically bounded coefficients form a union of cyclic subgroups of $(\mathbb{Z}/D\mathbb{Z})^\times$.*

Proof. Let $p, q > D$, $p \notin S(\alpha_i; \beta_k)$ and $q \equiv p^m \pmod{d}$. By Theorem 3.2.6, with respect to p every column of the parameters $\{\alpha_i\}, \{\beta_k\}$ are semi-interlaced; we want to show that every column is also semi-interlaced with respect to q.

Take any index j. Then for each $\gamma \in \{\alpha_i\} \cup \{\beta_k\}$, by Lemma 2.3.2, $\gamma_j(q) = \lfloor \{-q^{M-1-j}\gamma\} q \rfloor$, where as usual M is the period of the data. Now since $q \equiv p^m \pmod{D}$, we have that $q^{M-1-j} \equiv (p^m)^{M-1-j} = p^{mM-m-jm} \pmod{D}$. Furthermore, since M is the order of p in $\mathbb{Z}/D\mathbb{Z}$, there exists some k such that $p^{mM-m-jm} \equiv p^{M-1-k}$. We will show that any (strict) inequality which holds in the kth column of the parameters with respect to p holds in the jth column of the parameters with respect to p; then since the kth column with respect to p is semi-interlaced and, by Lemma 2.3.3 all the inequalities involved are strict, the jth column with respect to q must be semi-interlaced as well.

Suppose $\alpha, \beta \in \{\alpha_i\}, \{\beta_k\}$, and $\alpha_k(p) < \beta_k(p)$. Then by Lemma 2.3.2, this is equivalent to $\lfloor \{-p^{M-1-k}\alpha\} p \rfloor < \lfloor \{-p^{M-1-k}\beta\} p \rfloor$. Then by our assump-

tions, we have

$$\begin{aligned} & \{-q^{M-1-j}\alpha\} = \{-p^{M-1-k}\alpha\} < \{-p^{M-1-k}\beta\} = \{-q^{M-1-j}\beta\} \\ \Rightarrow\ & \lfloor\{-q^{M-1-j}\alpha\}q\rfloor \leqslant \lfloor\{-q^{M-1-j}\beta\}q\rfloor \\ \Rightarrow\ & \alpha_j(q) \leqslant \beta_j(q). \end{aligned}$$

Since by Lemma 2.3.3 we know that the inequality must be strict, we have $\alpha_j(q) \leqslant \beta_j(q)$, that is, the inequality is preserved. □

It follows from this corollary that if there is any congruence class of primes for which ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ is bounded, then the congruence class of $p \equiv 1 \pmod{D}$ must also be bounded. By Proposition 3.3.2, this means that if $S(\alpha_i;\beta_k)$ contains any prime $p \equiv 1 \pmod{D}$, then ${}_nF_{n-1}(\alpha_i;\beta_k;z)$ can have at most finitely many primes for for which its coefficients are p-adically bounded. This will form part of our more general characterization of when $S(\alpha_i;\beta_k)$ contains all but finitely many primes.

3.4 The Cases of Finitely Many Bounded or Unbounded Primes

In Theorem 4.14 of [11], the following is proven for the case of ${}_2F_1$.

Theorem 3.4.1. *Let $\{a,b\},\{c\}$ be admissible hypergeometric parameters whose denominators have least common multiple D. The following are equivalent:*

1. *$c < a$ and $c < b$;*
2. *$S(a,b;c)$ contains all but finitely many primes;*
3. *$S(a,b;c)$ contains infinitely many primes p such that $p \equiv 1 \pmod{D}$.*

We generalize this as follows for ${}_nF_{n-1}$. Once again it is no longer sufficient for a single parameter to be the smallest, and we must instead bring in the concept of being semi-interlaced. We have also strengthened the final condition.

Theorem 3.4.2. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters, D the least common multiple of their denominators. Then the following are equivalent:*

1. *The parameters are not semi-interlaced upwards;*
2. *$S(\alpha_i; \beta_k)$ contains all but finitely many primes;*
3. *There exists some $p \equiv 1 \pmod{D}$ such that $p \in S(\alpha_i; \beta_k)$*
4. *Every prime $p \equiv 1 \pmod{D}$ is in $S(\alpha_i; \beta_k)$.*

Proof. We mostly follow the proof of Theorem 4.12 in [11]; the same technique works here.

We first note that (3) and (4) are equivalent by the proof of Proposition 3.3.2, since all primes congruent to 1 modulo D must be larger than D.

We begin by showing that (1) implies (2). Suppose that the parameters are not semi-interlaced. Let $A := \{\alpha_i\} \cup \{\beta_k\}$. Take any prime $p > \max\left\{\frac{1}{\gamma_1-\gamma_2} | \gamma_1, \gamma_2 \in A\right\} \cup \{D\}$, and let M the period of the data. By Lemma 2.3.2, for each $\gamma_1, \gamma_2 \in A$ we have $\gamma_1^{M-1} = \lfloor\{-\gamma_1\}p\rfloor = p + \lfloor -p\gamma_1\rfloor$ and, similarly, $\gamma_2^{M-1} = p + \lfloor -p\gamma_2\rfloor$. Suppose that $\gamma_1 > \gamma_2$. By assumption, $p > \max\left\{\frac{1}{\gamma_1-\gamma_2} | \gamma_1, \gamma_2 \in A\right\}$, or equivalently, $-\gamma_1 + \frac{1}{p} < -\gamma_2$. Thus, $\lfloor -p\gamma_1\rfloor < \lfloor -p\gamma_2\rfloor$, that is, $\gamma_1^{M-1} < \gamma_2^{M-1}$.

Stated differently, this says that any strict inequality which holds for the parameters holds with the inequalities reversed for the $M-1$st column of those parameters. Then if the parameters are not semi-interlaced upwards, their $M-1$st column cannot be semi-interlaced downwards. By Lemma 3.2.6, this implies that coefficients of ${}_nF_{n-1}(\alpha_i; \beta_k; z)$ are p-adically unbounded for any sufficiently large p. Using Proposition 3.3.2 and Dirichlet's theorem on primes in arithmetic sequences, we now have that all good primes $p > D$ must be in $S(\alpha_i; \beta_k)$; in particular, $S(\alpha_i; \beta_k)$ contains all but finitely many primes.

That (2) implies (3) is clear. It remains to show that (3) implies (1). By Lemma 2.3.1, in the case where $p \equiv 1 \pmod{D}$ we have that the period of the data M is 1, and clearly $p > D$ (else $p = 1$). Then by the above argument, any inequality which holds for the parameters holds with the inequalities reversed

for the 0th, and thus for every, column with respect to p. In particular, if the parameters were semi-interlaced upwards, then every column of the parameters would be semi-interlaced downwards with respect to p. By Lemma 3.2.5, this would imply that $p \notin S(\alpha_i; \beta_k)$, contradicting our assumption. $\square$

Example 4. We previously showed that $13 \equiv 1 \pmod{12}$ is in $S(\frac{1}{6}, \frac{3}{4}, \frac{5}{6}; \frac{1}{4}, \frac{2}{3})$. Indeed, these parameters are not semi-interlaced downwards; and moreover, by checking the first prime in each of the other congruence groups modulo D as we did for 13, one can confirm that it has at most finitely many bounded primes.

The following theorem was proven for the case of ${}_2F_1$ in Theorem 4.12 of [11].

Theorem 3.4.3. *Let $\{a, b\}, \{c\}$ be admissible hypergeometric parameters whose denominators have least common multiple D. The following are equivalent:*

1. *the monodromy group of the corresponding hypergeometric differential equation is finite;*

2. *the set $S(a, b; c)$ is finite;*

3. *for every integer u coprime to D, the fractional parts are interlaced, ie $\{uc\}$ lies between $\{ua\}$ and $\{ub\}$.*

Recall that the monodromy representation, roughly speaking, is the data of a group and a transformation which describe how its solutions transform as one travels around its branch points. For instance, take the equation $\frac{df}{dz} = rf$; its monodromy is $e^{2\pi ir}$ for the counter-clockwise circle around 0. Monodromy reps of hypergeometric equations are similar, but with 3 branch points instead of 2.

We generalize this for ${}_nF_{n-1}$ as follows. Note that we have added an additional equivalent condition.

Theorem 3.4.4. *Let $\{\alpha_i\}, \{\beta_k\}$ be admissible hypergeometric parameters, D the least common multiple of the denominators. Then the following are equivalent:*

1. *the monodromy group of the corresponding hypergeometric differential equation is finite;*

2. *The set $S(\alpha_i;\beta_k)$ is finite;*

3. *The set $S(\alpha_i;\beta_k)$ contains no good primes $p > D$;*

4. *For every integer u coprime to D, the fractional parts $\{u\alpha_i\}$ and $\{u\beta_k\}$ are interlaced.*

Proof. As discussed in the introduction of [11], a well-known theorem of Eisenstein implies for that any function which is a solution of an ordinary differential equation, if it has a finite monodromy group and rational Taylor coefficients, then it has p-adically bounded coefficients for all but finitely many primes p. In particular, this holds for hypergeometric equations, that is, (1) implies (2). The equivalence of (2) and (3) follows immediately from Proposition 3.3.2. The equivalence of (1) and (4) follows from Theorem 4.8 in [5]. Thus it is sufficient to show (3) implies (4).

