Kashiwa Lectures on New Approaches to the Monster

by John McKay[*],
edited and annotated by Yang-Hui He[†‡§¶]

Dedication. It is with a heavy heart that I put the finishing touches on this editorial. On April 19th, 2022, Professor John K. S. McKay passed away. Though his soul is recalled to that eternal, infinite, Platonic world of mathematics, his thoughts and words shall remain with us. Over the years, he has been a hero, a colleague, a friend, and above all, a grandfather, intellectually and emotionally. It has been a great honour to have explored mathematics with John and it is a fitting tribute to his life that this last work is in his own words, on one of his most influential discoveries.

In piam memoriam J. K. S. McKay.

Abstract. These notes stem from lectures given by the first author (JM) at the 2008 "Moonshine Conference in Kashiwa" (organized by the Institute for the Physics and Mathematics of the Universe (IPMU) under the support of the Graduate School of Mathematical Sciences, the University of Tokyo) and contain a number of new perspectives and observations on Monstrous Moonshine. Because many new points have not appeared anywhere in print, it is thought expedient to update, annotate and clarify them (as footnotes), an editorial task which the second author (YHH) is more than delighted to undertake. We hope the various puzzles and correspondences, delivered in a personal and casual manner, will serve as diversions intriguing to the community.

Contents

[*] CICMA & Department of Mathematics and Statistics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, H3G 1M8, Canada
E-mail: mckay@encs.concordia.ca
[†] London Institute for Mathematical Sciences, Royal Institution of Great Britain, 21 Albemarle Street, Mayfair, London W1S 4BS, UK
[‡] Merton College, University of Oxford, OX14JD, UK
E-mail: hey@maths.ox.ac.uk
[§] Department of Mathematics, City, University of London, EC1V 0HB, UK
[¶] School of Physics, NanKai University, Tianjin, 300071, P.R. China

1. Introduction

I am very honoured to be able to attend and participate in this conference which I would very much appreciate to be a relevantly informal business where some mathematics gets done. It is said that there's a very short time between being the youngest member of a conference, and becoming the oldest. I don't know quite about Harada-San, but I'm 68, so I'm probably there.

My background is computer science, basically. I started off computing character tables because of a remark by a professor that computing character tables was more of an art than a science, and I thought that should not be the case, and I was very fortunate in starting off in the 60s, at about the same time as the discovery of the modern sporadics. Janko's first group was discovered in 1964.

I'm sure you will have some delightful word that would express the contents of moonshine, which means something dubious, among other things, and moonshine, of course, is illicitly produced liquor.[1]

1.1 Resources

Books on Moonshine: There is a book by Mark Ronan, the very popular book [Ron], and there is a series of four fifteen minute talks by him on the BBC3 radio in Europe, at the beginning of this month. They have been recorded. If you want to know more about this recording I can tell you.[2]

Using the Web and Moonshine's Web Page: One of the things that I would like to emphasize is the use of the computer and the web today (Internet). I was giving a talk recently and several people come up to me and said: how do you manage to make these astonishing connections between things? Well, it's not that difficult, really. You have immense resources available on the web, and they grow all the time, and what happened to me, one of the connections I made, was after having got knocked down by a car, maybe again because I was thinking too much about problems. I was in bed for two months, and I just searched on the web for phrases in papers, and providing there are not more than a 100 or so papers you can actually go through these papers and see where they are relevant to your interests; a very effective and quite successful way of finding things.

If you really want to know how to get information if you are a graduate student, there is something in the literature: there's Bruce Reznick [Rez] who has written an article on extracting information, and there are some techniques which don't seem to be too well known.

We have had web pages in the past on moonshine groups. One of them was started up by Chi-Han Sah from Stony Brook, and we had quite a nice little group on that, but he died after an operation, and the whole business folded after that. Then Chris Cummins, who is a colleague of mine, put up a moonshine page (this was several years ago) with the latest papers and things, and that seems to have disappeared. So maybe the time is right to start up something again, where we can discuss things.[3]

1.2 Talk Outline

Now, I have three ideas which I think worth pursuing, to the extent that you can show that they are not worth pursuing if necessary, and I will talk about them and explain what little there is to be said about them.

I particularly want to emphasize a few things which are not as well known as they should be. One of them is the action of the Hecke operator and its connection with some very classical objects called *Faber polynomials*. Faber was, I believe, a numerical analyst (an analyst), and in 1903, in Mathematische Annalen, he wrote a paper on solving an approximation problem which was of interest to some fairly eminent people, including Hilbert, and that's how they started. But in fact that's not quite true. They go back to a man called Francesco Faà di Bruno, and he was an Italian. Probably the only Italian mathematical saint. He was beatified in 1988. He died in March 27, 1888.

So these polynomials are of some interest, and one can now look at them from a rather different

[1] The Kanji, or Chinese characters, for the word is 密造酒, literally meaning "secretly made alcohol". Of course, the word does not quite capture the sense of "madness" in English which Conway originally used to express the incredible nature of the Moonshine Conjectures. However, in classical Chinese poetry, numerous allusions are made to drinking accompanied by moon-light. The great poet Li Po (701–762) supposedly drowned himself, in his habitual state of inebriation, trying to grasp the reflections of moonlight in a lake. Thus perhaps 月光酒, or "moon-light liquor" is a more fitting translation.

[2] Since its incipience [Con-Nor] and proof [Bor] (of course, there remains many more things to be understood, including even Ogg's initial mysterious observation of the supersingular primes [Ogg] – q.v. recent accounts in [DO>, San>]), Moonshine has developed into a vast field. The reader is referred, for example, to Ronan's book and interview [Ron], du Sautoy's recent account [duS>], as well as nice technical progress reports of [Gan] and [DGO>]. In parallel, there has been much activity in the physics community extending Moonshine to beyond the Monster, with special focus on Mathieu 24 and its relation to the elliptic genus of K3 surfaces [EOT>, Che>, CDH>, GHV>, CDHKW>, HMR>, HeMcK>, DGO2>], to matrix models [HeJe], dessins d'enfants [HM>, THM>] and to exceptional Lie algebras [HeMcK3>].

[3] With the growth of blogs, especially in mathematics, there are several emergent sites which are useful [Blog>]. However, it would be useful to consolidate these resources, collect comments and have them maintained professionally; much in the spirit of the PolyMath projects [PolyMath>].

light. Using these polynomials we can define what we call *replicable functions*. This is a finite class of functions of about several hundred of them, amongst which the 171 functions which arise in the Monster's context as which we call monstrous moonshine today. These Faber polynomials describe the Hecke action, and that's part of the game.

I spoke to an eminent number theorist a year or two back and he told me that everything was over and we didn't need to think any more about moonshine and we understood everything about it. But that's very far from being the case, in my view. I think there are several things which are worth thinking about. One of them is Witten's idea that there might be some 24-dimensional manifold which would explain this moonshine by looking at the action of the Monster group on the free loop space of the manifold. I can't find much by Witten on this, but maybe he's written something. That would be a very nice goal to either establish the existence of the manifold and the Monster's action, or to show that such a thing does not exist. I think Borcherds, for example, doesn't think that an action on $\mathbb{M}$ exists, but I don't think you should necessarily take much notice of experts; my experience has been rather negative in that respect.

Then finally, as a sort of dream, it would be whether one could gather together all the finite simple groups. Let's initially see whether we can put the Monster within a better framework, presumably generalizing the Chevalley work in the Tohoku Journal in 1955 [Chev], and maybe one can pick up the Monster by generalizing, and it's conceivable that you might be able to pick up other groups; the other six pariahs (J_1, J_3, J_4, Lyons, O'Nan, Rudvalis) in this way. John Duncan has found some more moonshine attached to two of the pariahs.

Let me make another remark. I think that it's quite useful if one finds mathematical objects in other contexts, to find whether there is a connection between them, and I'll say a little bit more about that later. I'll give an example of it shortly.

1.3 Where to Start?

Galois 1832 As of a starting point of the talk, one can start with Galois, who died in 1832, and Galois' work in recognizing the notion of simplicity of a group, normal subgroups, and the other result that $\mathrm{PSL}_2(\mathrm{p})$ is realizable on the cosets of a subgroup of index p providing that p is not bigger than 11. So there are certain cases of that, which at least initially, were believed to be related to the Monster (see [Con-Nor]).

Mathieu 1861, 1873 One can start with Mathieu in 1861. He wrote a paper in 1861 [Mat] in which he said he had found five new groups as transitive extensions of classical linear groups.[4] In the 1861 paper he describes the smaller Mathieu groups M_{11} and M_{12}, and he says that others exist. Then in a paper 12 years later, he writes that his friends had a bit of trouble seeing how to construct his big groups, and so in 1873 he gives a description of the big groups M_{24}, M_{23} and M_{22}. Now, it's notable, and this is true throughout, that the Schur multiplier[5] of groups associated with the sporadic groups is larger than one would expect. The Schur multiplier of $\mathrm{PSL}_3(4)$ is an exceptionally big group; a group of order 48. Now, $\mathrm{PSL}_3(4)$ is a group of size $2^6 \cdot 3^2 \cdot 5 \cdot 7 = 20160$ and is the Mathieu group M_{21}. So M_{21} is *not* a sporadic group, but it is a classical group which starts the chain of sporadic groups, M_{21}, M_{22}, M_{23}, M_{24}.
I don't know how much skepticism there was when Mathieu wrote this, but there is a paper 25 years after the second Mathieu group by a man called G. A. Miller, who delighted in writing about the problems of other people's work, and his paper attempts to show that M_{24} doesn't exist. In 1900 he wrote a paper [Mil] in French correcting himself. So that's a history of the Mathieu groups, and we'll come back to the Mathieu groups later.

Janko 1964 We could start with the Janko groups. Janko was a hard worker. He did an enormous amount of work attempting to find sporadic groups, and he ended up with four groups, which are called J_1, J_2, J_3 and J_4. So he worked very hard, and the first successful outcome was in 1964. This is the start of modern era for the sporadic groups.

Plato 400 B.C. I could start with around the Plato's date, around 400 B.C., with the description of the Platonic solids (cf. part (a) of Figure 1). Why are we interested in them? There's this very curious bijection between the Platonic solids and their symmetry groups inside SU(2), and the A-, D- and E-type Lie structures [McK]. So that's the reason for that. Predating Plato there's an interesting guy called Empedocles.[6] I'll say perhaps a bit more at the end about him. He was interest-

[4] We recall that k-transitive means the following. Let G be a permutation group on n points and $\{a_1, a_2, \ldots, a_k\}$, $\{b_1, b_2, \ldots, b_k\}$ are two sets of points with a_i distinct and b_i distinct. If there is an element $g \in G$ mapping each a_i to b_i for $i = 1, \ldots, k$, then G is k-transitive. The only 4-transitive groups are the symmetric group $S_{k \geq 4}$, the alternating group $A_{k \geq 6}$, and the Mathieu groups M_{24}, M_{23}, M_{12} and M_{11} [Cam].

[5] We recall that for a finite group G, the Schur multiplier is the finite Abelian group whose exponent – the LCM of the order of all elements – divides $|G|$. More generally, the Schur multiplier of a group G is the second group homology $H_2(G;\mathbb{Z})$.

[6] Empedocles (circa 490–430 BC), pre-Socratic Greek philosopher, known as the originator of the cosmogenic theory of the four elements.

