5

Vertex operator algebras

Vertex operator algebras (VOAs) are a mathematically precise formulation of the notion of *chiral algebra* (Section 4.3.2), the symmetry algebra of conformal field theory. They constitute the simplest expression we have of the machine that associates the Monster $\mathbb{M}$ with the Hauptmoduls. VOAs were first defined by Borcherds, and their theory has since been developed by a number of people. We begin with the rather complicated definition, before turning to our greatest interest: their representation theory. The final section sketches some relations of vertex algebras to geometry. See, for example, [**201**], [**330**], [**197**], [**376**] for more complete treatments; a more physically minded introduction is provided in [**242**].

Vertex operator algebras are not a type of operator algebra; rather, they are an algebra of *vertex operators*. Vertex operators arose first in string theory back in the early 1970s as a device for computing string amplitudes. They appeared independently in the mathematical literature (starting with [**377**]) in order to realise affine Kac–Moody algebras and their modules as algebras of differential operators. Today, just as we define a 'vector' to be an element of a vector space, we define a 'vertex operator' to be a formal power series $Y(u, z)$ appearing in a vertex algebra.

5.1 The definition and motivation

5.1.1 Vertex operators

In bosonic string theory, the vertex operator (Section 4.3.1) corresponding to the absorption of a tachyon with momentum $k = (k^\mu)$ at world-sheet position z and space-time position $X(z) = (X^\mu(z))$ is the normal-ordered expression $V(k, z) = :e^{\mathrm{i}k\cdot X(z)}:$. Write

$$X^\mu(z) = x^\mu - \mathrm{i}p^\mu \log(z) + \mathrm{i}\sum_{n\neq 0} \frac{1}{n}\alpha_n^\mu z^{-n},$$

where x^μ and p^μ are classically the position and momentum of the string's centre-of-mass, and α_n^μ its oscillation coordinates. Then the vertex operator is (chapter 2.2 of [**261**])

$$V(k, z) = \exp\left(k \cdot \sum_{n\geq 1} \frac{\alpha_{-n}}{n} z^n\right) z^{k\cdot p-1} e^{\mathrm{i}k\cdot x} \exp\left(-k \cdot \sum_{n\geq 1} \frac{\alpha_n}{n} z^{-n}\right). \qquad (5.1.1\mathrm{a})$$

Independently, Lepowsky and Wilson realised the affine algebra $A_1{}^{(1)}$ using differential operators (they tried to do this because finite-dimensional Lie algebras often act

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

as differential operators, for example, on the space of functions on an associated Lie group):

Theorem 5.1.1 [**377**] *A basis for the affine algebra ${A_1}^{(1)}$ consists of the operators*

$$1,\ y_n,\ \frac{\partial}{\partial y_n},\ Y_k \qquad \forall n \in \left\{\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots\right\},\ k \in \frac{1}{2}\mathbb{Z},$$

thought of as operators on the space $\mathbb{C}[y_{1/2}, y_{3/2}, y_{5/2}, \ldots]$ *of polynomials in the* y_n *(we are ignoring the derivation in* ${A_1}^{(1)}$*). The differential operators* Y_k *are the homogeneous components of the formal generating function*

$$Y(z) = \sum_{k\in\frac{1}{2}\mathbb{Z}} Y_k\, z^k = \exp\left(\sum_n \frac{y_n}{n} z^n\right) \exp\left(-2\sum_n \frac{\partial}{\partial y_n} z^{-n}\right). \tag{5.1.1b}$$

In particular, (ignoring the derivation ℓ_0) ${A_1}^{(1)}$ is spanned by a central term C, as well as $e \otimes t^m$, $f \otimes t^m$, $h \otimes t^m$ for each $m \in \mathbb{Z}$ (Section 3.2.2). In Theorem 5.1.1, 1 corresponds to C. For each $n \in \mathbb{N} + 1/2$, the operators y_n and $\partial/\partial y_n$ correspond (up to numerical proportionality factors) respectively to $e \otimes t^{\mp n-1/2} + f \otimes t^{\mp n+1/2}$, and $Y_{\pm n}$ corresponds to $-e \otimes t^{\pm n-1/2} + f \otimes t^{\pm n+1/2}$. For $k \in \mathbb{Z}$, the operator Y_k corresponds to $h \otimes t^k$ (for $k \neq 0$) and $h \otimes 1 - C/2$ for $k = 0$.