By Theorem 3.2.6 and Lemma 2.3.3, every column of the parameters is strictly semi-interlaced downwards for all primes $p > D$. Then by Lemma 2.3.2, the sets $\{\lfloor\{-p^{M-1-j}\alpha_i\}p\rfloor \mid 1 \leqslant i \leqslant n\}$ and $\{\lfloor\{-p^{M-1-j}\beta_k\}p\rfloor \mid 1 \leqslant k \leqslant n-1\}$ are strictly semi-interlaced downwards for all j and for all but finitely many primes p. This in turn implies that the sets $\{\{-p^{M-1-j}\alpha_i\} \mid 1 \leqslant i \leqslant n\}$ and $\{\{-p^{M-1-j}\beta_k\} \mid 1 \leqslant k \leqslant n-1\}$ are strictly semi-interlaced downwards.

Note that for any $\frac{x}{D} \in \mathbb{Q}$, $x, u, v \in \mathbb{Z}$, if $v \equiv u \pmod{D}$, then $\{v\frac{x}{D}\} = \{u\frac{x}{D}\}$. By Dirichlet's theorem on primes in arithmetic progressions, for any u coprime to D, there exist infinitely many primes congruent to $-u$ modulo D, and thus such that $-p \equiv u \pmod{D}$; thus by varying p and j, using the result above we get that for any u coprime to D, the sets $\{\{u\alpha_i\} \mid 1 \leqslant i \leqslant n\}$ and $\{\{u\beta_k\} \mid 1 \leqslant k \leqslant n-1\}$ are strictly semi-interlaced downwards.

We now prove that $\{u\alpha_i\}$ and $\{u\beta_k\}$ are strictly semi-interlaced upwards if $\{-u\alpha_i\}$ and $\{-u\beta_k\}$ are strictly semi-interlaced downwards. Since $-u$ will be coprime to D whenever u is, by the above argument know that $\{-u\alpha_i\}$ and $\{-u\beta_k\}$ are indeed strictly semi-interlaced downwards, so this will complete the argument.

Let $\frac{x}{D}, \frac{y}{D} \in \mathbb{Q}$, where $x, y \in \mathbb{Z}$ but the fractions are not necessarily in lowest terms. By the definitions of strictly semi-interlaced upwards and downwards, it will in fact be enough to show that if $\{-u\frac{x}{D}\} < \{-u\frac{y}{D}\}$ then $\{u\frac{x}{D}\} > \{u\frac{y}{D}\}$, that is, that taking the negative still reverses inequalities when done within the fractional part operator.

Suppose $\{-u\frac{x}{D}\} < \{-u\frac{y}{D}\}$. Let $-ux = a_x D + r_x$, $-uy = a_y D + r_y$, where r_x, r_y are the usual remainders satisfying $0 \leqslant r_x, r_y < D$. Then $\{-u\frac{x}{D}\} = \frac{r_x}{D}$, $\{-u\frac{y}{D}\} = \frac{r_y}{D}$, so that $\frac{r_x}{D} < \frac{r_y}{D}$. We also have $ux = -(a_x D + r_x) = -(a_x + 1)D + (D - r_x)$, and similarly $uy = -(a_y + 1)D + (D - r_y)$, so that $\{u\frac{x}{D}\} = \frac{D - r_x}{D}$ and $\{u\frac{y}{D}\} = \frac{D - r_y}{D}$. But $\frac{r_x}{D} < \frac{r_y}{D} \Longrightarrow \frac{D - r_x}{D} > \frac{D - r_y}{D}$, proving our claim.

Thus for every u coprime to D, $\{u\alpha_i\}$ and $\{u\beta_k\}$ are strictly semi-interlaced both upwards and downwards, that is, they are interlaced. □

Example 5. It is known (see [5]) that the monodromy group corresponding to ${}_3F_2(\frac{1}{6}, \frac{2}{3}, \frac{5}{6}; \frac{1}{4}, \frac{3}{4}; z)$ has finite monodromy. We have already checked that all primes larger than 12 congruent to 1 (mod 12) are bounded; one can similarly check that the same will be true for all other primes larger than 12.

Chapter 4

Modular Forms and the Atkin-Swinnerton-Dyer Conjecture

4.1 Modular Forms and the Modular Group

Although some might find the question of unbounded primes for the generalized hypergeometric function interesting of its own volition, to explain our interest in it we must first introduce a new concept, that of modular forms, and especially vector valued modular forms (or automorphic forms). We give here only a brief introduction; for a more thorough introduction to modular forms, see for example [26] or [23]. More information about vector valued modular/automorphic forms can be found, for example, in [4], [15] or [3].

Modular forms show up throughout mathematics, from number theory where they originated in connection to elliptic curves, to areas such as complex analysis, algebraic topology, and vertex operator algebras. For example, Zhu famously proved that the character of every "nice enough" vertex operator algebra is a vector valued modular form (see [29]), although he did not use that terminology.

To define modular forms in their original and simplest incarnation, we first need to define the modular group and its action on the upper half plane.

Definition 4.1.1. *The* upper half plane $\mathbb{H}$ *is the set of complex numbers* τ *with imaginary part* $Im(\tau) > 0$. *We define the action of* $SL_2(\mathbb{R})$, *the group of 2 by 2 matrices* $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ *with real coefficients and determinant 1, on* $\mathbb{H}$, *and indeed on* $\mathbb{C} \cup \{\infty\}$, *as follows: let* $\tau \in \mathbb{C} \cup \{\infty\}$, $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{R})$. *Then*

$$\gamma \cdot \tau := \frac{a\tau + b}{c\tau + d},$$

where if $\tau = \infty$ *this is defined by taking the limit as* $\tau \to \infty$. *This action is called the* Möbius transformation.

Note that $\mathbb{H}$ is stable under this action, that is, this action maps $\mathbb{H}$ to itself. Moreover, note that

$$\begin{aligned} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \tau &= \frac{-\tau + 0}{0 - 1} \\ &= \tau. \end{aligned}$$

Hence, it actually makes more sense to look at $PSL_2(\mathbb{R}) := SL_2(\mathbb{R})/\{\pm I\}$. The modular group is the subgroup of this whose matrices have entries in $\mathbb{Z}$. Formally, we have the following.

Definition 4.1.2. *The* modular group *is defined as* $\Gamma(1) := SL_2(\mathbb{Z})/\{\pm I\}$.

We will usually refer to elements of $\Gamma(1)$ by a single representative of its pre-image in $SL_2(\mathbb{Z})$, with it implicit that by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we mean $\pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

It is well known (see for example [26]) that the modular group is generated by two elements, commonly referred to as S and T:

$$\begin{aligned} S &:= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ T &:= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

These elements transform the upper half plane as follows: For any $\tau \in \mathbb{H}$

$$S \cdot \tau = -\frac{1}{\tau}$$
$$T \cdot \tau = \tau + 1$$

They also satisfy the following relations $S^2 = (ST)^3 = 1$; for this reason we sometimes refer to the element ST as U.

Unlike the group $SL_2(\mathbb{R})$, the modular group does not act transitively on the upper half plane. Instead we have what is called the *fundamental domain*. This is a concept which we will later extend to Fuchsian groups in general.

Definition 4.1.3. *A* fundamental domain *of* $\Gamma = \Gamma(1)$ *is an open connected set (that is, a domain)* $D \subset \mathbb{H}$ *such that*

1. *For every* $\tau \in \mathbb{H}$*, there exists some* $\gamma \in \Gamma$ *and some* $\tau' \in \overline{D}$*, that is in the closure of* D*, such that* $\gamma \cdot \tau' = \tau$*.*

2. *No two elements of* D *are in the same* Γ*-orbit, that is, for any* $\tau, \tau' \in D$*,* $\gamma \cdot \tau \neq \tau'$ *for all* $\gamma \in \Gamma$*.*

Note that some sources use the term fundamental domain to refer to what we call the closure of the fundamental domain; our definition here follows that of [3].

The usual fundamental domain for $\Gamma(1)$ is

$$D = \left\{ \tau \in \mathbb{H} \mid \ |\tau| > 1, |Re(\tau)| < \frac{1}{2} \right\};$$

its closure is

$$\overline{D} = \left\{ \tau \in \mathbb{H} \mid \ |\tau| \geqslant 1, |Re(\tau)| \leqslant \frac{1}{2} \right\}.$$

We are now ready to define modular functions and forms.