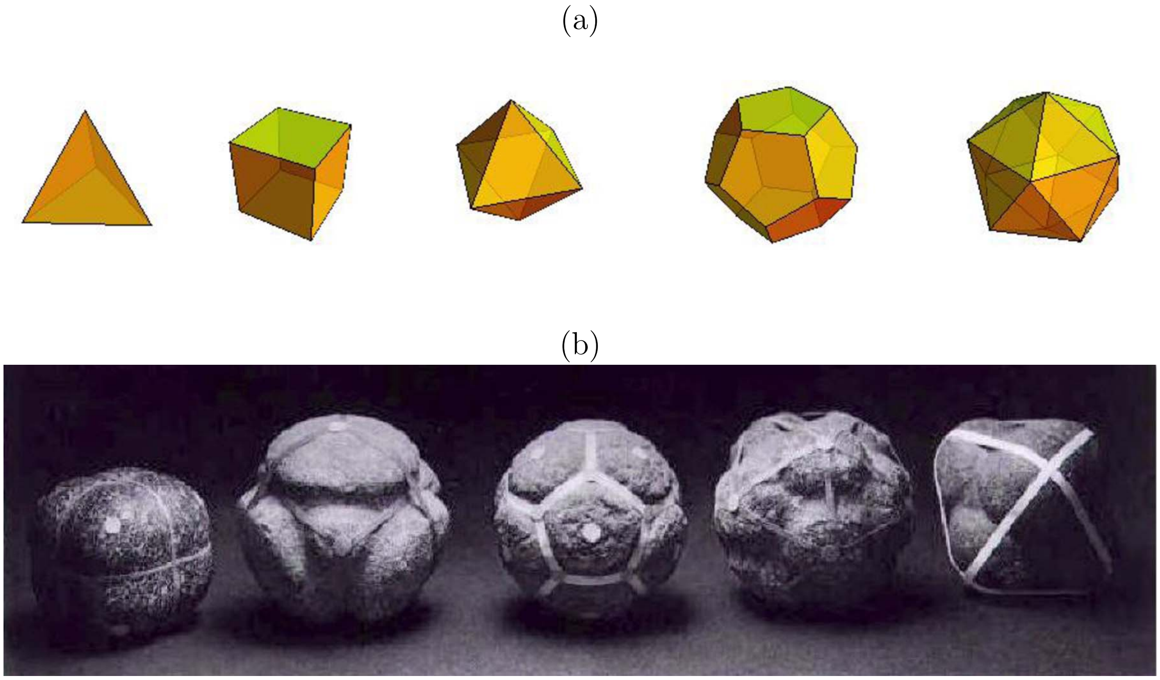

Figure 1. (a) The 5 Platonic solids, the tetrahedron (T), cubic (C), octahedron (O), dodecahedron (D) and icosahedron (I); C-O and D-I are graph duals and T is self-dual. (b) The Neolithic carved stones from Skara Brae, Scotland (in the Ashmolean Museum, Oxford), circa 3200BC.

ing because he forecast and predicted the finite speed of light, which I think was quite good for about 500 B.C.

Skara Brae 3000 B.C. And then there is something less well-known: Skara Brae. Skara Brae is a settlement in the Orkneys[7] which was discovered in about 1850. It's called *late neolithic* (that's a cultural date), but it is about 3200 B.C. This settlement, in a very isolated part of the world, contained some carved stones, and I'll show you some pictures of them (cf. part (b) of Figure 1). Now these stones are all about the same size, and nobody has any idea what they were for. The belief is they were not weapons because they're not damaged, and the possibility is that they gave the opportunity or permission to speak if you held one of these balls.I don't know whether the dodecahedron was above the cube or not, but anyhow there are these things around. And if you go to the Ashmolean Museum in Oxford,[8] they have them there. What was rather fun was that I mentioned this to Nigel Hitchin, the (emeritus) Savillian Professor of Geometry at Oxford, and is about 400 yards from these things, and he'd never heard of them.

2. Monstrous Moonshine

What I would like to do is make some remarks, and see where we get going from here. Conway and Norton's paper [Con-Nor] was published at the end of October 1979. And the story behind that you've probably all heard, Fischer was visiting me in Montreal, I wrote a letter to Thompson saying that one of the coefficients of the elliptic modular function j was 1 larger than the dimension of the smallest faithful representation of the Monster. Fischer took that back to Princeton. I think they all laughed at the concept of there being any connections, but there are. Borcherds has reminded me that what had happened was that I was reading a paper by Swinnerton-Dyer and Oliver Atkin [Atk-Swi], and in that paper they give the q-expansion for the j-function.

Oliver Atkin used to be a next door neighbour of mine in the ATLAS computer lab, and he is a number theorist. I was working on these big finite groups, like the Janko group whose order is $175,560$, and he was working on groups associated with two-by-two matrices. I was sure they would be much simpler things than I was working on, but I turned out to be wrong, in retrospect.

This is the order of the monster[9]

$$(2)\quad |\mathbb{M}| = 2^{46}.3^{20}.5^9.7^6.11^2.13^3.17.19.23.29.31.41.47.59.71\ .$$

[7] Orkney Islands, northern Scottland, GB.

[8] Ashmolean collection AN1927. 2727–2731, Oxford University, q.v. http://www.ashmolean.org/ash/britarch/highlights/stone-balls.html

[9] The Monster, largest of the 26 sporadic finite simple groups, is a 2-generated group, according to the ATLAS

There are 15 primes there. I don't quite know what name they should be given, but anyhow these are the *Monstrous primes*, or the *Monstrous supersingular primes*. We will return to these primes shortly. Thompson makes the remark. He says the order of a finite group is a very strong invariant. These primes appear elsewhere[10] in Erdenberger [Erd] (see later).

That is indeed true, and if you're trying to construct the groups, as we were in the early days, when these sporadic groups were sprouting so that every few weeks there would be a new one, one of the first things was getting hold of the order, and then using Sylow's Theorem to build up some structure, and perhaps guessing a subgroup, and then using that subgroup and the character table, then using characters building up the character table for the group, and announcing that, and then someone would come along and say the group doesn't exist because the character table doesn't satisfy some property or other. Then that property was eventually corrected. You had a correct table as far as you knew, and the question was trying to construct the group from the character table, and if that could be done, that was usually done by computer coset enumeration, and then using some technique to prove that the subgroup that you had made exists on the basis of the character table, did indeed exist.

2.1 Primes in the Monster's Order

It would be very useful to know more about these primes. I don't think there is so much that can be said about them, except for a remark that Ogg made, when he was attending at talk by Serre, I believe, at the Collège de France [Ogg]. This would be in the early 70's. One takes

$$\Gamma_0(p) = \left\{ \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod\ p \right\} \subset PSL(2;\mathbb{Z}) \ , \tag{3}$$

together with the Fricke involution $\alpha_p := \left(\begin{smallmatrix} 0 & 1 \\ -p & 0 \end{smallmatrix}\right)$. Consider the group

$$\Gamma_0(p)^+ = \langle \Gamma_0(p),\ \alpha_p \rangle \ , \tag{4}$$

and think of it as acting on the upper-half plane.[11] Then the genus of the Riemann surface $\Gamma_0(p)^+\backslash\mathbb{H}$ is zero precisely when p is one of the 15 supersingular primes that appear in the Monster's order. That's one number theoretic characterization of these primes.[12] We still don't know why, and Thompson regarded that as one of the major questions to be answered in connection with the Monster.

There is another way of saying it: for elliptic curves defined in characteristic p, then all the supersingular j-invariants of these curves (being a priori in $\mathbb{F}_{p^2}$) are lying in the base field $\mathbb{F}_p$, rather than in $\mathbb{F}_{p^2}$, precisely if p is one of the above 15 primes.

Now, rummaging through the contents of the preprints on the arXiv.org every weekday you look through and see if there's anything of interest. We found a paper by Cord Erdenberger [Erd], who is a student of Klaus Hulek from Hannover. And these 15 primes come up in his work. His title is "The Kodaira Dimension of Certain Moduli Spaces of Abelian Surfaces" (MR20923323 (2004)). He considers Abelian surfaces $(1,p)$ polarized (p a prime), and uses Jacobi cusp forms of weight 2 and level p, and these apparently exist just when p does not divide the order of the Monster. So in a sense, one might say that they are related to the existence of this Monster group. This is one connection which needs some explanation. Here you are working with a subgroup of the symplectic group rather than the modular group.

I've contacted Erdenberger and his supervisor Hulek, and nobody seems to know quite whether this is really saying something new, or whether it can be interpreted in terms of these supersingular elliptic curves that I mentioned earlier. It's something that should be pursued, at least try to find out whether there is a connection or not.

I would suggest if you want to follow this up, look in Math review for the Math Review number I have given above. The reason being that Sankaran reviewed it, and he does mention this in the review. You won't find any paper about it, and certainly Erdenberger was not aware of it. I don't know whether anyone is pursuing this; I don't know of any pursuit of this fact.[13]

[ATLAS]:

$$\begin{aligned} M = \langle a,b | a^2 = b^3 = (ab)^{29} = u^{50} = (au^{25})^5 \\ = (ab^2(b^2a)^5 b(ab)^5 b)^{34} = 1; \\ u := (ab)^4(abb)^2 \rangle \ . \end{aligned} \tag{1}$$

[10] There are marvelous recent expositions on how these 15 primes appear in 5 different contexts by Sankaran [San>], as well as how they can be explained from Moonshine [DO>].

[11] Indeed, a classical fact is that the upper-half plane, $\mathfrak{H} := \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$, when adjoining appropriate compactification points known as *cusps* which live in $\mathbb{Q} \cup \infty$, quotients the full modular group to give the Riemann sphere, of genus 0.

[12] Genus 0 congruence subgroups are very rare. For example, there are only 33 which are torsion free [Seb1] and the relation of these with elliptic surfaces, especially with K3 surfaces, is discussed in [Seb2, MS, HM>, HMR>].

[13] The reader is referred to the recent works of [DO>] and [San>] for various explanations.

2.2 Balance

This is typical of the sort of fact that you can gather, and one can formalize it as something (I don't know if it's a great thing to do so) maybe the words are "retro-syntactic retrieval", or something like that, but the game is very simple.[14] You have a bunch of people working on different subjects, and then if you study the phrases that are used in common by these people, or you find people who use a common phrase, there's a good chance that there is some related activity going on between the people that use this phrase. I didn't do that in this case. I think I was just looking through the arXiv.org and found it.

Now let me say something about balance. If we go back to the first paper, the word used is "seminal", certainly it was the only paper for a long time on the subject by Conway and Norton [Con-Nor]. Conway is here, and Simon Norton is not here. I don't think he should be forgotten. He is very often the motivating force between a lot of activity, some of which never gets published.

In this paper there is a list of observations which are introductory to the business of moonshine and one of them is that elements of the group M_{24} are *balanced*. So, what does this mean, and what is its significance? If you take a permutation in terms of disjoint cycle lengths you have a bunch of numbers which form a partition of the degree. A permutation is *balanced* if the product of the lengths of pairs from the outside-in is constant. Here is an example,[15] a permutation with cycle lengths 1, 2, 7 and 14 is balanced, since $N = 1.14 = 2.7 = 14$. The number N is called the *balance number*, if it exists.

In 1980 or thereabouts, we had a conference called "The Coming of Age of the Finite Groups", and some of the people were here then. Dummit and Kisilevsky and myself classified all permutations of degree 24 that are balanced [DKM]. Why choose 24? Well, we're going to replace k by $\eta(q^k)$, $q = e^{2\pi i\tau}$, for each cycle of length k, and thus form the product, we call it *eta products.*

We found all the η-products which are weakly multiplicative in the coefficients.[16] There are exactly 30 of them, and all the permutations in the Mathieu group M_{24} are balanced. And being balanced and weakly multiplicative is of the same thing as a theorem of Bryan Birch and Morris Newman on that [New]. And you have a cusp form for each balanced permutation, with what's called a *Grossen-character*, whose weight is half the number of parts.

Now we can generalize this to eta-quotients instead of eta-products by writing fractions. Here is an example: $2^{24}/1^{24}$ which means $\eta(2\tau)^{24}/\eta(\tau)^{24}$. The multiplicative η-quotients have been classified too[17] by Yves Martin [Mar]. These appear in [Apo].

The products are straight-forward because there are only finitely many partitions to look at, so you just go through them, find out the ones which look like they're multiplicative by looking at the first few coefficients and checking, and then filter them out and then you have to prove something. But for the quotients that's a different matter. The quotients are much more difficult. There is potentially infinitely many of them. The guy who had done a paper on them is Yves Martin. He hasn't completely done it. He made an assumption that both the eta-quotient and the eta-quotient with the above Fricke involution action on it are weakly multiplicative and that's not asked for. So, the general question about what eta-quotients are multiplicative is not known.