It was Garland who first recognised the formal resemblance between these transcendental expressions (5.1.1a) and (5.1.1b). Note that when expanded out they both involve a sum over powers of z, unbounded in both the positive and negative directions. Doubly-infinite series scream of convergence difficulties. The fractional indices n, k in Theorem 5.1.1 are a signature of what we today call twisted vertex operators.

The geometric meaning of the vertex operator is perhaps best explained in the context of the loop group (Section 3.2.6). Suppose the loop group $\mathcal{L}S^1$ acts on some space $\mathcal{H}$. For each $0 \leq s \leq 2\pi$ and $\epsilon > 0$, consider the loop $\gamma_s^\epsilon \in \mathcal{L}S^1$ defined by

$$\gamma_s^\epsilon(t) = \begin{cases} 1 \in S^1 & \text{for } |s-t| \geq \epsilon \\ \exp\left(\pi \mathrm{i} \frac{s-t}{2\epsilon}\right) \in S^1 & \text{for } s-\epsilon < t < s+\epsilon \end{cases},$$

for all $0 \leq t < 2\pi$. In words, γ_s^ϵ stays at the identity $1 \in S^1$ for all time t, except for a small interval around $t \approx s$ where the loop rapidly winds around S^1 once. This loop corresponds to some operator on $\mathcal{H}$; the limit (appropriately taken) as $\epsilon \to 0$ is an operator-valued distribution on $\mathcal{H}$ called a vertex operator (see chapter 13 of [**465**] for details).

5.1.2 Formal power series

As we saw last chapter, the basic object of quantum field theory is the quantum field. It is tempting to think of it as a choice of operator $\hat{A}(x)$ at each space-time point x, but 'function' (or 'section of a vector bundle' for that matter) is too narrow a concept even in free theories.

The *analytic* way to make sense of 'functions' like quantum fields is through distributions, and this was the approach taken in Section 4.2.4. We will describe now the *algebraic* alternative. These two approaches are not equivalent: you can do some things in one approach that you can't do in the other, at least not without difficulty (Section 5.4.1). But as always the algebraic approach is considerably simpler technically – there are no convergence concerns to address – and it is remarkable how much can still be captured. It was first created around 1980 by Garland and Date–Kashiwara–Miwa to make sense of doubly-infinite series like (5.1.1), and is now the language of VOAs. Good introductions to the material in this subsection are [**201**], [**330**], [**376**].

Keep in mind that in CFT we are trying to capture operator-valued 'functions' on two-dimensional Euclidean space-time (Section 4.3.2). Locally space-time looks like $\mathbb{C}$; as explained in Section 4.3, we like to compactify the external legs – for example, for an incoming string tracing a cylindrical world-sheet, the space-time point (x, t) is associated with the complex number $z = e^{t+\mathrm{i}x}$, so time $t = -\infty$ corresponds to $z = 0$.

Let $\mathcal{W}$ be any vector space. Define $\mathcal{W}[[z^{\pm 1}]]$ to be the set of all formal series $\sum_{n=-\infty}^{\infty} w_n z^n$, where the coefficients w_n lie in our space $\mathcal{W}$. We don't ask here whether a given series converges or diverges; z is merely a formal place-keeping variable. We will also be interested in the space $\mathcal{W}[z^{\pm 1}]$ of Laurent polynomials, that is, expressions of the form $\sum_{n=-M}^{N} w_n z^n$. $\mathcal{W}[[z^{\pm 1}]]$ itself forms a vector space, using the obvious addition and scalar multiplication.

Our aim here is to describe quantum fields, so we want our formal series to be operator-valued. To do this, choose $\mathcal{W}$ to be a vector space of operators (matrices if you prefer): $\mathcal{W} = \mathrm{End}(\mathcal{V})$, for some space $\mathcal{V}$. We are actually interested in $\mathcal{V}$ being the infinite-dimensional state-space of the theory, but in the following examples we take $\mathcal{V} = \mathbb{C}$, that is power series with numerical coefficients.