Definition 4.1.4. *Let* f *be a meromorphic function on* $\mathbb{H}$*.* f *is* weakly modular of weight k*, where for our purposes* $k \in 2\mathbb{Z}$*, if it satisfies*

$$f(\tau) = (c\tau + d)^{-k} f(\gamma \cdot \tau), \quad \textit{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1).$$

Note that, since $\Gamma(1)$ is generated by S and T, it is sufficient to check the relations for S and T. That is to say:

Proposition 4.1.1. *Let f be a meromorphic function on $\mathbb{H}$. Then f is weakly modular of weight k for $\Gamma(1)$ if and only if it satisfies*

$$f(\tau+1) = f(\tau)$$
$$f\left(\frac{-1}{\tau}\right) = \tau^k f(z).$$

The first implies that, using the Fourier series, we can write $f(\tau)$ as a function of $q = e^{2\pi i\tau}$, which will be meromorphic in the disk $0 < |q| < 1$. By abuse of notation, we will also call this function $f(q)$ or even $f(\tau)$. This change of variables maps the real line to the boundary of the unit circle, and sends $i\infty$ to zero. If this function is holomorphic in the punctured disc $0 < |q| < r$ for some $r \leqslant 1$ and has a pole at $q = 0$, we call $f(q)$ *meromorphic at infinity.* If $q = 0$ is at worst a removable singularity, then we call it *holomorphic at infinity.*

Definition 4.1.5. *Let f be a weakly modular form of weight k. If additionally it is meromorphic at infinity, f is called a* modular form of weight k. *A modular form of weight* 0 *is called a* modular function.

If a modular form f is holomorphic everywhere except infinity, it is called a weakly holomorphic modular form*; if it is holomorphic everywhere, including at infinity, then it is called a* holomorphic modular form.

Note that these different possibilities are reflected in the Laurent expansions of the function around the origin: namely, a modular form will have a Laurent expansion of the form

$$f(\tau) = \sum_{n=N}^{\infty} a_n q^n;$$

for a holomorphic modular form we have $N \geqslant 0$.

Some common, important examples of modular forms are the Dedekind eta function (raised to the 24th power) and the Eisenstein series. The 24th power

of the Dedekind eta function is

$$\eta^{24}(\tau) := q\prod_{n=1}^{\infty}(1-q^n)^{24},$$

which converges for $|q| < 1$. It is well known that the 24th power of the Dedekind eta function is a holomorphic modular form of weight 12. Additionally, the expansion will give a power series with integral coefficients, and clearly the coefficient of q^0 is zero. Moreover, the eta function has no zeroes in $\mathbb{H}$.

The Eisenstein series are defined as

$$G_{2k}(\tau) = \sum_{(n,m)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m\tau+n)^{2k}}.$$

It is well known that these are holomorphic modular forms of weight $2k$ for all $k > 1$. Especially important are G_4 and G_6, which are often rescaled into E_4 and E_6:

$$E_4(\tau) = 1 + 240q + 2160q^2 + 6720q^3 + \dots$$
$$E_6(\tau) = 1 - 504q - 16632q^2 - 122976q^3 + \dots.$$

As explained in more detail in Section 4.4, these freely generate the space of holomorphic modular forms for $\Gamma(1)$. (For details and a proof of this, see for example page 88-89 of [26].)

(As an aside, for an appropriate definition of p-adic modular forms, G_2 is in fact a p-adic modular form; however this is beyond the scope of this thesis. For more information, see [27]. G_2 is also quasi-modular, and enters in to the modular derivative, which we will discuss later.)

4.2 Fuchsian Groups and $\Gamma(2)$

None of the above definitions are necessarily restricted to the modular group, although this is where they originated. In fact they are easily generalized to a class of groups called Fuchsian groups. We provide here a brief introduction

to Fuchsian groups and the definitions of modular forms for them; however this is largely a digression, and our main interest will be the group $\Gamma(2)$.

Definition 4.2.1. *Let Γ be a subgroup of $SL_2(\mathbb{R})$. We call Γ a* discrete *or* Fuchsian group *if, taking the infimum over all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma, \gamma \neq I$, we have*

$$\inf\{(a-1)^2 + b^2 + c^2 + (d-1)^2\} > 0.$$

For a more in-depth discussion of Fuchsian groups, see for example [20].

The modular group is, of course, an example of such a group (or at least its pre-image $SL_2(\mathbb{Z})$ is). So are any subgroups thereof, such as the commonly referenced congruence subgroups. We usually identify Fuchsian groups with their images in $PSL_2(\mathbb{R})$, and from this point forward will deal only with their images in $PSL_2(\mathbb{R})$. While there is no need for Fuchsian groups to be subgroups of the modular group, all those we deal with will be.

Definition 4.2.2. *We define*

$$\Gamma(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) | a \equiv d \equiv 1 \pmod{N}, b \equiv c \equiv 0 \pmod{N} \right\}.$$

Let Γ be a subgroup of $SL_2(\mathbb{Z})$. It is called a congruence subgroup *if it contains some $\Gamma(N)$.*

Especially important to us is the group

$$\Gamma(2) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) | a \equiv d \equiv 1 \pmod{2}, b \equiv c \equiv 0 \pmod{2} \right\}.$$

It is freely generated by the matrices

$$A = T^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$$B = ST^2S^{-1} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$$

.

Some important examples of congruence subgroups, other than the $\Gamma(N)$ themselves, are

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) | c \equiv 0 \pmod{N} \right\}$$
$$\Gamma_1(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) | a \equiv 1 \pmod{N}, c \equiv 0 \pmod{N} \right\}$$

The definition of the fundamental domain extends in the obvious way to Fuchsian groups. Namely, we have

Definition 4.2.3. *Let Γ be a Fuchsian group. A* fundamental domain *of Γ is an open connected set (that is, a domain) $D \subset \mathbb{H}$ such that*

1. *For every $\tau \in \mathbb{H}$, there exists some $\gamma \in \Gamma$ and some $\tau' \in \overline{D}$, that is in the closure of D, such that $\gamma \cdot \tau' = \tau$.*

2. *No two elements of D are in the same Γ-orbit, that is, for any $\tau, \tau' \in D$, $\gamma \cdot \tau \neq \tau'$ for all $\gamma \in \Gamma$.*

Definition 4.2.4. *Let Γ be a Fuchsian group. It is called a* Fuchsian group of the first kind *if it has a fundamental domain of finite hyperbolic area. Otherwise, it is called a* Fuchsian group of the second kind.

Fuchsian groups of the first kind are the general objects of interest; the definition of Fuchsian groups of the second kind in essence means that they are too sparse. As an example, take the group generated by $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. It has as a fundamental domain the entire strip $\left\{\tau \in \mathbb{H} || Re(\tau)| < \frac{1}{2}\right\}$, which has infinite hyperbolic area. In fact, almost all Fuchsian groups Γ appearing in the literature are *commensurable* with $SL_2(\mathbb{Z})$, that is to say, that $\Gamma \cap SL_2(\mathbb{Z})$ has finite index in both $SL_2(\mathbb{Z})$ and Γ. These are automatically of the first kind. For more information, see for example [14] or [3]. We will assume from here forward that our Fuchsian group satisfy these conditions.

$\Gamma(2)$ is a Fuchsian group of the first kind, commensurable with $\Gamma(1)$. A fundamental domain for $\Gamma(2)$ can be obtained by taking the union of six appropriately chosen fundamental domains for $\Gamma(1)$; see [20][pg 141-142] for details.

If we take the fundamental domain of a Fuchsian group and identify the edges of the boundary appropriately, that is if we take $\mathbb{H}$ modulo the action of a Fuchsian group Γ, this will give us a Riemann surface with some number of "punctures". These punctures correspond to the Γ-orbits with non-trivial stabilizers; we remove these orbits as otherwise they'd correspond to conical singularities on the Riemann surface. The genus of this surface is what we call the *genus* of the Fuchsian group Γ. For instance, in the case of $\Gamma(1)$, we have a Riemann sphere with three punctures, corresponding to the $\Gamma(1)$ orbits of ∞ (the cusps), and those of i and $e^{2\pi i/3}$ (stabilized by S and U respectively).

Fuchsian groups with genus zero and three punctures are called *triangle groups*. They are the Fuchsian groups directly related to the hypergeometric function. $\Gamma(2)$ is another example of a triangle group; in fact, any other triangle group is a homomorphic image of $\Gamma(2)$. It is for this reason that $\Gamma(2)$ is our main focus. For more information and other examples see [14] or [10].

To generalize modular forms to Fuchsian groups, we need a growth condition on the function f. These definitions follow [3].

Definition 4.2.5. *Let $f(\tau)$ be a (scalar-valued) meromorphic function on $\mathbb{H}$. We say that f has* moderate growth at ∞ *if there exists some $z \in \mathbb{C}$ and some $Y \in \mathbb{R}$ such that*

$$|f(x+iy)| < e^{Im(z\tau)} \quad \textit{for all } y > Y.$$

where $\tau = x + iy$. Let $c \in \mathbb{R}$. Then there exists some $\gamma \in PSL_2(\mathbb{R})$ such that $\gamma \cdot c = \infty$. If $f(\gamma \cdot \tau)$ has moderate growth at infinity, then we say that $f(\tau)$ has moderate growth at c.