Being multiplicative means that you have some Euler product through the (inverse) Mellin transform,[18] and here is the Euler product (we use the subscript g to identify the particular eta-product):

$$\eta_g(s) = \prod_p \left(1 - \frac{a_p(g)}{p^s} + \frac{b_p(g)}{p^{2s}}\right)^{-1} \tag{5}$$

with a_p and b_p integers; we have

$$a_p^2 - a_{p^2} = b_p = \chi\left(\frac{\cdot}{p}\right) p^{w_g - 1} \ , \tag{6}$$

where χ is some character and w_g is the weight of the eta-quotient, which is half the number of parts. In particular, on the identity, the partition is 1^{24} and the weight is $w_g = 24/2 = 12$, which yields p^{11}. We have a degree p^{11} generalized character. These are proper characters which can be checked this directly. There is some anomalous behavior for $p = 3$, but other than that everything is clear.

[14] This was mentioned earlier in the introduction about how one could retrieve information and establish correspondences.

[15] Thus, an element of the permutation group S_{24} of degree 24, would have cycle notation $(a_1)(a_2,a_3)(a_4,\ldots,a_{10})(a_{11},\ldots,a_{24})$, which is indeed the shape of one of the 1575 conjugacy classes of S_{24} (1575 is the number of unrestricted partitions of 24).

[16] The simplest case is the famous $\Delta(q) = \eta(q)^{24}$, which is the modular discriminant function, with q-expansion $\Delta(q) = \sum_{n=1}^{\infty} \tau(n)q^n$ with $\tau(n)$ being the Ramanujan tau-function. This is weakly multiplicative in the sense that $\tau(m\ n) = \tau(m)\tau(n)$ if $\gcd(m,n) = 1$. The multiplicative eta-products appear in physics, especially in partition functions in string theory and are discussed in [GK>, Che>, HeMcK>].

[17] Cf. also [Kil, MarOno] for relations to elliptic curves.

[18] That is, we can form the Dirichlet series for the coefficients c_n of the q-expansion $\sum_n c_n q^n$ for these eta-products to give $\eta_g(s) = \sum_n c_n n^{-s}$. This can then be taken as product over primes as in (5).

2.3 Conjugacy Classes

There are curious connections with physics and conjugacy classes, and I think someone raised this the other day at the EWM meeting,[19] that in studying elliptic genera you look at commuting pairs of elements, and that's related to the number of conjugacy classes in the group when working with a finite group. And I just make the passing remark, the class number – the number of conjugacy classes- of M_{24} is 26, which should ring a bell with some physicists.[20] You have 5 quadratic boxes of irrationalities, so there are 21 classes of cyclic subgroups, or if you like, 21 rationally irreducible representations. The group M_{12}, believe it or not, has 10 instead of 26.

If you want to know more about this remark about conjugacy classes, there is recent publication, a large book in fact called "From Number Theory to Physics", in Les Houches proceedings of 2002, containing a paper with Sebbar and myself [McK-Seb]. More recently, in a paper in a conference proceedings on things to do with moonshine, there's a paper by Anda Degeratu and Katrin Wendland [Deg-Wen>], and they look again at one of these situations that you discover by reading and saying "my goodness, it's the same number there." They are looking at a situation where the appropriate number replacing 26 is 194, and 194 is the number of conjugacy classes in the Monster.[21] So they are looking at a situation that is interesting if anything comes out of it.

2.4 Frame Shape

Now, these shapes we've been talking about – the partition of n where n is the degree, or the more general situation when you divide one term by another is called *Frame shapes*. Now "Frame" is the name of a person, J. S. Frame, not an abstract notion, and he was an interesting man. His thesis was on character tables in the 1930's [Fra]. He did character tables like other people do crosswords. So that's who Frame was.

Basically speaking, what you are doing is that you are describing the eigenvalues within an orthogonal group. So if you have a fraction, you want to make sure that the eigenvalues that you take away from the denominator are already in the numerator, to make any sense. There's a paper by Takeshi Kondo [Kon], who wrote a very nice paper about the Frame shapes of elements in the automorphism group of the Leech lattice, $\mathrm{Aut}(L)$. If you go down from there to Conway's group $\cdot 1$, which is this group modulo the action on diameters – i.e., $\mathrm{Aut}(L)/\langle \pm 1\rangle$ – you get a mixture of the functions that describe the elements on the various classes and they are not as consistent as those for the monster $\mathbb{M}$.

By the way, there's a very nice survey by Masao Koike [Koi], I will say more about him later on. Again this was in Sugaku, in Japanese that has been translated into English by the AMS translations #160. This is one of the few early surveys on moonshine, so that's a useful paper too.

2.5 Faber Polynomials

If we take a Riemann map[22] from the exterior of some region in the complex plane containing two points at least, by the Riemann mapping theorem we can map the exterior of this region with some conditions at infinity to the exterior of a disk of radius d, and we can normalize this

$$z = \phi(w) = dw + d_0 + \sum_{k\geq 1} d_k w^{-k} \ . \tag{7}$$

If you are doing analysis you don't have to worry about the constant term and if you are doing moonshine you put it equal to 0 and the radius equal to 1. We get an inverse of the same forms as (7)

$$w = \phi^{-1}(z) = z/d + g_0 + \sum_{k\geq 1} g_k z^{-k}, \tag{8}$$

and similarly we can take $d = 1$ and $g_0 = 0$. The *Faber polynomial* F_n is the part of $(\phi^{-1}(z))^n$ with non-negative powers of z. So you're picking up the polynomial part of this series $(\phi^{-1}(z))^n$.

You may not be familiar with this, but if you're looking at pseudo-differential operators that's a standard procedure to pick the plus part of the operator. That's what the Faber polynomial does for you, and that how it's used and I'll say quite a lot more about that.[23]

[19] Encounters with Mathematics, Chuo University, May, 2008, http://www.math.chuo-u.ac.jp/ENCwMATH/45.shtml.

[20] In string theory, the critical space-time dimension of the bosonic string is 26 and that of type II super-string is 10.

[21] In other words, as remarked in [McK-Seb], the number of conjugacy classes of M_{12} and M_{24} are respectively 10 and 26, the critical dimension of the supersymmetric and the bosonic string theories. Moreover, 194, the number of conjugacy classes of the Monster, is the Picard number of the base of an elliptically fibred Calabi-Yau threefold in an extremal case of heterotic-F-theory duality as studied in [AKM]. Furthermore, of these 194 classes, considered as column-vectors in the character table, only 163 are linearly independent; and of course, 163 is a famous Heegner number where the exponential assumes an almost-integer value: $\exp(\pi\sqrt{163}) \sim 640320^3 + 744$.

[22] This approach of looking at Moonshine from the perspective of geometric function theory, in terms of the shape of the analytic functional form of $j(q)$ and generalizations, is very much the spirit of the current lecture notes, and is also summarized in [McK2].

[23] In other words, we consider a meromorphic function and its inverse with Laurent expansion of the form (7) and (9).

What we do is we compose with the map $z \mapsto 1/q = e^{-2\pi i z}$, to get

$$f(z) = q^{-1} + \sum_{k\geq 1} a_k q^k, \qquad \Im(z) > 0 \ . \tag{9}$$

The a_k's are general coefficients and we'll take them to be integers, but they need not be integers, and Simon Norton has classified the functions we are interested in (the replicable ones to come later) even when they have complex coefficients. If these coefficients are not integers, they do lie in a field whose Galois group is an elementary 2-group over the rationals. In other words, the a_k's, lie in a composite of quadratic fields. I don't think we really know which quadratic fields and why, but anyhow that's where they lie when we're talking about replicable functions, that's the ones we are interested in.

But for our purposes and for all the stuff here we work with the a_k's being integers. That avoids any problems with Galois theory and is convenient. The functions of the form (9) are typical functions we shall study, and we are going to study them first of all slightly more generally than the connection with the Monster, and then specialize to functions that are attached to $\mathbb{M}$.

2.6 Grunsky Coefficients

We can define the elliptic modular function j by the above property, because, if this holds for some other function $f(z)$ for all positive integers n, the level of this function $f(z)$ must be equal to 1, and is therefore a rational function of $j(z)$, and so, using the fact that j is normalized at infinity as in (9), we have that $f(z) = j(z)$.

For f as in (9) we write

$$F_n(f) = q^{-n} + n \sum_{m\geq 1} h_{m,n}(f) q^m \ , \tag{10}$$

the coefficients $h_{m,n}$ are called *Grunsky coefficients*, see [Grun, Pomm]. These are in fact symmetric in the m,n indices.[24] They have a remarkable connection with the Bieberbach conjecture.[25] The Bieberbach conjecture is a bound on the coefficients of functions univalent on the unit disk, and there's a very nice book [de Bra], an AMS publication, on the solution to the Bieberbach conjecture by de Branges, and these Grunsky coefficients played the major role in the establishment of this conjecture in its early days.[26]

In our definition (10), we have written $F_n(f)$ in this way with an n in front in order to take advantage of the symmetry of $h_{m,n}$ in m and n. In fact that is the definition of the Grunsky coefficients

$$h_{m,n}(f) = T_n(f)|_{q^m} \ . \tag{13}$$

There will be other introductions to the Faber polynomials and these Hecke operators T_n later.[27]

It's not obvious from the above expression that $h_{m,n}$ is symmetric in m and n, but it is, and we can see

This is clearly inspired by the form of the q-expansion of the j-invariant, as we shall shortly see. We emphasize that the constant term is 0, so henceforth, by the j-invariant, we mean the normalized one $j(q) - 744$. In [HeJe], this shape of a Laurent series was interpreted as the master-field of a large N matrix model, whereby giving a modular matrix model.

[24] We can in fact define the Grunsky coefficients and Faber polynomials in the following way. We will encounter some of the ensuing expressions in due course. Let $g(z)$ be a holomorphic univalent (i.e., one-to-one on the open set) function on the unit disk $|z| < 1$, normalized so that $g(0) = 0, \ g'(0) = 1$. Then the function $f(z) = g(1/z)^{-1}$ is a non-vanishing univalent function outside the unit disk with simple pole at ∞ with residue 1; that is,

$$f(z) = z + a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots \ .$$

The expansion coefficients c_{nm} of

$$\log \frac{f(z) - f(w)}{z - w} = - \sum_{m,n>0} c_{nm} z^{-m} w^{-n} \tag{11}$$

are the *Grunsky coefficients*. Definition (11) implies, upon $z\frac{\partial}{\partial z}$ on both sides, that $\frac{zg'(z)}{f(z)-f(w)} - \frac{z}{z-w} = \sum\limits_{m,n>0} m c_{nm} z^{-m} w^{-n}$. Thus we define the *Faber polynomials* $F_n(w)$, as

$$\frac{zg'(z)}{f(z) - w} = \sum_{n\geq 0} F_n(w) z^{-n} \ . \tag{12}$$

It is non-trivial that, thus defined, $F_n(w)$ are monic polynomials of degree n. In fact, F_n are themselves polynomials in the coefficients a_i in the definition of $f(z)$. This is seen as follows. Definition (12) implies, upon applying $\int_0^\infty dz \frac{1}{z}$ on both sides, that $\log \frac{f(z)-w}{z} = - \sum\limits_{n\geq 1} \frac{1}{n} F_n(w) z^{-n}$. Expanding out $f(z)$ order by order and comparing with (12) then gives the recursion

$$F_n(0) = 1 \ , \quad F_n(w) = (w - a_0) F_{n-1}(w) - n a_n - \sum_{i=0}^{n-1} a_{n-i} F_i(w) \ .$$

Furthermore, combining (12) and (11) we have that $\sum\limits_{n\geq 0} F_n(g(\zeta)) z^{-n} = \frac{z}{z-\zeta} + \sum\limits_{m,n>0} m c_{nm} z^{-m} \zeta^{-n} = \sum\limits_{n>0} \left(\frac{w}{z}\right)^n + \sum\limits_{m,n>0} m c_{nm} z^{-m} \zeta^{-n}$ so that

$$F_n(g(z)) = z^n + \sum_{m\geq 1} c_{nm} z^{-m} \ ,$$

which is (10) in our definition, up to the factor of n which will be more convenient for our succeeding discussions.