We can now multiply our formal series in the obvious way. For example, consider $\mathcal{V} = \mathbb{C}$, and take $c(z) = z^{21} - 5z^{100}$ and $d(z) = \sum_{n=-\infty}^{\infty} z^n$. Then

$$c(z)\, d(z) = \sum_{n \in \mathbb{Z}} z^{n+21} - 5 \sum_{n \in \mathbb{Z}} z^{n+100} = \sum_{n \in \mathbb{Z}} z^n - 5 \sum_{n \in \mathbb{Z}} z^n = -4d(z).$$

This simple calculation tells us many things.

(i) *We can't always divide*: $c(z)\, d(z) = -4d(z)$ shows that the cancellation law fails and that $\mathbb{C}[[z^{\pm 1}]]$ isn't even an integral domain.

(ii) Try to compute the square $d(z)^2$: we get infinity. That is, you can't always multiply in $\mathcal{W}[[z^{\pm 1}]]$.

(iii) Working out a few more multiplications of this kind, we find that $f(z)\, d(z) = f(1) \times d(z)$ for any f for which $f(1)$ exists (e.g. any Laurent polynomial $f \in \mathcal{W}[z^{\pm 1}]$). Thus $d(z)$ is what we have called the Dirac delta $\delta(z - 1)$ centred at $z = 1$. (You can think of it as the Fourier expansion of the Dirac delta, followed by a change of variables.) So of course it makes perfect sense that we couldn't work out $d(z)^2$ – we were trying to square the Dirac delta, which we know is impossible!

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

There is a certain divergence of notations here: should δ be written additively (i.e. $\delta(z-1)$), in the familiar way, or should it be written multiplicatively (i.e. $\delta(z)$), in the more honest way? Throughout this chapter *we use the multiplicative notation*. So we get

$$\delta(z) := \sum_{n=-\infty}^{\infty} z^n. \tag{5.1.2}$$

In fact, the best notation of all would be the awkward $\delta(z)\,\mathrm{d}z$, since the Dirac delta centred at $z=a$ is $\sum_n z^n a^{-n-1} = a^{-1}\delta(z/a)$.

Making contact with Section 1.3.1, the Laurent polynomials $(\mathrm{End}\,\mathcal{V})[z^{\pm 1}]$ play the role here of the smooth functions C^∞_{cs} with compact support, and the formal power series $(\mathrm{End}\,\mathcal{V})[[z^{\pm 1}]]$ play the role of its dual. So these power series $f \in (\mathrm{End}\,\mathcal{V})[[z^{\pm 1}]]$ are formal distributions – this is why $f(z)$ usually diverges. The evaluation $f(p)$ of a distribution $f \in (\mathrm{End}\,\mathcal{V})[[z^{\pm 1}]]$ on the test function $p \in (\mathrm{End}\,\mathcal{V})[z^{\pm 1}]$ is given by the 'formal residue' $\mathrm{Res}_z(f(z)\,p(z)) \in \mathrm{End}\,\mathcal{V}$, where

$$\mathrm{Res}_z\left(\sum_{n\in\mathbb{Z}} b_n z^n\right) = b_{-1}. \tag{5.1.3a}$$

The idea is that, up to a factor of $(2\pi\mathrm{i})^{-1}$, this would equal the contour integral of $g(z) = \sum b_n z^n$ around a small circle about $z=0$, at least for meromorphic g. Hence Res_z obeys many of the familiar properties of integrals, such as integration by parts:

$$\mathrm{Res}_z(g\,\partial_z f) = -\mathrm{Res}_z(f\,\partial_z g), \tag{5.1.3b}$$

where $\partial_z f$ is the formal (term-by-term) derivative of $f(z)$. For example, the formal distribution $a^{-k-1}(\partial_z^k\delta)(z/a)$ takes the test function $f(z)$ to the value $(-1)^k(\partial_z^k f)(a)$. Because of the usefulness of the notion of residue, we write

$$f(z) = \sum_{n\in\mathbb{Z}} a_n z^n =: \sum_{m\in\mathbb{Z}} a_{(m)} z^{-m-1}, \tag{5.1.3c}$$

where $a_{(m)} = \mathrm{Res}_z(z^m f(z)) = a_{-m-1}$ is called a *mode*.