To complete our definition we will need the notion of cusps. Roughly speaking, we can think of cusps as being orbits of where the boundary of the fundamental domain hits the real line together with ∞. More formally, we have the following definition, following [3]:

Definition 4.2.6. *Let Γ be a Fuchsian group of the first kind. An element $\gamma \in \Gamma$ is called* parabolic *if its trace is equal to 2. A point in $\mathbb{R} \cup \{\infty\}$ is called a* cusp *if it is fixed by some parabolic element of Γ. In the case of Fuchisan groups commensurable with $\Gamma(1)$, the cusps are always $\mathbb{Q} \cup \{\infty\}$.*

Definition 4.2.7. *Let Γ be a Fuchsian group of the first kind with a cusp at infinity, $k \in 2\mathbb{Z}$, ρ a representation. A* modular form of weight k for ρ *is a function on $\mathbb{H}$ such that*

1. *$f(\gamma\tau) = (c\tau + d)^k f(\tau)$ for all $\tau \in \mathbb{H}$ and all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$,*
2. *$f(\tau)$ has finitely many poles in $\overline{D} \cap \mathbb{H}$, where $\overline{D}$ is the closure of the fundamental domain of Γ,*
3. *$f(\tau)$ has moderate growth at all cusps c of Γ.*

A modular function *is a modular form of weight 0.*

In particular, we can use this to define modular forms for $\Gamma(2)$. Their Fourier expansions will be in $\tilde{q} := e^{\pi i \tau}$, as this is now the variable which is fixed by the stabilizer of infinity. $\Gamma(2)$ has three inequivalent cusps, namely $0, 1$ and ∞, as opposed to $SL_2(\mathbb{Z})$'s single cusp, slightly complicating definitions of holomorphic and weakly holomorphic. Namely, meromorphic at infinity is as for $\Gamma(1)$ except using the $\tilde{q}$-expansion, meromorphic at 0 means $f(-\frac{1}{\tau})$ is meromorphic at infinity, and meromorphic at 1 means $f(-\frac{1}{\tau-1})$ is meromorphic at infinity. Holomorphicity is defined similarly.

In the case of Fuchsian group Γ of genus zero, we have a uniformizing function, called a *Hauptmodul*, which maps our surface (with its conical singularities) to the Riemann sphere. This will be a modular function for our group Γ, and in fact any modular function for Γ can be written as a rational function in the Hauptmodul. As an example, the standard Hauptmodul for $\Gamma = \Gamma(1)$ is the well-known j-function,

$$j(\tau) = \frac{E_4^3}{\eta^2 4} = q^{-1} + 744 + 196884q + \dots$$

For more information, once again see [14].

To construct a Hauptmodul for $\Gamma(2)$, we begin with the well-known classical theta series,

$$\begin{aligned}
\theta_2(\tau) &= \sum_{n=-\infty}^{\infty} q^{(n-1/2)^2/2} &&= 2q^{1/8}\prod_{n=1}^{\infty}(1-q^n)(1+q^n)^2 \\
\theta_3(\tau) &= \sum_{n=-\infty}^{\infty} q^{n^2/2} &&= \prod_{n=1}^{\infty}(1-q^n)(1+q^{n-1/2})^2 \\
\theta_4(\tau) &= \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2/2} &&= \prod_{n=1}^{\infty}(1-q^n)(1-q^{n-1/2})^2.
\end{aligned}$$

A Hauptmodul for $\Gamma(2)$ is $-\frac{16\theta_3^4(\tau)}{\theta_2^4(\tau)}$. The holomorphic modular forms will be generated freely by the weight 2 forms θ_2^4 and θ_3^4. Note that $\theta_3^4 = \theta_2^4 + \theta_4^4$.

4.3 Vector Valued Modular Forms

This definition can be further generalized to allow for multidimensional representations, giving rise to the notion of vector valued modular forms. In this case, we usually need some additional admissibility conditions on our representation. For more information, see [15]. Namely, we usually assume some diagonalization condition on the stabilizer of $i\infty$, depending on our group. In the case of $\Gamma(1)$, this means $\rho(T)$ is diagonal; in the case of $\Gamma(2)$, our main concern later, we will assume that $\rho(T^2)$ is diagonal. This allows for a q- or $\tilde{q}$-series, as explained for example in [4]. Although we define vector valued modular forms only for even integer weight, a more general definition is possible; however this definition will be sufficient for our purposes.

Definition 4.3.1. *Let Γ be a genus zero Fuchsian group of the first kind with a cusp at infinity, and $\rho : \Gamma \to GL_n(C)$ be a rank n representation with associated weight $k \in 2\mathbb{Z}$ for some $n \in \mathbb{Z}_{>0}$. Let $\overline{F}$ be the closure of the fundamental domain of Γ. A* vector valued modular form of weight k *is a meromorphic function $\mathbb{X} : \mathbb{H} \to \mathbb{C}^d$ such that*

1. $\mathbb{X}(\gamma\tau) = \rho(\gamma)(c\tau + d)^k \mathbb{X}(\tau)$ *for all* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$,

2. $\mathbb{X}(\tau)$ *has finitely many poles in* $\overline{F} \cap \mathbb{H}$,

3. *Each component of* $\mathbb{X}(\tau)$ *has moderate growth at all cusps* c *of* Γ.

Such a form is called weakly holomorphic *if it is holomorphic everywhere except the orbit of* ∞*; if it is holomorphic including at* ∞*, then it is called* holomorphic. *A vector valued modular form of weight* 0 *is called a* vector valued modular function.

For brevity, we will from here on sometimes refer to vector valued modular forms simply as modular forms.

Definition 4.3.2. *Let* $\Gamma = \Gamma(1)$ *or* $\Gamma(2)$. *Let* $\mathbb{X}(\tau)$ *be a vector valued modular form for* Γ. *Then we can write*

$$\mathbb{X}(\tau) = q^\lambda \sum_{m=M}^{\infty} X[m] q^m$$

(if $\Gamma = \Gamma(1)$*) or*

$$\mathbb{X}(\tau) = \tilde{q}^\lambda \sum_{m=M}^{\infty} X[m] \tilde{q}^m$$

(if $\Gamma = \Gamma(2)$*), for some* $M \in \mathbb{Z}$ *and* $X[m], \lambda \in \mathbb{C}^d$, *where* λ *is a diagonal matrix. The* $X[m]$ *are called the* Fourier coefficients *of* $\mathbb{X}(\tau)$.

As mentioned previously, such vector valued modular forms show up in places such as the characters of vertex operator algebras. A more concrete example is the θ series, however since they form a vector valued modular form of half-integral weight, which we have not defined, we will gloss over many of the details. However, as their relations will be important to us later, we will nevertheless mention them. The reader may refer to for example [4] for more information on modular forms of half-integral weight.

Let $\Theta(\tau) = \begin{pmatrix} \theta_2(\tau) \\ \theta_3(\tau) \\ \theta_4(\tau) \end{pmatrix}$. Since the modular group is generated by S and T,

it is sufficient to find the representation for these. It can be shown that we have

$$\Theta(T\tau) = \Theta(\tau+1) = (1)^{1/2} \begin{pmatrix} \varepsilon_8 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \Theta(\tau),$$

where ε_8 is the 8th root of unity $\varepsilon_8 = e^{2\pi i/8}$, and

$$\Theta(S\tau) = \Theta(-\frac{1}{\tau}) = (\tau)^{1/2} \begin{pmatrix} 0 & 0 & \sqrt{\frac{1}{i}} \\ 0 & \sqrt{\frac{1}{i}} & 0 \\ \sqrt{\frac{1}{i}} & 0 & 0 \end{pmatrix} \Theta(\tau).$$

(We gloss over the subtleties of complex square roots here, as they will be unimportant to us.) Hence we have a vector valued modular form of weight $\frac{1}{2}$. These also form a vector valued modular form for $\Gamma(2)$, as will be important for us later. Recall that $\Gamma(2)$ is generated by $A = T^2$ and $B = ST^2S^{-1}$; $\rho(A)$ and $\rho(B)$ are easily calculated from the above computations. In particular, $\rho(T^2)$ is diagonal.

4.4 Theory of Vector Valued Modular Forms and Fuchsian Differential Equations

We now briefly describe the space and theory of vector valued modular forms, focusing on those for $\Gamma(1)$ and $\Gamma(2)$, which are essentially the same. We give only a brief overview here, largely without proof; for more information, see for example [4] and [15] for the case of $\Gamma(1)$, and [10], where $\Gamma(2)$ is referred to as a triangle group of type (∞, ∞, ∞), for the case of $\Gamma(2)$.

Definition 4.4.1. *Let $\Gamma = \Gamma(1)$ or $\Gamma(2)$. We define $\mathcal{M}^!_{k,\Gamma}(\rho)$ to be the space of weight k weakly holomorphic vector valued modular forms, and $\mathcal{M}^{hlm}_{k,\Gamma}(\rho)$ to be the space of weight k holomorphic vector valued modular forms. We may omit the Γ in the subscript when the group in question is obvious from context.*

These are always vector spaces, and $\mathcal{M}^{hlm}_{k,\Gamma}(\rho)$ is of finite dimension. In the case of $\Gamma(1)$, $\oplus_{k\in 2\mathbb{Z}} M^{hlm}_{k,\Gamma(1)}(\rho)$ is a free module of rank $\dim \rho$ over $\mathbb{C}[E_4, E_6]$, where the E_i are the Eisenstein series defined previously, and $\mathcal{M}^!_{k,\Gamma(1)}(\rho)$ is a free module of dimension $\dim \rho$ over $\mathbb{C}[j]$, where the j function is the usual Hauptmodul for $\Gamma(1)$. The case of $\Gamma(2)$ is similar, but with different functions; E_4 and E_6 are replaced by θ_2^4, θ_3^4, and the Hauptmodul j is replaced by $\frac{\theta_3^4}{\theta_2^4}$, where the θ_i are the classical theta series defined previously.