[25] We recall the statement of the Bieberbach Conjecture, proven by de Branges. For univalent holomorphic function with Taylor series of the form $f(z) = z + \sum\limits_{n\geq 2} a_n z^n$ (such functions are called *Schlicht*, or simple/plain), the coefficients have the property that $|a_n| \leq n$ for all $n \geq 2$.

[26] There is a very recent paper on the appearance and relevance of Bieberbach/de Brange as well as Grunsky coefficients in scattering amplitudes in quantum field theories [HSZ$^>$].

[27] In [McK-Seb], a particularly nice characterization of the Faber polynomial is as given. Consider, as always, a function $f(q) = q^{-1} + \sum\limits_{n\geq 1} a_n q^n$ for nome $q = \exp(2\pi i z)$ with $\mathrm{Im}(z) > 0$, as in (9). Then, for each $n \in \mathbb{Z}_{>0}$, there is a *unique* monic poly-

it with a slightly different generating function for the Grunsky coefficients.

3. Hecke Operators and Faber Polynomials

I now turn to Part 2 of my lectures, having alluded to the Hecke operators. There are Hecke operators, and often in the books they assume that the function on which the Hecke operator is acting has a weight greater than zero, whereas the functions we are interested in all have weight zero. You have a group action acting linear fractionally on z, and on the modular function j this is given by, for all $n \geq 1$

$$T_n(j(z)) = \frac{1}{n} \sum_{\substack{ad=n\\ 0\leq b<d}} j\left(\frac{az+b}{d}\right) = \frac{1}{n} F_n(j(z)) \ . \tag{14}$$

The effect of the Hecke action is to replace the pole of order 1 of the j-function with a pole of order n at infinity. Moreover, the action of the Hecke operator preserves the space of modular functions. Hence, $T_n(j(z))$ is a rational function of $j(z)$, and so this going to be a polynomial in $j(z)$. $T_n(j(z))$ can be expressed both as a q-series and as a polynomial in $j(z)$. In fact, we can define the j-function by the fact that there is an action of the Hecke operator defined in terms of sum over the function valued on sublattices. Let's have a quick look at this, which is standard.

In part (a) of Figure 2, this is the fundamental region, up to orientation and homothety. The Hecke operator T_n maps the lattice $\Lambda = \mathbb{Z} + z.\mathbb{Z}$, $z = \omega_1/\omega_2$ such that $\Im z > 0$, to sublattices $\{\Lambda_i\}$ of index n, so induces an action on functions defined on these lattices.[28] For example, for $n = 2$, there are 3 lattices of fundamental region twice that of Λ and the Hecke action is given by

$$T_2 : \ f(z) \mapsto \frac{1}{2}\left(f(2z) + f\left(\frac{z}{2}\right) + f\left(\frac{z+1}{2}\right)\right) \ . \tag{15}$$

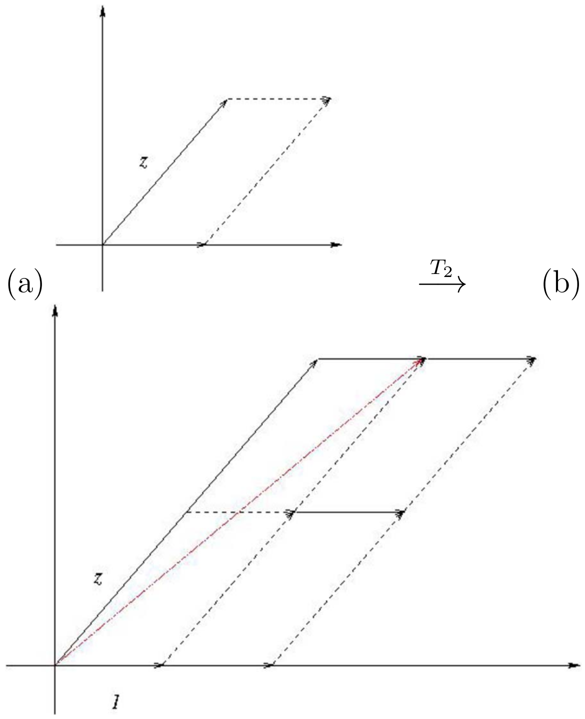

Figure 2. (a) Fundamental region of the lattice $\Lambda = \mathbb{Z} + z\mathbb{Z}$. (b) The 3 sublattices of index $n = 2$, viz., $\mathbb{Z} + 2z\mathbb{Z}$, $\mathbb{Z} + (1+2z)\mathbb{Z}$ and $2\mathbb{Z} + z\mathbb{Z}$.

This is drawn in part (b) of Figure 2. I've used z here because τ is used later.

For each of the functions we are interested in there will be a discrete subgroup G_f of $\mathrm{PSL}_2(\mathbb{R})$ with respect to which f is modular, and in [Con-Nor] there is a discussion on how much is being fixed. The action of G_f is linear fractional. For us, the transformation as an element of G_f

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \tag{16}$$

is not affected by the automorphy factor $(cz+d)^k$ because we are working with functions of *weight* $k = 0$, and this is rather important since the behavior is different.[29] There is also the notion of the *level*, which is the smallest N such that $\Gamma(N) \subset G_f$.[30]

Classically, for all $\gcd(n,N) = 1$, $T_n(f)$ is a polynomial in f (f as in (9)). This last statement is the criti-

nomial F_n such that

$$F_n(f(q)) = q^{-n} + \mathcal{O}(q) \ , \quad \text{as } q \to 0 \ .$$

These are the Faber polynomials. Depending on the Taylor series of $f(q)$, the first few are

$$F_0(z) = 1, \ F_1(z) = z, \ F_2(z) = z^2 - 2a_1, \ F_3(z) = z^3 - 3a_1 z - 3a_2 \ .$$

More generally, we have

$$F_n(z) = \det(z\mathbb{I} - A_n) \ , \qquad A_n := \begin{pmatrix} a_0 & 1 & & & & \\ 2a_1 & a_0 & 1 & & & \\ \vdots & \vdots & \vdots & & & \\ (n-2)a_{n-3} & a_{n-4} & a_{n-5} & \dots & 1 & \\ (n-1)a_{n-2} & a_{n-3} & a_{n-4} & \dots & a_1 & 1 \\ na_{n-1} & a_{n-2} & a_{n-3} & \dots & a_1 & a_0 \end{pmatrix}$$

[28] The fundamental fact here is that sublattices of index n are in one-one correspondence with integer matrices $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$ with $a > 0$, $b = 0, 1, \dots, d-1$ and $ad = n$. For example, at $n = 2$, we have 3 such matrices, $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, corresponding to the lattices $\mathbb{Z} + 2z\mathbb{Z}$, $\mathbb{Z} + (1+2z)\mathbb{Z}$ and $2\mathbb{Z} + z\mathbb{Z}$.

[29] That is, the j-function is an absolute invariant $j\left(\frac{az+b}{cz+d}\right) = j(z)$ for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in PSL(2;\mathbb{Z})$. Indeed, for weight k objects, the standard definition [Ser] of the Hecke operator is

$$T_n(f(z)) = n^{k-1} \sum_{\substack{ad=n\\ 0\leq b<d}} d^{-k} f\left(\frac{az+b}{d}\right) \ ,$$

for all $n \in \mathbb{Z}_{\geq 1}$.

[30] The congruence groups are defined with some modulo N relation. For example, the principal congruence subgroup is $\Gamma(N) := \{A \in PSL(2;\mathbb{Z}) \mid A \equiv I \bmod N\}$.

cal one that enable us to generalize the action of the Hecke operator, and the generality is that if one defines the j-function with the normal Hecke operator acting as a polynomial then what we do is we preserve the action of the (Faber) polynomial and redefine the Hecke operator. Indeed, since for all $n \geq 1$ we have

$$T_n(f(z)) = \frac{1}{n} \sum_{\substack{ad=n \\ 0\leq b<d}} f\left(\frac{az+b}{d}\right) = \frac{1}{n} F_n(f(z)) \ , \tag{17}$$

the modular level is one, and therefore is a rational function of j. Hence, using (9) where f is normalized to have a simple pole at infinity and a zero constant term, implies that $T_n(f(z))$ is a polynomial in $j(z)$, and thus $f = j$. This enables us to define the j-function from the property that the Hecke operator on j averaging over sublattices of index n is the Faber polynomial associated with j.

I want to say quite a bit about this important polynomial. It's easy to find this polynomial from a computational and algorithmic points of view. Start off with (9) and look at various powers of f you'll get various negative powers of q on the right hand side of (17), and by forming linear combinations of these powers of f on the left hand side of (17) we can eliminate all but the largest negative power of q, and if there is a constant term we can put it into the polynomial on the left hand side. That makes it clear that this polynomial is very simply determined from (9) and is unique. What this polynomial is doing is replacing the simple pole at infinity in (9) by an order n pole also at infinity in (17). And if we write, for f as in (9),

$$F_n(f) = q^{-n} + n \sum_{m\geq 1} h_{m,n}(f) q^m, \tag{18}$$

the coefficients $h_{m,n}$ are the *Grunsky coefficients*. They have a lot of interesting properties too.[31]

3.1 Replicable Functions: Norton's Basis

We can define replicable functions $\{f^{(m)}\}$, say, by the same q-expansion as in (9)

$$f^{(m)}(z) = \frac{1}{q} + \sum_{k\geq 1} h_k^{(m)} q^k \ , \tag{19}$$

I am using the superscript (m) here, and I was talking to John Conway in the breakfast about this. I think the notation has to be changed. I've been using small letters for coefficients and character, and I've used a little h_k here instead of a_k. You can use a_k if you wanted to. But the I use the h_k to remind you of the characters of the Monster in the special case when you restrict to the Monster. We can define a collection of functions here by, for all $n \geq 1$

$$\widehat{T}_n(f) = \frac{1}{n} \sum_{\substack{ad=n \\ 0\leq b<d}} f^{(a)}\left(\frac{az+b}{d}\right) = \frac{1}{n} F_n(f) \ . \tag{20}$$

What we've done is that we kept the Faber action in the above and replaced the sum over representatives of the sublattices of the function f by the functions $f^{(a)}$. Those of you who want a glimpse of the future, I can give you the relation between f and $f^{(a)}$ with reference to the Monster.

To an element of the Monster g there corresponds a function f_g, corresponding for the moonshine for this function on the element g, then raising f to the a-th replicate power $f^{(a)}$ is the same as replacing g by g^a. You will get all this in a short time. There is no need to make reference to the Monster in order to define this.[33]

There is an inductive definition, and there is only going to be one term, which is $f^{(n)}$, in the sum in (20), and so you can take out the rest of this sum and put it on the right hand side and that with will involve only $f^{(a)}$ with $a < n$ and you have an inductive definition of these functions. I'll say more about these Hecke operators.

Now, Norton did all this unaware that it had all been done 70 years before and before that. Perhaps the easy definition of the Grunsky coefficients $h_{m,n}$ is in terms of this generating function

$$\ln\left(\frac{f(p)-f(q)}{\frac{1}{p}-\frac{1}{q}}\right) = -\sum_{m,n\geq 1} h_{m,n} p^n q^m \tag{21}$$

with $p = e^{2\pi i s}$, $q = e^{2\pi i z}$. You take the logarithm of the difference of the function evaluated at two different arguments p and q. Remember that $f(p)$ (resp. $f(q)$) will start with $1/p$ (resp. $1/q$), so dividing through by

[31] Let us re-iterate this point. We saw in the footnotes above that the Faber polynomials are the unique degree n monic polynomials bringing $q^{-1} + \mathcal{O}(q)$ to $q^{-n} + \mathcal{O}(q)$ as $q \to 0$. Now, our Hecke operator on j of weight 0 as in (17), thus all $T_n(j(z))$ are invariant under $PSL(2;\mathbb{Z})$ since the sum gets permutated by the action of the modular group. Whence they must be rational functions in $j(z)$ since the j-invariant, being a Hauptmodul[32], generates the function field of invariants. However, since it has no poles in the upper half plane, they must in fact be polynomials. In fact, we find that $T(j(q)) = q^{-n} + \mathcal{O}(q)$ as $q \to 0$. By uniqueness then, $T_n(j)$ must be (up to overall normalization) the Faber polynomials!