Similar remarks hold for several variables z_i. The distributions are the formal series

$$f(z_1,\ldots,z_k) = \sum_{n_i\in\mathbb{Z}} a_{n_1,\ldots,n_k} z_1^{n_1}\cdots z_k^{n_k} = \sum_{m_i\in\mathbb{Z}} a_{(m_1,\ldots,m_k)} z_1^{-m_1-1}\cdots z_k^{-m_k-1}$$

in $\mathcal{W}[[z_1^{\pm 1},\ldots,z_k^{\pm 1}]]$, and the test functions $f(z_1,\ldots,z_k) \in \mathcal{W}[z_1^{\pm 1},\ldots,z_k^{\pm 1}]$ consist of those power series with only finitely many nonzero terms. The Dirac delta centred at $z_1 = z_2$ is given by $z_2^{-1}\delta(z_1/z_2) = z_1^{-1}\delta(z_2/z_1)$.

But we must not get overconfident:

Paradox 5.1 *Consider the following product:*

$$\delta(z) = \left[\left(\sum_{n\geq 0} z^n\right)(1-z)\right]\delta(z) = \left(\sum_{n\geq 0} z^n\right)[(1-z)\,\delta(z)] = \left(\sum_{n\geq 0} z^n\right)[0\,\delta(z)] = 0.$$

When physicists are confronted with 'paradoxes' such as this, they respond by treading with care when they are involved in a calculation reminiscent of the paradoxes,

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

and otherwise trusting their instincts. Mathematicians typically over-react: after kicking themselves for walking head first into a 'paradox', they devise a rule absolutely guaranteeing that the paradox will always be safely avoided in the future. We will follow the mathematicians' approach, and in the next few paragraphs describe how to avoid Paradox 5.1 by forbidding certain innocent-looking products.

Recall that we are actually interested in the space $\mathcal{W} = \mathrm{End}(\mathcal{V})$. We call infinitely many linear maps $w^{(i)} \in \mathrm{End}(\mathcal{V})$ *algebraically summable* if for every vector $v \in \mathcal{V}$, only finitely many values $w^{(i)}v \in \mathcal{V}$ are different from 0. In other words, fixing a basis for $\mathcal{V}$, only finitely many of the matrices $w^{(i)}$ have a nonzero first column, only finitely many have a nonzero second column, etc. The usual notation '$\sum_i w^{(i)}$' will denote the well-defined endomorphism sending each $v \in \mathcal{V}$ to that effectively finite sum $\sum_i w^{(i)}v$.

Consider a family (possibly infinite) of formal series $w^{(i)}(z) \in \mathcal{W}[[z^{\pm 1}]]$. We certainly have a well-defined sum $\sum_i w^{(i)}(z)$ if for each fixed n, the set $\{w_n^{(i)}\}$ (as i varies) of maps is algebraically summable. We shall call such a sum *algebraically defined*, and write

$$\sum_i w^{(i)}(z) = \sum_{n\in\mathbb{Z}} \left(\sum_i w_n^{(i)}\right) z^n .$$

All other sums are forbidden. Likewise, we certainly have a well-defined product $\prod_{i=1}^m w^{(i)}(z)$ of finitely many formal power series if for each n, the set

$$\left\{w_{n_1}^{(1)}\, w_{n_2}^{(2)} \cdots w_{n_m}^{(m)}\right\}_{\sum n_i = n}$$

(vary the n_i subject to the constraint $\sum_i n_i = n$) is algebraically summable. Again, call such a product *algebraically defined* and set it equal to

$$\prod_{i=1}^m w^{(i)}(z) = \sum_{n\in\mathbb{Z}} \left(\sum_{n_1+\cdots+n_m=n} w_{n_1}^{(1)}\, w_{n_2}^{(2)} \cdots w_{n_m}^{(m)}\right) z^n ,$$

where the second sum is over all m-tuples (n_i) obeying $\sum_i n_i = n$. All other products (e.g. all infinite ones) are forbidden. An algebraically defined product is necessarily associative.