Also important to us are differential operators, from which we construct our Fuchsian differential equations. We mentioned these briefly previously, and will here add only slightly more detail and a slightly different emphasis. Although a definition exists in more generality, we will here use a restricted definition that will be sufficient for our purposes. For more on Fuchsian differential equations and equations of hypergeometric type, see for example [5] and section 9.6 of [19].

Definition 4.4.2. *A* Fuchsian differential equation *is a linear ordinary differential equation living on the Riemann sphere with only regular singularities, including at ∞.*

By necessity a Fuchsian differential equation has only finitely many singularities. We are most interested in Fuchsian differential equations of hypergeometric type, which we briefly mentioned when talking about hypergeometric equations. Recall that they are of the following form:

Definition 4.4.3. *Let $\theta = z\frac{d}{dz}$, and let P be the differential operator $P := \theta^n + p_1\theta^{n-1} + \ldots + p_{n-1}\theta + p_n$. A Fuchsian differential equation with regular singularities only at the points $z = 0, 1, \infty$ is called a differential equation of* hypergeometric type*, or simply a* hypergeometric equation*, if each p_i is of the form $p_i(z) = p_{i,0} + p_{i,1}(z-1)^k$, where the $p_{i,k} \in \mathbb{C}$. We call n the* order *of the hypergeometric equation.*

We can rearrange such equations to be of the following form:

$$\begin{aligned} &D(\alpha_1, \ldots, \alpha_n; \beta_1, \ldots, \beta_n) \\ &= (\theta + \beta_1 - 1)\cdots(\theta + \beta_n - 1) - z(\theta + \alpha_1)\cdots(\theta + \alpha_n) = 0 \end{aligned}$$

It is known that we have a basis of solutions for such equations which are generalized hypergeometric functions up to some multiple of z^β. It is known that every Fuchsian differential equation of order 2 with three regular singularities is, up to a change of variables, the standard hypergeometric equation for ${}_2F_1$. We can recover the standard order 2 hypergeometric equation from this one by setting $\beta_2 = 1$; then, setting $\alpha_1 = a$, $\alpha_2 = b$, $\beta_1 = c$ we have

$$\begin{aligned}
&D(a,b;c,1)f(z) \\
&= \left[\left(z\frac{d}{dz} + c - 1\right)\left(z\frac{d}{dz}\right) - z\left(z\frac{d}{dz} + a\right)\left(z\frac{d}{dz} + b\right)\right] f(z) \\
&= z^2 f''(z) + czf'(z) - z(zf'(z) + z^2 f''(z) + bzf'(z) + azf'(z) + abf(z)) \\
&= z[z(1-z)f''(z) + [c - z(a+b+1)]f'(z) + abf(z)] = 0
\end{aligned}$$

Hence removing the factor of z, which we may do since $zg(0) = 0$ if and only if $g(z) = 0$ for any meromorphic function $g(z)$, we have obtained the standard hypergeometric equation.

Another formulation of Fuchsian and hypergeometric equations is also useful to us. Choose k distinct points z_i on the Riemann sphere, and k distinct constant matrices A_i. Then the first-order matrix differential equation

$$\frac{d}{dz}\Xi = \sum_i \frac{A_i}{z - z_i}\Xi$$

is called *Fuchsian*. Any order-n Fuchsian differential equation can be recast in this way (but not conversely, for $n > 2$). When $k = 3$, and $z_i = 0, 1, \infty$, we say this differential equation is of *hypergeometric type*.

As mentioned previously, if we assume that $\rho(T)$ is diagonal in the case of $\Gamma(1)$, or $\rho(T^2)$ is diagonal in the case of $\Gamma(2)$, we get that our vector valued modular forms will have Fourier coefficients. From here on we will assume this is always the case. Note that in the case of $\Gamma(2)$, our Fourier coefficients will be for a series in $\tilde{q} = e^{\pi i \tau}$ instead of the usual $q = e^{2\pi i \tau}$.

In the following let $\Gamma = \Gamma(1)$ or $\Gamma(2)$.

Definition 4.4.4. *Let $\mathbb{K}$ be some subring of $\mathbb{C}$. We call a modular form $\mathbb{K}$-*Fourier *if all Fourier coefficients of each component lie in $\mathbb{K}$. In particular, if*

all Fourier coefficients of each component lie in $\mathbb{Q}$, we call the modular form rational*; if they lie in $\mathbb{Z}$, we call the modular form* integral*.*

Definition 4.4.5. *A vector valued modular form is called* full *if its components are linearly independent.*

We will want our vector valued modular forms to be full, otherwise it will see only part of the representation. Due to this and the nature of our problem, we will mostly be dealing with the following types of modular form:

Definition 4.4.6. *A vector valued modular form is called a* full rational *modular form if it is both full and $\mathbb{Q}$-Fourier; it is called a* full integral *modular form if it is both full and $\mathbb{Z}$-Fourier.*

Let ρ be a representation. It is known that the following are equivalent:

1. There is at least one full rational (resp. full integral) weakly holomorphic vector valued modular form;

2. There exists a free basis over the ring $\mathbb{C}[E_4, E_6]$ (in the case of $\Gamma(1)$) or $\mathbb{C}[\theta_2^4, \theta_3^4]$ (in the case of $\Gamma(2)$) of full rational (resp. full integral) vector valued modular forms for $\oplus_{k\in 2\mathbb{Z}} M_k^{hlm}(\rho)$;

3. There exists a free basis over the ring $\mathbb{C}[E_4, E_6]$ (in the case of $\Gamma(1)$) or $\mathbb{C}[\theta_2^4, \theta_3^4]$ (in the case of $\Gamma(2)$) of full rational (resp. full integral) vector valued modular forms for $\oplus_{k\in 2\mathbb{Z}} M_k^!(\rho)$.

It is known that if ρ satisfies any of these conditions, $\rho(T)$ (for $\Gamma(1)$) or $\rho(T^2)$ (for $\Gamma(2)$) must have finite order, as shown for instance in [1].

Definition 4.4.7. *If ρ satisfies any (and therefore all) of the above three conditions, we call it* rational *(or, respectively,* integral*).*

Recall that we previously defined a congruence subgroup as one containing some $\Gamma(N)$.

Definition 4.4.8. *Let ρ be a representation. We call ρ a* congruence representation *if its kernel is a congruence subgroup, that is, contains some $\Gamma(N)$.*

The classification of congruence representations is known for dimensions at least 1 through 5 for $\Gamma(1)$ and dimensions $1, 2$ and 3 for $\Gamma(2)$. Our main concern is $\Gamma(2)$. Since $\Gamma(2)$ is the free group on two generators, an n-dimensional rep is a pair of invertible matrices a, b, corresponding to $\rho(A) = a$, $\rho(B) = b$, where A, B are the generators of $\Gamma(2)$ defined previously. We denote such a representation by $\rho_{a,b}$. In particular, a 1-dimensional congruence representation of $\Gamma(2)$ is a pair of numbers $a, b \in \mathbb{C}^\times$. By [16], $\rho_{a,b}$ is a 1-dimensional congruence representation if and only if $a^{24} = 1$ and $a^8 = b^8$, which implies that there are precisely 192 1-dimensional congruence representations for $\Gamma(2)$. Likewise, there are exactly 912 irreducible congruence representations of $\Gamma(2)$ in dimension 2, and 832 in dimension 3; see [16] for more details.

4.5 The Atkin-Swinnerton-Dyer Conjecture

We are now ready to sketch the conjecture which motivated our work in Chapter 3 on hypergeometric functions.

In [2], Atkin-Swinnerton-Dyer made an observation which has since been elevated to a conjecture. It can be generalized to the setting of vector valued modular forms of arbitrary Fuchsian groups in the obvious way, and call this generalization vASD.

Conjecture 4.5.1 (vASD)**.** *Suppose that ρ is a rational representation for some Fuchsian group commensurable with $\Gamma(1)$. Then ρ has a full integral vector valued modular form, if and only if ρ is congruence.*

It has been known for at least a hundred years that any rational congruence representation is necessarily an integer representation; hence one direction of vASD is a theorem. However, although the Atkin-S-D conjecture is 50 years old, and its generalization vASD has been studied for at least the last decade, it can be argued that we still have little evidence in support of it. See [28] and [13] for recent results and some discussion of the literature.