[33] Historically, the concept of replicability came about from Conway-Norton's initial observation [Con-Nor] that the moonshine functions (McKay-Thompson series) obeyed certain functional identities, which they called replication. This is the reason for studying the type of recurrences in §3.5. The motivation in defining it in the manner of the present section is to generalize the remarkable fact that action of the n-th Hecke operator on j is the n-th Faber polynomial in j. Thus a function of the expansion type (9) is *replicable* if there exists a family of function $\{f^{(a)}\}$, called replicable functions of f such that the generalized Hecke operator on these $\frac{1}{n} \sum_{\substack{ad=n \\ 0\leq b<d}} f^{(a)}\left(\frac{az+b}{d}\right)$ is the Faber polynomial in f, i.e., $\frac{1}{n}F_n(f)$.

$1/p-1/q$ gets rid of the singularities and you end up with a quite nice series, namely the series on the right hand side of (21). That's the generating function for the $h_{m,n}$ and you can expand that to have

$$\text{(22)}\qquad Eq.\ (21) = \ln\left(1 - pq\sum_{k\geq 1} h_k \frac{p^k-q^k}{p-q}\right)$$

and there are certain consequences of the expansion one of them is that the $h_{m,n}$'s are polynomial in the h_k's with $k \leq m+n$; the value $m+n$ provides you with the grading. I'm using the convention $h_k = a_k = h_{k,1}$. When we compare coefficients in the last expansion with the $\ln$ term, we get this recursive expression
(23)

$$h_{r,s} = h_{r+s-1} + \frac{1}{r+s}\sum_{m=1}^{r-1}\sum_{n=1}^{s-1} h_{m+n-1}(r+s-m-n)h_{r-m,s-n}.$$

Indeed, in (22), the terms $\frac{p^k-q^k}{p-q}$ are part of p^sq^t where $s+t=k-1$, but you are multiplying by pq so you get a term involving the expression involving the h_k's and the $h_{m,n}$ corresponding to the appropriate exponent. If we call $r+s$ the grade then all $h_{r-m,s-n}$ in (23) have a lower grade. The $h_{r,s}$ are not integers, but the largest denominator is (r,s). I believe that's correct in a sense. Suppose we want to compute a q-coefficient of the function, then that is typically given by the term h_{r+s-1}. So we have a choice: if we want to compute h_k then we can choose $r+s=k+1$ so that $k=r+s-1$, and we can do that what will give us the initial step, and the game is to try to find a pair which has a reduced sum.

Simon Norton[34] has a definition of a replicable function which is that a function is replicable if $h_{m,n} = h_{m',n'}$ whenever $mn = m'n'$ and $\gcd(r,s) = \gcd(m,n)$. This is an important definition and we can take advantage of this in computing the coefficients of a replicable function from (23).

The following picture shows how to compute the coefficient of q^k. With k fixed this gives us a choice of $r,\ s$, so we can draw the line $x+y=r+s=k+1$, and then the game is to find a point (r,s) on that line that dominates, if we are lucky, some other point (r',s') with the same $\gcd$ and lcm. We then have the hyperbola $xy=rs$. So if we can find this (r',s') on a lower line then we start with, we can go to this line and proceed to do the same thing again, and each time we do that there are two possibilities – there is or there is not a line below it. If the point exists we carry on. If it does not exist we mark the parameter for the line. So here there is a bunch of lines here, and it turns out – this is Norton's straight theorem – that there are 12 values for which you can't reduce them further. These are

$$23,\ 19,\ 17,\ 11,\ 9,\ 8,\ 7,\ 5,\ 4,\ 3,\ 2,\ 1\ .$$

These 12 values are the values of k so that every coefficient of a replicable function is a polynomial in these 12 values of h_k. This is called *Norton basis*, and that's a very fundamental result. Now, Conway was talking yesterday about the work of Atkin, Fong and Smith [AFS], and by the way, Borcherds. Well, if Atkin, Fong and Smith had got this theorem at the time they did their computations they would have only needed at worst 24 coefficients to establish the result of the moonshine conjecture and these modular functions. But they didn't have it at the time and so that wasn't accessible to them. This is the general version, and this provides a basis for all replicable functions.

I want to say a bit more about this, something special, in a minute. If the function has odd level, that means for the moonshine functions and that means that for the conjugacy class containing the element has odd order, you don't need more than these few at the bottom, in fact, $1,\ 2,\ 3,\ 5$ are sufficient to do the thing. There is something [CohMck] rather special when you have odd level, but in general you need all the above 12 elements. I'll show a quite neat proof of the theorem in a minute. This is a restatement of the condition of replicability by Norton saying that

$$\text{(25)}\qquad h_{m,n} = h_{\mathrm{lcm(m,\ n),gcd(m,n)}}\ .$$

Think of the Smith normal form of a 2×2 matrix perhaps. You have a matrix with m and n on the diagonal and it's equivalent to a matrix with $\gcd(m,n)$ and $\operatorname{lcm}(\mathrm{m,\ n})$ on the diagonal. If you check the coefficients $h_i^{(k)}$, where $f^{(k)}(z) = \sum h_i^{(k)} q^i$, you can see that

$$\text{(26)}\qquad h_i^{(k)} = k\sum_{d|k}\mu(d)h_{k/d,dki}, \quad i>0,\ h_{-1}^{(k)}=1,\ h_0^{(k)}=0\ .$$

From this, inverting the above, you deduce for all $r,s\in\mathbb{N}$

$$\text{(27)}\qquad h_{r,rs} = \sum_{d|r}\frac{1}{d}h^{(d)}_{r^2s/d^2}\ .$$

[34] Following [McK-Seb], we can proceed with this formal definition of a replicable function. Consider a function the form (9), and write its corresponding Faber polynomial, with Grunsky coefficients $h_{m,n}$ as in (10). Then f is replicable if $h_{m,n}=h_{r,s}$ whenever $\gcd(m,n)=\gcd(r,s)$ and $\operatorname{lcm}(m,n)=\operatorname{lcm}(r,s)$. Equivalently, we can define replicable functions using the Hecke operators (for weight 0). The function of our form (9) is replicable, if for each positive integer n and positive divisor $a|n$, there are functions $f^{(a)}$ of the form of (9) such that $\widehat{T}_n(f) := F_n(f(q)) = \sum_{\substack{ad=n\\ 0\leq b<d}} f^{(a)}\left(\frac{az+b}{d}\right)$. The functions $f^{(a)}$ are called replication powers and have the property that

$$\text{(24)}\qquad f^{(k)}(q) = q^{-1} + \sum_{i\geq 1}\left(k\sum_{d|k}\mu(d)h_{dki,\frac{k}{d}}\right)q^i\ ,$$

where μ is the standard Möbius μ-function.

This can be rewritten as

$$h_{m,n} = \sum_{d|(m,n)} \frac{1}{d} h^{(d)}_{mn/d^2} \,. \tag{28}$$

This is the final result.

Now those of you who have read the useful book *A course in Arithmetic* by Serre [Ser], the second half of the book is devoted to things of interest to us (modular forms), and mentions the Leech lattice and various things to do with theta functions, you will find a formula very like this without the superscript (d), and if you follow Serre's proofs they go through pretty well word for word in this more general situation.

I'd like to just show you the proof of the Norton basis theorem [Nor] done by Cummins [Cum] because it's very neat, and maybe other versions around are not as neat as this.[35]

Theorem 3.0 (Norton Basis Theorem). *The q-coefficients of a replicable function are polynomials in h_k, $k \in B = \{1,\ 2,\ 3,\ 4,\ 5,\ 7,\ 8,\ 9,\ 11,\ 17,\ 19,\ 23\}$.*

So what we want to do is to prove that for everything that is not in the basis we can actually reduce the sum $m+n$. So you need to prove that there exists $m,\ n,\ m',\ n' \in \mathbb{N}$ such that

1. $m+n = N$, where N is given;
2. $\operatorname{lcm}(m,n) = \operatorname{lcm}(m',\ n')$;
3. $\gcd(m,n) = \gcd(m',n')$;
4. $m'+n' < m+n$.

Proof. The first remark is that if we find some result which is true for any number N then the result can be true for kN. That's a useful thing to look at, and that means that we don't need to have common factors in the subscripts. Here are the cases to go through one by one:

i. $N = 2^k$ not 2, 4 or 8; we can always look at what we believe to be the basis and see what we need not worry about. For $N = 16$, here is a pair $m = 1,\ n = 15,\ m' = 3,\ n' = 5$. All else is a power of $2 = 16t$.
ii. For N odd and $2^a + 1,\ a \geq 4,\ N \geq 17$, we have $m = 2^a - 2,\ n = 3,\ m' = 2^{a-1} - 1,\ n' = 6$.
iii. N odd $N \neq 1 + 2^k$, you subtract 1 from it to get $m = N-1,\ n = 1,\ m' 2^{-r}(N-1),\ n' = 2^r$ with $2^r | (N-1)$ is a whole divisor of $N-1$. You can follow that through $N - 1 > 2^r \ \Rightarrow\ N > 2^r + 1 \ \Rightarrow\ N(2^r - 1) > 2^{2r} - 1 \ \Rightarrow\ N > 2^{-r}(N-1) + 2^r$.
iv. Then N even not a power of 2. In this situation N is a going to be a product of 2, 4 or 8 with 3, 5 or 9. And these cases we look at individually. For 40 take $m = 1,\ n = 39,\ m' = 3,\ n' = 13$. For 36, $m = 1,\ n = 35,\ m' = 5,\ n' = 7$ also works for 72.

This reduces all the ones that can be reduced, and what you have left over is in the Norton basis, and this proves what the Norton basis actually is. □

3.2 Elastica

Now something of interest that maybe someone throw some light on here. If you write down the Norton basis in this way

	1	2	3	4	5		7	8	9		11						17		19				23
23				19		17						11		9	8	7		5	4	3	2	1	

we have a symmetry[36] in the Norton basis with the exception of the boxes that are supposed to contain 4 and 10. This was noticed by Matsutani who is in Yokohama, and he wondered whether this has to do with Weierstrass gaps [KMP>]. Some of you might know about Weierstrass gaps and the implications of it if there is some connection but that I don't know.[37] A remark to make is that when you are working with odd level functions, these 1, 2, 3 and 5 are the relevant entries that you need for the basis. Now what about 4 and 10? Well, you can write down expressions for 4 and 10. Basically speaking you can express a_{10} in terms of a_4, and I think there might well be an argument for putting 10 in the above tables and leaving things as they are otherwise, as you can express a_4 in terms of a_{10}. There is a little complication which I don't want to talk about.

This is an interesting remark and this leads to some other work of Matsutani [Mats] which suggests that there might be some connection with the variational problem of Euler and these replicable functions. The variational problem of Euler is what you get when you take an *elastica*, and elastica is Euler's word. It's about 1750 or so.

[35] Cf. [McKSev] for more discussions on the algorithms. Of course, one sees the beginning of the supersingular primes here.

[36] In that a number is accompanied by a gap, and vice versa, except for positions 4 and 10.

[37] We recall the statement of the Weierstraß gap theorem: For a compact genus $g > 0$ Riemann surface $X \ni x$, there exists exactly g numbers $1 = n_1 < n_2 < \ldots < n_g < 2g$ such that there does not exist a holomorphic function on $X \backslash x$ with a pole of order n_i at x.

What is an elastica? Take a metal ruler and push it in from the ends it will bend, and the question is what is the curve that you get when you bend it and that is a variational problem on the integral on the square of the curvature over the arc length. That was solved completely by Euler.[38] But, if you generalized it a bit into what Matsutani calls a quantized version [Mats2] you get some interesting objects, and genus 0 functions come up, rather than replicable functions. But whether there is a connection or not, I don't know.[39] So that's a curiosity which might be worth pursuing.

3.3 Faber Polynomials and Symmetric Functions

Let me come back to the Faber polynomials. They haven't been much studied, really. I'd like to mention some things here. I think the first of the following identities is perhaps the most important one to remember because it's easy to remember

$$1+\sum_{n\geq 1} h_n t^n = \prod_i (1-x_i t)^{-1} = \exp\left(\sum \frac{t^n}{n}\cdot\sum_i x_i^n\right) . \tag{29}$$

The h_n's are called the complete homogeneous symmetric functions.[40] Symmetric functions can be expressed in terms of $x_1,\ x_2\cdots$, and the h_n is of degree n and is a sum over the x_i's whether or not they are equal. So you are looking at sums of x_i's where $x_{i_1} \geq x_{i_2} \geq \cdots$. The functions on the right hand side of (29) are generating functions for these homogeneous symmetric functions.[41]

McDonald [McD] on symmetric functions shows that there is an involution that transforms the homogeneous symmetric functions to the elementary ones, and thus involution changes the generating functions. So there is a relation between

$$1+\sum_{n\geq 1} b_n t^n = \prod_i (1+x_i t) = \exp\left(\sum \frac{-t^n}{n}\cdot\sum_i x_i^n\right) \tag{30}$$

and (29).

Now, what are the Faber polynomial doing? Well, there is a completely different notion from what I've been talking about and that is that the Faber polynomials are related to a change of basis for a symmetric function. There are six standard bases for symmetric functions, five of which are well known, one of which is the Doubilet basis, and is called the forgotten symmetric functions [Dou]. Anyhow, one could take one of the above functions and multiply it by $1/t$ so that everything is shifted by 1. You take the following matrix (you have to be careful about these Faber polynomials)

$$A_n = \begin{pmatrix} b_1 & 1 & 0 & \cdots & 0 \\ 2b_2 & b_1 & 1 & \ddots & \vdots \\ 3b_3 & b_2 & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ nb_n & b_{n-1} & & & b_1 \end{pmatrix} . \tag{31}$$

We have $F_n(b_1,\cdots,b_n) = \det(A_n)$.[42] Write $F_n(z) = F_n(z,b_2,\cdots,b_n)$ with $F_0(f)=1$, $F_1(f)=f$, $F_2(f)=f^2-2b_2$, $F_3(f)=f^3-3b_2f-3b_3$. One can think of the Faber polynomial as $F_n(z)$, and remember that I shifted everything by 1 in (30) (by dividing by q); the function f we are dealing with here is

$$f(q)=\frac{1}{q}\left(1+\sum_{n\geq 1} b_n q^n\right) .$$

For replicable functions $b_1=0$ and $b_k=h_{k-1}$. Moreover, the $F_n(z)$ are isobaric, meaning homogeneous in the subscripts, in the sense that if you replace z by b_1 these are isobaric polynomials in the b_i's. It's very easy to get signs wrong, for me anyhow, and if you get things correct for the third one I think you are ok. These are Faber polynomials, and they come about from solving the Newton relations which I just described in terms of the roots of the polynomial $F_n(z)$ and its coefficients.[43] So that's what these Faber polynomials really are.

As I said yesterday, historically it's quite interesting that they were described by Faber in 1903 and mathematicians know them, and there are certainly due to someone earlier than Faber, but it might be one of this fairly folklorist things that goes back and that might be even predates Newton. There is also Girard,[44] but I don't know quite what role he had to play. Anyhow, Faà di Bruno is the person who predates Faber, and I don't know what date we are talking about, probably 1857; he's well known for the n-th derivative of the composition of two functions.[45]

[38] q.v. Leonhard Euler, "Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici lattissimo sensu accepti, chapter Additamentum 1", eulerarchive.org, E065, 1744.

[39] The reader is referred to [CumGan] for a discussion on the significance of the genus zero property.

[40] It is also called the plethystic exponential and has been the key to a programme of counting gauge-invariant operators in quantum field theories [FHH].

[41] This has been interpreted as fugacity-inserted plethystic exponential of a Hilbert series in the context of D-brane gauge theories [BFHH] and as Witt vectors in [FM>].

[42] That's the matrix, and I got into trouble when talking to Serre about this because we were using different notations and he objected very strongly to this. Note that the notation in footnote 27 is the unshifted version.

[43] That is, the recursion relations for the Faber polynomials described in footnote 24.

[44] Albert Girard (1595–1632), worked on fundamental theorem of algebra, symmetric polynomials, Fibonacci numbers, inter alia.

[45] Francesco Faà di Bruno (1825–1888), cf. "Sullo sviluppo delle Funzioni", Annali di Scienze Matematiche e Fisiche, 6:

3.4 Norton's Conjecture on Replicable Functions

Now, I'd like to make a statement on this main outstanding conjecture about replicable functions and nobody, as far as I know, has tried to solve it, but it's not quite as simple as one might wish. This is Norton's main conjecture and it is that:

Conjecture 1 (Norton's Conjecture). *A function f of the form* (9) *is replicable if either*

0. *$f(q) = 1/q + cq$ with $c \in \{0, -1, 1\}$ as we are working with integer coefficients. Respectively for these values of c we get the $\exp$, $\sin$ and $\cos$ functions,*[46] *which we call "Modular Fictions" and to be ignored henceforth.*
   Or, surprisingly using the modular polynomial you can get some results about these things, which are consequences of the modular polynomial for them. Things like the $\cos$ of twice the angle is a polynomial of the $\cos$ whereas the $\sin$ of twice the angle is not a polynomial in the $\sin$. This was proved by C. Cummins again [Cum], that this is all there is.[47]
1. *There exists $N: \ \Gamma_0(N) \subset G_f \subset \mathrm{Nor}(\Gamma_0(N))$, where Nor is the normalizer inside $\mathrm{PSL}_2(\mathbb{R})$, such that the compact Riemann surface $\widehat{G_f/\mathfrak{H}}$ obtained by adding a finite set of inequivalent cusps has genus 0, f is a principal modulus of this Riemann surface, and G_f is commensurable*[48] *with $\mathrm{PSL}_2(\mathbb{Z})$.*
   The maximal groups on the right side are called Helling groups, *and there is a paper by Conway [Con] called "Understanding groups like $\Gamma_0(N)$" which was referred to by J. Duncan. It's really a pretty piece of work in that he also proves Helling's theorem.*

What is needed is a proof of this result. I think the proper way to do this is to use a two-sided decomposition with respect to GL_2 of the adèles, and in that way one should be able to pick up the primes dividing the Monster's order, which are the primes you find inside the above levels N, namely the 15 supersingular primes.

What has been done to now, I might say something about computations if there is time, it's been done by some people; I had some visitors that did some work, C. Cummins did some, but the bulk of this stuff was done by S. Norton and I could describe it. Basically speaking, it was a very local picture that was used to do the computations to find these replicable functions. There are one or two functions that you know classically from Weber's [Web] and Schläfli's [Sch] work, and from there you can build up other functions which are closely related to replicable functions. You just keep looking and making sure that the genus is zero. But it is conceivable that something has been left out, it's unlikely but it's possible.

The number 616 is the number of replicable functions there are (with integer coefficients). That in itself is a quite interesting number, and there is amusing reference to J. Conway's remark. Remember that he was talking about[49] a group Y_{555} or perhaps Y_{666}. Well, 666 as you know occurs in the Bible, it's sort of a bad number, and I was giving this talk in Norway to the mathematical department and I mentioned this 616 as the number of replicable functions and a person in the back went out and he came back to me with a transparency on which there is something called the "Oxyrhynchus Papyrus" from few hundred A. D.[50] In the papyrus, which is written in Greek – it's very readable, it says that the real number shouldn't be 666; it should be 616. So that was very quite amusing. I don't know why or how he came across to this.

Here is another piece of numerology, entirely frivolous but nevertheless grounded on some deep mathematics, which may amuse you. The Leech lattice Λ_{24}, the famous even unimodular lattice in 24 dimensions, can be constructed from the even Lorentzian unimodular lattice $II_{25,1}$ in 26 dimensions using a Weyl vector $w = (0, 1, 2, 3, \ldots, 23, 24; 70)$. Then Λ_{24} is realized as $w^{\perp}/w$. That w is indeed an integral vector in $II_{25,1}$ follows from the remarkable Diophantine condition[51]

$$1^2 + 2^2 + 3^2 + \ldots + 23^2 + 24^2 = 70^2 \ . \tag{32}$$

Now, consider the first 24 q-series coefficients of the (normalized) modular J-function, viz,

479–480, 1855 and "Note sur une nouvelle formule de calcul differentiel", The Quarterly Journal of Pure and Applied Mathematics, 1: 359–360, 1857.

[46] Recall that the nome $q = \exp(2\pi i \tau)$.

[47] Indeed, these are the Chebyshev polynomials; the reader is referred to [McK-Seb] for discussions on how these famous polynomials are the simplest replicable functions.

[48] Recall that G_f was defined in (16) as the modular subgroup for which f is invariant (weight $k = 0$).

[49] The group is defined as follows. Consider a graph $G_{p,q,r}$ with a single tri-vertex, say a and 3 strands, each consisting of respectively p, q and r nodes joined up; call these $b_{1,\ldots,p}$, $c_{1,\ldots,q}$ and $d_{1,\ldots,r}$. Thus there are $p+q+r+1$ nodes in total. The group $Y_{p,q,r}$ is a Coxeter-type group with one generator associated to each of the nodes and presentation

$$Y_{p,q,r} = \langle a, b_{1,\ldots,p}, c_{1,\ldots,q}, d_{1,\ldots,r} | (gg')^{O(g,g')} = (ab_1b_2ac_1c_2ad_1d_2)^{10} = 1 \rangle \ ,$$

where g and g' are any of the a, b, c, d generators and that $O(g, g') = 3$ (respectively 2) if g and g' are adjacent (respectively, not adjacent). It was shown [Iva] that $Y_{5,5,5} \simeq Y_{4,4,4} \simeq \mathbb{M} \wr C_2$, the wreath product of the monster with the cyclic group of order 2.

[50] Manuscripts discovered in the C19th near Oxyrhynchus, Egypt, dating from 1st to 6th century AD., mostly housed in the Ashmolean in Oxford.

[51] Discussions on this equation and the emergence of 42 from $j(q)$ are presented in [HeMcK2>], in a volume in honour of J. H. Conway. The number of pages of the present notes is, of course, 42.

$196884, 21493760, 864299970, 20245856256, \ldots$ (that is we do not include the constant term 744 nor $1/q$ and start from $\mathcal{O}(q)$). Sum the squares of these 24 numbers and compute it modulo 70, you will get 42, which we all know to be the answer to the ultimate question of life, the universe and everything. So, here again, you see how Moonshine encodes all things.

3.5 Mahler's Recurrence Relations

Now, there is another interesting paper [ACMS] where we used recurrences to build up these series. What is not very well known is the recurrence that we use which people call Borcherds' recurrence,[52] which is fair enough, is really due to Mahler and Mahler wrote his paper in 1974, published in 1976 in the journal of the Australian Math. Soc. [Mah].

He was a very remarkable number theorist. He was a cripple and his father knew Carl Ludwig Siegel. So he got to get a rather good education without going through the usual processes of university and then he became a refugee in Manchester for 30 years and then he went from an assistant professor to a personal Chair, in Canberra where he lived and died not a long ago.