We will need the following definition:

Definition 4.5.1. *Let ρ be a rational representation for $\Gamma(1)$ or $\Gamma(2)$. A prime p is called an* unbounded prime *for ρ if it appears to arbitrarily high powers in*

the denominators of the coefficients of one (and hence all) full vector valued modular form for ρ, that is, if that vector valued modular form's coefficients are p-adically unbounded. For clarity, we say the coefficients $X[m]$ are p-adically unbounded *if the infimum over all i and m of $\nu_p(X[m]_i) = -\infty$. That is, it is sufficient that one component be p-adically unbounded.*

In dimensions 1 and 2 for $\Gamma(1)$, a representation is congruence if and only if it is finite image. We now know that any rational representation of $\Gamma(1)$ or $\Gamma(2)$ with infinite image is necessarily p-unbounded for almost all primes p. Hence in hindsight, the verification of vASD for $\Gamma(1)$ in dimensions 1 and 2 is not deep. However, $\Gamma(2)$ is another story!

4.6 vASD for $\Gamma(2)$ in 1 Dimension

In the 1-dimensional case for $\Gamma(2)$, we have the classification of congruence forms given above. We will now show that vASD holds for $\Gamma(2)$ in 1 dimension.

First we will need the following two simple facts.

Lemma 4.6.1. *Let $h(x) = 1 + \sum_{n=N}^{\infty} h_n x^n$, where $N \geqslant 1$, and the $h_n \in \mathbb{Q}, h_N \neq 0$. Let p be any prime which is coprime to the denominators of all the coefficients h_n. Then,*

1. *p is also coprime to the denominators of all the coefficients of $\frac{1}{h(x)}$; and*

2. *Let r be a rational number, p a prime which divides the denominator of r, and such that $v_p(r) < -v_p(h_N)$. Then p appears to arbitrarily high powers in the denominator of $h(x)^r$ (ie p is an* unbounded prime *for $h(x)^r$).*

Proof. Recall Newton's binomial formula, $(1+a)^y = 1+ya+\ldots+\frac{y(y-1)\cdots(y-k+1)}{k!}a^k+\ldots$. Subbing in $a = \sum_{n=N}^{\infty} h_n x^n$, $y = -1$, we get

$$1 - \left(\sum_{n=N}^{\infty} h_n x^n\right) + \left(\sum_{n=N}^{\infty} h_n x^n\right)^2 - \ldots \pm \left(\sum_{n=N}^{\infty} h_n x^n\right)^k \mp \ldots$$

Clearly any prime which is coprime to all the denominators of the h_n will still be coprime to the denominators of this sum.

For part 2, we can use the same formula, now with $a = \sum_{n=N}^{\infty} h_n x^n$, $y = r$. Rather than focusing on the whole thing, we will focus on the coefficient of each x^{Nn}. This will be a polynomial in r of degree k, whose leading term comes from $(\sum_{n=N}^{\infty} h_n x^n)^k$ and hence is $\frac{r^k h_N^k}{k!}$. All other terms will come from equal or lower powers in the expansion of the binomial formula, and hence the coefficients of the other r^k will be polynomials in $\frac{1}{n}\mathbb{Z}[h_N, \ldots, h_{(n)N}]$. Hence, in particular, all these terms will have a strictly smaller power of p in their denominator than the leading term. Thus, we can find the power of p in the denominator of the coefficient of x^{Nn} as being precisely the power of p in the denominator of $\frac{r^n h_N^n}{(n)!}$. Clearly, since $v_p(r^n) < -v_p(h_N^n)$, this power increases monotonically as n increases; therefore, p is an unbounded prime for $h(x)^r$. $\square$

We recall the earlier introduced θ series. Their transformations under S, T are recorded in a matrix above, but we will restate them in a different format here for the fourth power of each of these series, which we will use in the coming calculations.

$$\theta_2^4(\tau+1) = -\theta_2^4(\tau); \qquad \theta_2^4(-1/\tau) = -\tau^2\theta_4^4(\tau);$$
$$\theta_3^4(\tau+1) = \theta_4^4(\tau); \qquad \theta_3^4(-1/\tau) = -\tau^2\theta_3^4(\tau);$$
$$\theta_4^4(\tau+1) = \theta_3^4(\tau); \qquad \theta_4^4(-1/\tau) = -\tau^2\theta_2^4(\tau).$$

It is known that none of these fourth powers have zeroes or poles in $\mathbb{H}$, which is a simply connected domain. Hence, they have a well-defined, holomorphic logarithm. We can multiply this logarithm by any complex number and exponentiate without disrupting holomorphicity. Hence, these $\frac{\theta_i^4}{\theta_j^4}$ can be raised to arbitrary complex powers in $\mathbb{H}$ and still be well-defined and holomorphic. Moreover, as stated previously, we can use elements of $\Gamma(1)$ to map any cusps to the cusp at infinity, and from the q or $\tilde{q}$ expansion of the transformed function read of the T or A action.

Recall from above that the 1-dimensional representations of $\Gamma(2)$ can be

classified by choosing a complex number for each of $A = T^2$ and $B = ST^2S^{-1}$. As before, let $\rho_{a,b}$ be the representation defined by $\rho_{a,b}(A) = a$, $\rho_{a,b}(B) = b$.

Lemma 4.6.2. *$\rho_{a,b}$ is rational (with $\rho(T^2)$ diagonal since we are in the one-dimensional case) if and only if a and b are both roots of unity, that is, if and only if $\rho_{a,b}$ has finite image.*

Proof. Let

$$f(\tau) = \frac{\theta_2^4(\tau)}{16\theta_3^4(\tau)} = \tilde{q} - 8\tilde{q}^2 + 44\tilde{q}^3 + \dots \ ,$$
$$g(\tau) = \frac{\theta_2^4(\tau)}{16\theta_4^4(\tau)} = \tilde{q} + 8\tilde{q}^2 + 44\tilde{q}^3 + \dots \ .$$

Note that A is the stabilizer of the cusp at infinity, and B is the stabilizer of the cusp at 0. We can send the cusp at 0 to the cusp at infinity by using the S matrix. Hence, we can examine the behaviour of powers of these functions with respect to the modular transformations by looking at the $\tilde{q}$-expansions at infinity. For example, looking at $f(\tau)$, we have

$$f^r(\tau) = \tilde{q}^r(1 + 8\tilde{q} + 44\tilde{q}^2 + \dots)^r.$$

By the binomial formula we have that the second part of this will still be a $\tilde{q}$-expansion, hence the part we care about is that first factor of $\tilde{q}$. This tells us that

$$\begin{aligned} f^r(A \cdot \tau) &= f^r(\tau + 2) \\ &= e^{\pi i(\tau+2)r}(1 + 8\tilde{q} + \dots) \\ &= e^{2\pi i r} f^r(\tau). \end{aligned}$$

Similarly, we have

$$g^s(A\tau) = e^{2\pi i s} g(\tau)$$

.

To get the answer for $B\tau$ we use a transformation sending the cusp 0 to the cusp at ∞, a transformation which due to earlier calculations we understand

well as it is simply the S matrix. We can then once again look at the expansion around this cusp to determine how this will act. Namely, we know that sending $\tau \mapsto -1/\tau$ sends

$$f(-1/\tau) = \frac{\theta_4^4(\tau)}{16\theta_3^4(\tau)},$$
$$g(-1/\tau) = \frac{\theta_4^4(\tau)}{16\theta_2^4(\tau)} .$$

Then for the $\tilde{q}$ expansion we have

$$f(-1/\tau) = \frac{\prod_{n=1}^{\infty}(1-\tilde{q}^{2n-1})^8}{16\prod_{n=1}^{\infty}(1+\tilde{q}^{2n-1})^8},$$

which clearly will have no pre-factor of $\tilde{q}$. Hence f transforms trivially under B. Similarly, the expansion for

$$g(-1/\tau) = \frac{\prod_{n=1}^{\infty}(1-\tilde{q}^{2n-1})^8}{16^2\tilde{q}\prod_{n=1}^{\infty}(1+\tilde{q}^{2n})^8}$$

has a pre-factor of $\tilde{q}^{-1}$, hence $g^s(B\cdot\tau) = e^{-2\pi i s}g^s(\tau)$.

Let us write $a = e^{2\pi i r}$, $b = e^{2\pi i s}$ for some $r, s \in \mathbb{C}$. Then by the above calculations, $\mathbb{X}_{r,s} := f(\tau)^{r+s}g(\tau)^{-s}$ is a vector valued modular form for $\rho_{a,b}$.