He wrote me a letter in 1982 or 1983 in which he said that he thought his paper has something to do with moonshine, and if I would come to Canberra and discuss it with him before he dies. I was in Montreal and he was in Australia. I went out there and unfortunately I got ill and I came back. So we never actually met, but Mahler recurrence is the recurrence that does the calculations.

We observe that (q.v. also [HeMcK3>])

$$\begin{aligned} 360+256 &= 616 \\ 120+2\times 248 &= 616\ . \end{aligned} \tag{33}$$

Maybe you can explain this. I'll give you the Mahler recurrence. I'm just going back a bit. I already remarked upon

$$f^{(n)}(nz) = \frac{1}{n}F_n(f(z)) - \frac{1}{n}\sum_{\substack{ad=n\\0\le b<d\\a<n}} f^{(a)}\left(\frac{az+b}{d}\right)\ ,$$

and that $f^{(n)}(nz)$ is invariant under $z \mapsto z+1/n$. Cummins and Norton have proved the replicability of rational Hauptmoduln [Cum-Nor], and remember that the replicability property is essentially combinatorial. You are just saying that things work nicely with Hecke operators whereas if you want to prove results about replicable being Hauptmodul you have to go from the combinatorial side to the analytic and it's a much more difficult thing to do.

There is some very remarkable work done by a student of Arne Meurman in Sweden. His name is Dmitry Kozlov [Koz]. Masao Koike is the first to realize what the generalized Hecke operator is. And if I have done it correctly. When $n = p$ is prime, the classical Hecke operator can be expressed in terms of the Atkin U_--operator and if you know it for p you know it for all integers because of the multiplicative property, and the Adams' V_--operator

$$\begin{aligned} T_p &= \frac{1}{p}V_p + U_p \\ V_p :\ f(q) &\mapsto f(q^p) \\ U_p :\ a_n q^n &\mapsto a_{pn} q^n\ . \end{aligned} \tag{34}$$

Koike recognizes this for the twisted Hecke operator in the very early days

$$\widehat{T}_p = \frac{1}{p}\Psi^p \circ V_p + U_p \tag{35}$$

with $\widehat{V}_p = \Psi^p \circ V_p :\ f(q) \mapsto f^{(p)}(q^p)$. Now if you go back to the moonshine to remember what this does for you, this means that the coefficients are traces which are the sums of the initial p-th power function. That's the context of moonshine but the rest is true for replicable functions generally.

There are some characteristic classes associated with this, and this is called the *Bott Cannibalistic Class*. So when I was in Harvard I asked Bott about it, but he told me he had forgotten so I don't think we'll get any further there. By the way there is a remarkable meeting in a week or two in Montreal, I think in June, a week on Bott's legacy with very good speakers including Witten and Atiyah among others in the University of Montreal.[53] I just make a passing remark is that $f \equiv f^{(p)} \bmod p$ because in $f^{(p)}$ the coefficients are sums of p-th powers of the coefficients of f. This is the work of Koike, and there is a nice survey article by him in Japanese in Sugaku, and there is an English version in number 160 of the AMS translations [Koi].

I think Mahler must be the first or the only mathematician, unless Conway has done the same thing, to publish a calculator program in the Royal Society of London Proceedings, but he did that in an subsequent paper where he does his computations. What's extremely interesting about Mahler is that for each prime p he has a recurrence relation to compute the coefficients of these functions and he is really com-

[52] The sort of recurrences which arise from Borcherds' proof is the remarkable one such as

$$j(p) - j(q) = \left(\frac{1}{p} - \frac{1}{q}\right)\prod_{m,n=1}^{\infty}(1-p^n q^m)^{c_{nm}}\ ,$$

where c_k are the q-expansion coefficients of $j(q)$. We will see more recurrences in §3.5.

[53] This is the conference "A Celebration of the Mathematical Legacy of Raoul Bott", CRM, June 9–13, 2008, Montreal. One can find the proceedings in *CRM Proceedings & Lecture Notes*, Volume 50, AMS 2010.

puting the q-coefficients of the j-function. There are several people who have done this for j, where the actual setup will work for replicable functions in general. This was true by Kozlov's thesis which was on the j-function, but in fact it applies to all replicable functions. The same with Mahler; he has this recurrence relations for all prime p for the j-function and he tries to use them in a broader context to other functions that have arisen in [Fri] and others and it doesn't work and he doesn't see why it doesn't work.

Today we know why it doesn't work. It's because the level of the function and the level of the Hecke operator are not coprime, and that's the modification that you need to go from Borcherds' formula to Mahler's formula. So I'm just going to give it to you here. For $p = 2$ Mahler has used

$$\begin{aligned} f\left(\frac{z}{2}\right) + f\left(\frac{z+1}{2}\right) + f(2z) &= f(z)^2 - 2a_1 \\ f\left(\frac{z}{2}\right) f(2z) + f\left(\frac{z+1}{2}\right) f(2z) + f\left(\frac{z}{2}\right) f\left(\frac{z+1}{2}\right) & \\ = 2a_2 f - f + 2(a_4 - a_1)\ , \end{aligned} \tag{36}$$

which might be look at as the elementary symmetric functions of degree 1 and 2 respectively in the function values $f\left(\frac{z}{2}\right)$, $f\left(\frac{z+1}{2}\right)$ and $f(2z)$. This is for the j-function originally, and then you have to make modifications. The modifications are simple to make; whenever you see a 2 in the f-value, say in $f(2z)$ you change it to $f^{(2)}(2z)$. These are the recurrence relations from Mahler

$$\begin{aligned} a_{4k} &= \sum_{j=1}^{k-1} a_j a_{2k-j} + \frac{1}{2}(a_k^2 - a_k^{(2)})\ , \\ a_{4k+} &= a_{2k+3} \sum_{j=1}^{k} a_j a_{2k+2-j} + \frac{1}{2}(a_{k+1}^2 - a_{k+1}^{(2)}) + \frac{1}{2}(a_{2k}^2 - a_{2k}^{(2)}) \\ &\quad - a_2 a_{2k} + \sum_{j=1}^{k} a_j^{(2)} a_{4k-4j} + \sum_{j=1}^{2k-1} (-1)^j a_j a_{4k-j}\ , \\ a_{4k+2} &= a_{2k+2} + \sum_{j=1}^{k-1} a_j a_{2k+1-j}\ , \\ a_{4k+3} &= a_{2k+4} \sum_{j=1}^{k+1} a_j a_{2k+3-j} + \frac{1}{2}(a_{2k+1}^2 - a_{2k+1}) - a_2 a_{2k+1} + \\ &\quad + \sum_{j=1}^{k} a_j^{(2)} a_{4k+2-4j} + \sum_{j=1}^{2k} (-1)^j a_j a_{4k+2-j}\ , \end{aligned} \tag{37}$$

which do come from

$$\begin{aligned} f\left(\frac{z}{2}\right) + f\left(\frac{z+1}{2}\right) + f^{(2)}(2z) &= f(z)^2 - 2a_1 \\ f\left(\frac{z}{2}\right) f^{(2)}(2z) + f\left(\frac{z+1}{2}\right) f^{(2)}(2z) + f\left(\frac{z}{2}\right) f\left(\frac{z+1}{2}\right) & \\ = 2a_2 f - f + 2(a_4 - a_1)\ , \end{aligned} \tag{38}$$

and these are Borcherds' relations. That's the difference between Mahler's and Borcherds'.

Notice that the original ones are universal, and if you have any function in this case of odd level these are the recurrence relations for its coefficients and if you have a function of even level then you have to know what $f^{(2)}$ is in order to make use them; you have to use coefficients from $f^{(2)}$ in order to build up these relations. They come in a group of four. They are extremely good for computing several hundreds coefficients quite easily, but if you speak to someone like Atkin few hundreds is nothing; he computes the first 10,000.

Now, to go back to some remarks a bit earlier on. We find $f^{(p)} = \sum_n h_n^{(p)} q^n$ and

$$h_n^{(p)} = p h_{pn,p} - p a_{p^2 n}\ . \tag{39}$$

On $\mathbb{M}$: $f_{\langle g \rangle} \longrightarrow f_{\langle g \rangle}^{(k)} = f_{\langle g^k \rangle}$, this is the replication formula I mentioned to you.[54] Since $|\mathbb{M}| < \infty$ there exists $k = k_0$ such that $g^{k_0} = id$ and so $f_{\langle g^{k_0} \rangle} = f_{\langle id \rangle} = j$. So in a sense you can think of the functions we start with as replication roots of the j-function.

Acknowledgements

JM is grateful to the organizers for the Kashiwa conference for their warm hospitality and that of the NSERC of Canada. YHH would like to thank the Science and Technology Facilities Council, UK, for grant ST/J00037X/1, and the Chinese Ministry of Education, for a Chang-Jiang Chair Professorship at NanKai University.

References

NB. We place a superscript $^>$ for all references which have appeared after the date of the delivery of the lectures, as they are added in the editorial process.

[ACMS] D. Alexander, C. Cummins, J. McKay, C. Simons, "Completely replicable functions," in Liebeck, Saxl, "Groups, Combinatorics and Geometry", LMS Lecture Note Series (CUP) 165, pp. 87–98.

[AFS] S. D. Smith, "On the head characters of the Monster simple group", Finite Groups – Coming of Age (Montréal, 1982), Contemp. Math. 45. American Mathematical Society, Providence, 1996.

[Apo] T. M. Apostol, "Modular functions and Dirichlet series in number theory", Second edition. Graduate Texts in Mathematics, 41. Springer-Verlag, New York, 1990.

[54] Thus we come full circle back to the Monster $\mathbb{M}$. Recall that for conjugacy classes, and in particular of cyclic groups $\langle g \rangle$, one can write down McKay-Thompson series, these obey replication identities. The question then is whether group structure in $\mathbb{M}$, such as the power map $g \mapsto g^k$ are reflected in the replicable functions $f \mapsto f^{(k)}$.

[AKM] P. S. Aspinwall, S. H. Katz, D. R. Morrison, “Lie groups, Calabi-Yau threefolds, and F theory”, Adv. Theor. Math. Phys. 4 (2000), 95 [arXiv:hep-th/0002012].

[Atk-Swi] A. O. L. Atkin, H. P. F. Swinnerton-Dyer, “Modular forms on noncongruence subgroups”, Combinatorics (Proc. Sympos. Pure Math., Vol. XIX, Univ. Californa, Los Angeles, Calif., 1968), pp. 1–25. AMS., Providence, R.I., 1971.

[ATLAS] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, R. A. Wilson, “Atlas of Finite groups”, Oxford University Press, Eynsham, UK, 1985. http://brauer.maths.qmul.ac.uk/Atlas/v3/

[BFHH] S. Benvenuti, B. Feng, A. Hanany, Y.-H. He, “Counting BPS Operators in Gauge Theories: Quivers, Syzygies and Plethystics,” JHEP 0711 (2007), 050 [arXiv:hep-th/0608050].

[Blog$^>$] Some blogs on Moonshine:
http://ncatlab.org/nlab/show/Moonshine,
http://www.neverendingbooks.org/index.php/monsters-and-moonshine-a-booklet.html

[Bor] R. E. Borcherds, “Vertex algebras, Kac-Moody algebras, and the monster”, Proc. Nat. Acad. Sci. 83 (1986), 3068; “Monstrous moonshine and monstrous Lie superalgebras”, Invent. Math. 109 (1992), 405–444.

[Cam] P. J. Cameron, Permutation Groups, LMS Student Texts 45. CUP, 1999.

[Chev] C. Chevalley, “Sur certains groupes simples”, (French) Tohoku Math. J. (2) 7 (1955), 14–66.

[CohMck] H. Cohn, J. McKay, “Spontaneous Generation of Modular Invariants”, Math. Comp. 65 (1996), 1295–1309.

[Con] J. H. Conway, “Understanding groups like $\Gamma_0(N)$”, Groups, difference sets, and the Monster (Columbus, OH, 1993), 327–343.

[Con-Nor] J. H. Conway, S. P. Norton, “Monstrous moonshine”, Bull. London Math. Soc. 11 (1979), no. 3, 308–339.

[Cum] C. J. Cummins, “Some comments on replicable functions”, Modern trends in Lie algebra representation theory (Queen’s Univ., Kingston, ON, 1994) 48–55; Queen’s Papers in Pure and Appl. Math. 94 (1994).

[CumGan] C. J. Cummins, T. Gannon, “Modular equations and the genus zero property of moonshine functions”, Inventiones mathematicae 129 (1997), no. 3.

[Cum-Nor] C. J. Cummins, S. P. Norton, “Rational Hauptmoduls are replicable”, Canad. J. Math. 47 (1995), no. 6, 1201–1218.

[CDH$^>$] M. C. N. Cheng, J. F. R. Duncan, J. A. Harvey, “Umbral Moonshine”, arXiv:1204.2779 [math.RT].

[CDHKW$^>$] M. C. N. Cheng, X. Dong, J. Duncan, J. Harvey, S. Kachru, T. Wrase, “Mathieu Moonshine and N=2 String Compactifications”, arXiv:1306.4981 [hep-th].

[Che$^>$] M. C. N. Cheng, “K3 Surfaces, N=4 Dyons, and the Mathieu Group M24”, Commun. Num. Theor. Phys. 4 (2010), 623 [arXiv:1005.5415 [hep-th]].

[de Bra] L. de Branges, “Underlying concepts in the proof of the Bieberbach conjecture”, A plenary address presented at the International Congress of Mathematicians held in Berkeley, California, August 1986. Introduced by Max M. Schiffer. ICM Series. American Mathematical Society, Providence, RI, 1988.

[Deg-Wen$^>$] A. Degeratu, K. Wendland, “Friendly giant meets pointlike instantons? On a new conjecture by John McKay”, Moonshine: the first quarter century and beyond, 55–127, London Math. Soc. Lecture Note Ser. 372. Cambridge Univ. Press, Cambridge, 2010.

[DGO$^>$] J. F. R. Duncan, M. J. Griffin, K. Ono, “Moonshine”, Res. in the Math. Sciences 2 (2015), 11 [arXiv:1411.6571 [math.RT]].

[DGO2$^>$] J. F. R. Duncan, M. J. Griffin, K. Ono, “Proof of the Umbral Moonshine Conjecture”, Res. in the Math. Sciences 2 (2015), 26 [arXiv:1503.01472 [math.RT]].

[DKM] D. Dummit, H. Kisilevsky, J. McKay, “Multiplicative products of eta functions”, in Finite groups – coming of age (Montreal, Que., 1982), Contemp. Math. 45, pp. 89–98. Amer. Math. Soc., Providence, RI, 1985. (The reviewer in Math Reviews points out and corrects a printing error in the paper.)

[Dou] P. Doubilet, “On the Foundations of Combinatorial Theory. VII: Symmetric Functions through the theory of distribution and occupancy”, Studies in Applied Maths LI (1972), 4.

[DO$^>$] J. F. R. Duncan, K. Ono, “The Jack Daniels Problem”, J. Number Theory 161 (2016), 230–239 [arXiv:1411.5354 [math.NT]].

[duS$^>$] M. du Sautoy, “Finding Moonshine: A Mathematician’s Journey Through Symmetry”, Harper Perennial, 2009. ISBN-13: 978-0007214624.

[EOT$^>$] T. Eguchi, H. Ooguri, Y. Tachikawa, “Notes on the K3 Surface and the Mathieu group M_{24}”, Exper. Math. 20 (2011), 91 [arXiv:1004.0956 [hep-th]].

[Erd] C. Erdenberger, “The Kodaira dimension of certain moduli spaces of abelian surfaces”, Math. Nachr. 274/275 (2004), 32–39.

[FHH] B. Feng, A. Hanany, Y.-H. He, “Counting gauge invariants: The Plethystic program”, JHEP 03 (2007), 090 [arXiv:hep-th/0701063 [hep-th]].

[FM$^>$] R. Friedrich, J. McKay, “Formal Groups, Witt vectors and Free Probability”, arXiv:1204.6522.

[Fra] J. S. Frame, “The Theory of Tables of Group Characteristics”, Harvard Univ. thesis, 1933.

[Fri] R. Fricke, “Die elliptischen Funktionen und ihre Anwendungen”, Zweiter Band, 1922; Reprint, Springer Verlag, 2011.

[GHV$^>$] M. R. Gaberdiel, S. Hohenegger, R. Volpato, “Mathieu Moonshine in the elliptic genus of K3”, JHEP 1010 (2010), 062 [arXiv:1008.3778 [hep-th]].

[Gan] T. Gannon, “Monstrous moonshine: The First twenty five years”, arxiv:math/0402345 [math-qa].

[GK$^>$] S. Govindarajan, K. Gopala Krishna, “BKM Lie superalgebras from dyon spectra in Z(N) CHL orbifolds for composite N”, JHEP 1005 (2010), 014 [arXiv:0907.1410 [hep-th]].

[Grun] H. Grunsky, “Koeffizientenbedingungen für schlicht abbildende meromorphe Funktionen”, Math. Z. 45 (1939), 29–61.

[HeJe] Y.-H. He, V. Jejjala, “Modular matrix models”, hep-th/0307293.

[HM$^>$] Y.-H. He, J. McKay, “N=2 Gauge Theories: Congruence Subgroups, Coset Graphs and Modular Surfaces”, J. Math. Phys. 54 (2013), 012301 [arXiv:1201.3633 [hep-th]].

[HeMcK$^>$] Y.-H. He, J. McKay, “Eta Products, BPS States and

K3 Surfaces", arXiv:1308.5233 [hep-th].

[HeMcK2$^>$] Y.-H. He, J. McKay, "Moonshine and the Meaning of Life", in Contemporary Mathematics 694, Ed. M. Bhagarva et al., 2017 [arXiv:1408.2083 [math.NT]].

[HeMcK3$^>$] Y.-H. He, J. McKay, "Sporadic and Exceptional", arXiv:1505.06742 [math.AG].

[HMR$^>$] Y.-H. He, J. McKay, J. Read, "Modular Subgroups, Dessins d'Enfants and Elliptic K3 Surfaces", arXiv:1211.1931 [math.AG].

[HSZ$^>$] P. Haldar, A. Sinha, A. Zahed, "Quantum field theory and the Bieberbach conjecture", arXiv:2103.12108 [hep-th].

[Iva] A. Ivanov, "Y-groups via transitive extension", J. Alg. 218 (1999), 412–435.

[Kil] L. J. P. Kilford, "Generating spaces of modular forms with η-quotients", arXiv:math/0701478.

[KMP$^>$] J. Komeda, S. Matsutani, E. Previato, "The sigma function for Weierstrass semigroups $\langle 3,7,8\rangle$ and $\langle 6,13,14,15,16\rangle$", arXiv:1303.0451 [math.AG].

[Koi] M. Koike, "Moonshine: a mysterious relationship between simple groups and automorphic functions" in Selected papers on number theory, algebraic geometry, and differential geometry, Amer. Math. Soc. Transl. Ser. 2, 160, pp. 33–45. AMS, Providence, RI, 1994.

[Kon] T. Kondo, "The automorphism group of Leech lattice and elliptic modular functions", J. Math. Soc. Japan 37 (1985), no. 2, 337–362.

[Koz] D. Kozlov, "On Functions Satisfying Modular Equations for Infinitely Many Primes", Canad. J. Math. 51 (1999), 1020–1034; – "On Completely Replicable Functions and Extremal Poset Theory", Masters Thesis, University of Lund, 1994.

[Mah] K. Mahler, "On a Class of Non-Linear Functional Equations Connected with Modular Functions", J. Aust. Math. Soc. 22A (1976), 65–118.

[Mar] Y. Martin, "Multiplicative η-quotients", Trans. Amer. Math. Soc. 348 (1996), no. 12, 4825–4856.

[MarOno] Y. Martin, K. Ono, "Eta-quotients and elliptic curves", Proc. Amer. Math. Soc. 125 (1997).

[Mat] E. Mathieu, "Mémoire sur l'étude des fonctions de plusieurs quantités, sur la maniére de les former et sur les substitutions qui les laissent invariables", J. Math. Pures Appl. (Liouville) (2) VI (1861), 241–323.

[Mats] S. Matsutani, "Euler's Elastica and Beyond", J. Geometry and Symmetry in Physics 17 (2010), 45–86.

[Mats2] S. Matsutani, "Relations in a quantized elastica", J. Phys. A. 41 (2008), no. 7, 075201.

[McD] I. G. Macdonald, "Symmetric functions and Hall polynomials", Second ed. Oxford Mathematical Monographs. OUP, 1995. ISBN 0-19-853489-2.

[McK] J. McKay, "Graphs, Singularities, and Finite Groups", Proc. Symp. Pure Math. 37 (1980), 183–186.

[McK2] J. McKay, "The Essentials of Monstrous Moonshine" Adv. Studies in Pure Maths 32 (2001), 347–353.

[MS] J. McKay, A. Sebbar, "Arithmetic Semistable Elliptic Surfaces", Proceedings on Moonshine and related topics (Montréal, QC, 1999), CRM Proc. Lecture Notes 30, pp. 119–130. Amer. Math. Soc., Providence, RI, 2001.

[McK-Seb] J. McKay, A. Sebbar, "Replicable functions: an introduction", Frontiers in number theory, physics, and geometry. II, pp. 373–386, Springer, Berlin, 2007.

[McKSev] J. McKay, D. Sevilla, "Decomposing replicable functions", LMS J. Comput. Math. 11, 146–171 [arXiv:0803.3419 [math.NT]].

[New] M. Newman, "Modular Forms Whose Coefficients Possess Multiplicative Properties", Annals of Mathematics 70 (1959), no. 3, 478–489.

[Mil] G. A. Miller, "Sur plusieurs groupes simples", (French) Bull. Soc. Math. France 28 (1900), 266–267.

[Nor] S. P. Norton, "More on moonshine", Computational group theory, pp. 185–193. London Academic Press, 1984.

[Pomm] Ch. Pommerenke, "Uber die Faberschen Polynome schlichter Funktionen", Mathematische Z. 85 (1964), 197–208.

[PolyMath$^>$] The polymath project, https://polymathprojects.org/

[Ogg] A. Ogg, "Modular Functions", in The Santa Cruz Conference on Finite Groups, pp. 521–532. Ed. B. Copperstein, G. Mason. Providence, RI: Amer. Math. Soc., June 25–July 20, 1979.

[Rez] B. Reznick, "Resources for Research (an always preliminary list)", http://www.math.uiuc.edu/~reznick/rfr.html

[Ron] M. Ronan, "Symmetry and the Monster: One of the greatest quests of mathematics", OUP, 2007. ISBN-13: 978-0192807236.

[San$^>$] G. Sankaran, "A supersingular coincidence", arXiv:2009.11379 [math.NT].

[Sch] L. Schläfli, J. H. Graf ed., "Theorie der vielfachen Kontinuität", Republished by Cornell University Library historical math monographs 2010 (in German), Zürich, Basel: Georg & Co., 1901 [1852]. ISBN 978-1-4297-0481-6.

[Seb1] A. Sebbar, "Classification of torsion-free genus zero congruence groups", Proc. Amer. Math. Soc. 129 (2001), 2517–2527.

[Seb2] A. Sebbar, "Modular subgroups, forms, curves and surfaces", Canad. Math. Bull. 45 (2002), no. 2, 294–308.

[Ser] J.-P. Serre, "Cours d'arithmétique", Springer, 1973.

[THM$^>$] V. Tatitscheff, Y.-H. He, J. McKay, "Cusps, Congruence Groups and Monstrous Dessins", Indagationes Mathematicae, 31 (2020), no. 6, 1015–1065 [arXiv:1812.11752 [math.NT]].

[Web] H. M. Weber, "Lehrbuch der Algebra" (in German) (3rd ed.), New York: AMS Chelsea Publishing, 1981 [1898]. ISBN 978-0-8218-2971-4.