Note that, using the binomial formula, if both r and s are rational, then clearly $\mathbb{X}_{r,s}(\tau)$ will have rational $\tilde{q}$-coefficients. Conversely, let's look at the first few coefficients of $\mathbb{X}_{r,s}(\tau)$. Note that the first few coefficients of f and g are

$$f(\tau) = \tilde{q} - 8\tilde{q}^2 + 44\tilde{q}^3 - \dots$$
$$g(\tau) = \tilde{q} + 8\tilde{q}^2 + 44\tilde{q}^3 + \dots$$

Using the binomial formula again, we can calculate the first few terms of $\mathbb{X}_{r,s}(\tau)$:

$$\mathbb{X}_{r,s}(\tau) = \tilde{q}^{r+s}(1+(-8\tilde{q}+44\tilde{q}^2-\dots))^{r+s}\tilde{q}^{-s}(1+(8\tilde{q}+44\tilde{q}^2+\dots))^{-s}$$

$$
\begin{aligned}
=&\tilde{q}^r\left(1+(r+s)(-8\tilde{q}+44\tilde{q}^2-\ldots)+\frac{(r+s)(r+s-1)}{2!}(-8\tilde{q}+\ldots)^2+\ldots\right)\\
&\cdot\left(1+(-s)(8\tilde{q}+44\tilde{q}^2-\ldots)+\frac{(-s)(-s-1)}{2!}(-8\tilde{q}+\ldots)^2+\ldots\right)\\
=&\tilde{q}^r(1-(16s+8r)\tilde{q}+(32r^2+128rs+128s^2+12r)\tilde{q}^2+\ldots).
\end{aligned}
$$

Now if $\mathbb{X}_{r,s}$ is $\mathbb{Q}$-Fourier, then clearly $16s+8r, 32r^2+128rs+128s^2+12r \in \mathbb{Q}$. But then also $12r = 32r^2+128rs+128s^2+12r-32(2s+r)^2 \in \mathbb{Q}$. Hence $r, s \in \mathbb{Q}$. □

Theorem 4.6.3. *$\rho_{a,b}$ is integral if and only if $a^{24}=1$ and $a^8=b^8$, that is to say, if and only if the kernel of $\rho_{a,b}$ contains a congruence subgroup.*

Proof. Suppose first that $a^{24}=1$ and $a^8=b^8$. Define r, s as in the proof of Lemma 4.6.2. Note that if we can show that if $f^{1/8}$, $g^{1/8}$ and $(fg)^{1/3}$ have no unbounded primes, then neither will $\mathbb{X}_{r,s} = f^{r+s}g^{-s}$. To see this, write $r = k/8 + l/3$,$s = k'/8 + l'/3$ for integers k, k', l, l'. Then $l \equiv l' \pmod 3$, so without loss of generality we can choose $l = l' \geq 0$. Likewise, we can insist that k, k' are both non-negative. Then for these modified but equally acceptable r, s, $\mathbb{X}_{r,s}(\tau) = (f^{1/8})^k(g^{1/8})^{k'}((fg)^{1/3})^l$. Thus we will know $\mathbb{X}_{r,s}$ is integral, if we know $f^{1/8}$, $g^{1/8}$ and $(fg)^{1/3}$ are.

To see that $f^{1/8}$ and $g^{1/8}$ have no unbounded primes, we use the product formulas for the theta series, in terms of $\tilde{q}$, which we defined previously. We have

$$
\begin{aligned}
f(\tau)^{1/8} &= \left(\frac{2\tilde{q}^{1/4}\prod_{n=1}^{\infty}(1-\tilde{q}^{2n})(1+\tilde{q}^{2n})^2}{2\prod_{n=1}^{\infty}(1-\tilde{q}^{2n})(1+\tilde{q}^{2n-1})^2}\right)^{1/2}\\
&= \tilde{q}^{1/8}\frac{\prod_{n=1}^{\infty}(1+\tilde{q}^{2n})}{\prod_{n=1}^{\infty}(1+\tilde{q}^{2n-1})}
\end{aligned}
$$

The Fourier expansions of both the numerator and the denominator clearly have integral coefficients; hence by Lemma 4.6.1 part 1, so does $f(\tau)$. The same is true for

$$
g(\tau)^{1/8} = \tilde{q}^{1/8}\frac{\prod_{n=1}^{\infty}(1+\tilde{q}^{2n})}{\prod_{n=1}^{\infty}(1-\tilde{q}^{2n-1})}.
$$

For $(fg)^{1/3}$, we use the well-known identity $\theta_2\theta_3\theta_4 = 2\eta^3$, where η is the

Dedekind eta function, which is integral by the previously shown product formula. We thus have

$$\begin{aligned}(fg(\tau))^{1/3} &= \left(\frac{\theta_2^8(\tau)}{(16)^2\theta_3^4(\tau)\theta_4^4(\tau)}\right)^{1/3} \\ &= \left(\frac{2^8(\eta^3(\tau))^8}{(16)^2\theta_3^{12}(\tau)\theta_4^{12}(\tau)}\right)^{1/3} \\ &= \frac{\eta^8(\tau)}{\theta_3^4(\tau)\theta_4^4(\tau)}.\end{aligned}$$

Once again, since all the coefficients of the expansions of both the numerator and denominator are integral, we can use Lemma 4.6.1 part 1 to say that the coefficients of $(fg(\tau))^{1/3}$ also have no unbounded primes.

Conversely, suppose that we have $\mathbb{X}_{r,s} = f^{r+s}g^{-s}$ is integral, that is, suppose that $\rho_{a,b}$ is integral. We know already that $r, s \in \mathbb{Q}$; we will show that they must satisfy the conditions that their denominators must divide 24 and $24r \equiv 24s \pmod 3$, that is, $a^{24} = 1, a^8 = b^8$.

Let p be some prime which divides the denominator of either r or s. Suppose first that $\nu_p(r+s) \leqslant \nu_p(-s)$. Note that we must have $\nu_p(r+s) < 0$. Hence we can find some $L \in \mathbb{Z}_{\geqslant 0}$ such that $L(r+s) \equiv \frac{1}{p} \pmod 1$; for instance, if $r+s = \frac{n}{p^k m}$, then we have $p^{k-1}m(r+s) = \frac{n}{p}$. Then there exists some $c \in \mathbb{Z}_{\geqslant 0}$ such that $cn \equiv 1 \pmod p$. Set $L = p^{k-1}mc$.

Now since $L \in \mathbb{Z}_{\geqslant 0}$, and $\mathbb{X}_{r,s}$ is integral, we must have that $\mathbb{X}_{r,s}^L = (f^{r+s}g^{-s})^L$ is integral. Then dividing by positive integer powers of f and g, which will similarly have no unbounded primes, we get that $f^{1/p}g^{l/p}$ must have no unbounded primes, where $0 \leqslant l < p$. Now, by our previous calculation,

$$fg^l = \tilde{q}^{l+1}(1 + 8(l-1)\tilde{q} + (32l^2 - 52l + 44)\tilde{q}^2 + \ldots).$$

Similarly, in the case of $\nu_p(s) < \nu_p(r+s)$ we get that for some $0 \leqslant l < p$, $f^{l/p}g^{1/p}$ has no unbounded primes, and

$$f^l g = \tilde{q}^{l+1}(1 + 8(1-l)\tilde{q} + (32l^2 - 52l + 44)\tilde{q}^2 + \ldots).$$

If $p > 3$, then by Lemma 4.6.1 part 2 with $N = 1$, unless $l = 1$ in which

case $N = 2$, we have that p is an unbounded prime, giving a contradiction. Hence no $p > 3$ can divide the denominators of r or s. Moreover, by the same lemma, the same is true for powers of 2 greater than 2^3, since 8 divides the first coefficient and the second with $l = 1$; similarly, the only way that $p = 3$ can divide the denominator of r or s is if $l = 1$, since then we have $32 - 52 + 44 = 24$ and $8(1 - l) = 0$. That is, $24r \equiv 24s \pmod 3$. Moreover, no higher power of 3 than the first power can divide the denominators. Thus we must have $a^{24} = 1, a^8 = b^8$, as desired. □

4.7 Higher dimensional modular forms and hypergeometric equations

The two portions of this thesis come together when we look at the vASD conjecture for higher dimensional representations of $\Gamma(2)$. Unfortunately, the details of this connection are beyond the scope of this thesis. For the story in the case of $\Gamma(1)$ see for example [3], [4] and [15]; we will give a very brief sketch here. The story for $\Gamma(2)$ will be similar, but has not yet been published.

Given a representation ρ for $\Gamma(1)$ satisfying certain extra conditions, we have the space of weakly holomorphic vector valued modular forms, $\mathcal{M}_0^!(\rho)$, where here and from now on $\Gamma = \Gamma(1)$. As we stated previously, this is a free module of rank $n = \dim(\rho)$ over $\mathbb{C}[j]$. We can thus find a free basis $\mathbb{X}^{(1)}, \ldots, \mathbb{X}^{(n)}$. Additionally, these free basis vectors satisfy differential relations. Specifically, these arise from the differential operator ∇ defined by

$$\nabla = \frac{E_{10}(\tau)}{2\pi i \cdot \eta^{24}(\tau)} \frac{d}{d\tau},$$

which maps $\mathcal{M}_0^!(\rho)$ to itself. This gives rise to differential relations on the canonical basis vectors.

In fact, we can arrange these canonical basis vectors into a matrix, which we call the *fundamental matrix*. Then, the differential relations on the basis vectors give rise to a first order matrix differential equation of hypergeometric type satisfied by the fundamental matrix. Additionally, this fundamental ma-

trix encodes all the information about $\mathcal{M}_0^!(\rho)$, as it stores the free basis. This differential equation means that the free basis can be completely recovered once a small number (approx n^2) of complex numbers are known.

Consider ρ defined by

$$\rho(T) = \begin{pmatrix} e^{11\pi i/30} & 0 \\ 0 & e^{-\pi i/30}) \end{pmatrix},$$

$$\rho(S) = \sqrt{\frac{2}{5+\sqrt{5}}} \begin{pmatrix} 1 & \frac{1+\sqrt{5}}{2} \\ \frac{1+\sqrt{5}}{2} & -1 \end{pmatrix}.$$

Then the free generators are

$$X^{(1)}(\tau) = \begin{pmatrix} q^{11/60}(1+q^2+\ldots) \\ q^{-1/60}(q^{-1}+1+q+q^2+\ldots) \end{pmatrix},$$

$$X^{(1)}(\tau) = \begin{pmatrix} q^{11/60}(q^{-1}-245-113239q-6029989q^2+\ldots \\ q^{-1/60}(26999+1820504q+\ldots) \end{pmatrix}.$$

In rank 2, this matrix differential equation can always be recast as a (standard) hypergeometric equation. This means that the components of the free basis for rank 2 can always be expressed as a hypergeometric function ${}_2F_1$ times a power of j. The same happens in rank 3, except ${}_3F_2$ is involved. In rank greater than 3, the hypergeometric functions only account for some of the vector valued modular forms (by definition the most accessible ones).

One way to express with ${}_2F_1$ the fundamental matrix for any 2-dim rep of $\Gamma(1)$ is given in Section 4.2 of [15]. Each component is written as $(j(\tau)/1728)$ to some power, times some ${}_2F_1$ for the choice $z = 1728/j(\tau)$.

[13] gives several explicit examples expressing holomorphic rank 3 vector valued modular forms for $\Gamma(1)$ in terms of ${}_3F_2$, for the same choice of z.

As mentioned previously, the story in the case of $\Gamma(2)$ is very similar, and more important in that there are far more representations in each dimension in which to probe vASD. For this reason, the case of $\Gamma(2)$ holds much more interest for the vASD conjecture than does $\Gamma(1)$. As touched on in the Conclusion, vASD in 2 and 3 dimensions for any triangle group, such as $\Gamma(2)$, will

reduce to knowing the unbounded primes for the hypergeometric functions ${}_2F_1$ and ${}_3F_2$, which we solved in this thesis. The higher order hypergeometric functions come into play in higher dimensions for $\Gamma(2)$ and other triangle groups. In particular, the work of this thesis opens the door to the largest test by far of the vASD conjecture to date.

Chapter 5

Conclusion

We have answered the question of unbounded primes for the generalized hypergeometric function, and have shown vASD to be true in the case of 1-dimensional vector valued modular forms for $\Gamma(2)$. Now, there is a clear next direction for this project. Namely, it remains to work out the detailed analysis of vASD for the 2- and 3-dimensional representations of $\Gamma(2)$. This amounts to combining the 1-dimensional analysis (applied to powers of j) with the hypergeometric analysis for ${}_2F_1$ and ${}_3F_2$. The unbounded primes grow much faster in the 1-dimensional case than in the 2-dimensional one, so this shouldn't constitute a major problem, but we get infinite families. Another direction would be to try this for other triangle groups, for example $\Gamma_0(2)$. In rank 4 and higher, can we find how to generalize the hypergeometric functions to recover all $\Gamma(2)$ modular forms?

In this thesis, we considered only those hypergeometric functions and vector valued modular forms with rational coefficients. This should be generalized to coefficients coming from any algebraic number field. The cyclotomic fields are the most important of these. Indeed, they are the proper home of vASD because a basis for the space of modular forms for $\Gamma(N)$ can only be made to be $\mathbb{Z}[e^{2\pi i/N}]$-integral.

Some other open questions in this area would be to look at the question of unbounded primes for other generalizations of the hypergeometric function. For instance, while the series ${}_nF_{n-1}$ are the most interesting as they converge only in a circle of radius 1 around the origin (we have to analytically extend

to the rest of the Riemann sphere, and this introduces branch singularities, from which ρ arises as monodromy), we can also ask the question for any ${}_pF_q$, which are defined as

$$ {}_nF_m(\alpha_1, \ldots, \alpha_n; \beta_1, \ldots, \beta_m; z) := \sum_{k=0}^{\infty} \frac{(\alpha_1)_k \cdots (\alpha_n)_k}{(\beta_1)_k \cdots (\beta_m)_k k!} z^k, $$

where $(\alpha)_k$ is the Pockhammer symbol defined earlier. When $p < q + 1$ these converge for any finite value of z, and hence define an entire function; when $p > q + 1$, the series diverges other than for $z = 0$. We could also ask the question for other generalizations, for example hypergeometric functions in multiple variables.

Bibliography

[1] Greg Anderson and Greg Moore. Rationality in conformal field theory. *Communications in Mathematical Physics*, 117(3):441–450, Sep 1988.

[2] A.O.L. Atkin and H.P.F. Swinnerton-Dyer. Modular forms on noncongruence subgroups. *Combinatorics (Proc. Sympos. Pure Math, Vol XIX, Univ. California, Los Angeles, Calif., 1968)*, pages 1–25, 1971.

[3] Jitendra Bajpai. *On Vector-Valued Modular Forms*. PhD thesis, University of Alberta, 2015.

[4] Peter Bantay and Terry Gannon. Vector-valued modular functions for the modular group and the hypergeometric equation. *Communications in Number Theory and Physics*, 1, 06 2007.

[5] F Beukers and G Heckman. Monodromy for the hypergeometric function ${}_nF_{n-1}$. *Inventiones Mathematicae*, 95:325–354, 1989.

[6] Gilles Christol. Fonctions hypergéométriques bornées. *Groupe d'Étude d'Analyse ultramétroqie*, 8:1–16, 1986.

[7] B. Dwork. On p-adic differential equations. iv generalized hypergeometric functions as p-adic analytic functions in one variable. *Annales scientifiques de l'École normale supérieure*, (4) 6:295–315, 1973.

[8] David Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, 2004.

[9] Harold Exton. *Multiple Hypergeometric Functions and Applications*. Ellis Horwood Limited, 1976.

[10] Charles F. Doran, Terry Gannon, Hossein Movasati, and Khosro Shokri. Automorphic forms for triangle groups. *Communications in Number Theory and Physics*, 7:689–737, 07 2013.

[11] Cameron Franc, Terry Gannon, and Geoffrey Mason. On unbounded denominators and hypergeometric series. *Journal of Number Theory*, in press.

[12] Cameron Franc and Geoffrey Mason. Fourier coefficients of vector-valued modular forms of dimension 2. *Canadian Math Bulletin 57*, pages 485–49, 2014.

[13] Cameron Franc and Geoffrey Mason. Three-dimensional imprimitive representations of the modular group and their associated modular forms. *Journal of Number Theory*, 160:186–214, 2016.

[14] Terry Gannon. *Moonshine Beyond the Monster*. Cambridge University Press, 2006.

[15] Terry Gannon. The theory of vector-valued modular forms for the modular group. *Conformal Field Theory, Automorphic Forms and Related Topics, W Kohnen and R WEissauer (eds)*, pages 247–286, 2014.

[16] Terry Gannon and Geoffrey Mason. Some low-dimensional congruence representations of $\Gamma(2)$. In Preparation.

[17] Fernando Q Gouvêa. *p-adic Numbers: An Introduction*. Universitext, Springer-Verlag, 2nd edition, 1997.

[18] Jeremy John Gray. *Differential Equations and Group Theory from Riemann to Poincaré*. PhD thesis, University of Warwick, 1981. Available http://go.warwick.ac.uk/wrap/35578.

[19] Einar Hille. *Ordinary Differential Equations in the Complex Domain*. John Wiley and Sons, 1976.

[20] Svetlana Katok. *Fuchsian Groups*. Chicago Lectures in Mathematics, 1992.

[21] Svetlana Katok. *p-adic Analysis Compared with Real*. AMS Universities Press, 2010.

[22] Neal Koblitz. *p-adic Numbers, p-adic Analysis,and Zeta-Functions*. Springer-Verlag, 2nd edition, 1984.

[23] Neal Koblitz. *Introduction to Elliptic Curves and Modular Forms*. Springer, 2nd edition, 1993.

[24] Serge Lang. *Algebra*. Springer, 3rd edition, 2002.

[25] A. M Legendre. *Thorie des Nombres*. Firmin Didot Frères, 1830.

[26] Jean-Pierre Serre. *A Course in Arithmetic*. Springer-Verlag, 1970.

[27] Jean-Pierre Serre. Formes modulaires et fonctions zêta p-adiques. *Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, 1972)*, pages 191–268. Lecture Notes in Math., Vol. 350, 1973.

[28] Winnie Li Wen-Ching and Ling Long. Fourier coefficients of non-congruence cusp forms. *Bull. Lond. Math. Soc.*, 44(3):591–598, 2012.

[29] Yongchang Zhu. Modular invariance of characters of vertex operator algebras. *Journal of the American Mathematical Society*, 9:237–302, 01 1996.