There are certainly more general ways to have a well-defined product or sum. For example, according to our rule, the series $\sum_n 2^{-n}$ would be forbidden. In this way we avoid the more complicated realm of convergence issues. In short, we are doing *algebra* here, and don't want to be distracted by the dust clouds kicked up by mere *analytic* concerns. Such restrictions are common in infinite-dimensional algebra (recall footnote 14 in chapter 1). The product of a distribution $f \in \mathcal{W}[[z_1^{\pm 1}, \ldots, z_k^{\pm 1}]]$ with a test function $p \in \mathcal{W}[z_1^{\pm 1}, \ldots, z_k^{\pm 1}]$ is always defined, and will be a distribution. The explanation of Paradox 5.1 is that although $(\sum z^n)(1-z)$ exists and equals 1, and $(1-z)\,\delta(z)$ exists and equals 0, the triple product $(\sum z^n)(1-z)\delta(z)$ is forbidden.

A consequence of our algebraic approach is that the product $z^{\frac{1}{2}}\delta(z)$ does not equal $1^{\frac{1}{2}}\delta(z) = \delta(z)$ – their formal power series are very different. In hindsight this 'failing' is understandable: it is artificial here to prefer the positive root of 1 over the negative root.

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

Proposition 5.1.2 *Let $\mathcal{W}$ be any vector space, and $f \in \mathcal{W}[[z_1^{\pm 1}, z_2^{\pm 1}]]$. Then $(z_1 - z_2)^N f(z_1, z_2) = 0$ for some integer $N \geq 1$, iff*

$$f(z_1, z_2) = \sum_{j=0}^{N-1} c_j(z_2)\, \partial_{z_2}^j \delta(z_1/z_2),$$

where $c_j(z_2) = \mathrm{Res}_{z_1}((z_1 - z_2)^j f(z_1, z_2)) \in \mathcal{W}[[z_2^{\pm 1}]]$.

Proof: First, $(z_1 - z_2)\, f(z_1, z_2) = 0$ iff $a_{m-1,n} = a_{m,n-1}$ $\forall m, n$, iff $a_{m,n} = a_{0,m+n}$ $\forall m, n$, iff

$$f(z_1, z_2) = \left(\sum_{n \in \mathbb{Z}} a_{0,n} z_2^{n+1} \right) \delta(z_1/z_2).$$

Also, for any $j \geq 1$,

$$\begin{aligned} (z_1 - z_2)\, \partial_{z_2}^j \left(z_2^{-1} \delta(z_1/z_2) \right) &= (z_1 - z_2) \sum_{n \in \mathbb{Z}} n\,(n-1) \cdots (n - j + 1) z_1^{-n-1} z_2^{n-j} \\ &= j\, \partial_{z_2}^{j-1} \left(z_2^{-1} \delta(z_1/z_2) \right). \end{aligned}$$

Hence

$$(z_1 - z_2)\, f(z_1, z_2) = \sum_{j=0}^{M} b_j(z_2)\, \partial_{z_2}^j \left(z_2^{-1} \delta(z_1/z_2) \right)$$

has general solution

$$f(z_1, z_2) = \sum_{j=0}^{M} \frac{1}{j+1} b_j(z_2)\, \partial_{z_2}^{j+1} \left(z_2^{-1} \delta(z_1/z_2) \right).$$

■

For reasons given next subsection, we call any formal distributions $a(z)$, $b(z)$ *mutually local* if $f(z_1, z_2) := [a(z_1), b(z_2)]$ satisfies the condition in Proposition 5.1.2. In a vertex algebra or VOA (Definition 5.1.3), all fields are mutually local.

We need ways to make new formal power series from old ones. First, for any $n \in \mathbb{Z}$, we define the binomial formula to hold:

$$(z_1 + z_2)^n := \sum_{k \in \mathbb{N}} \binom{n}{k} z_1^{n-k} z_2^k, \tag{5.1.4a}$$

where we define $\binom{n}{k} = n(n-1) \cdots (n - k + 1)/k!$ for any n. Equation (5.1.4a) lets us define, for any formal power series $f(z) = \sum_n a_n z^n \in \mathcal{W}[[z^{\pm 1}]]$,

$$f(z_1 + z_2) := \sum_{n \in \mathbb{Z}} \sum_{k \geq 0} a_n \binom{n}{k} z_1^{n-k} z_2^k \in \mathcal{W}\left[\left[z_1^{\pm 1}, z_2\right]\right]. \tag{5.1.4b}$$

Paradox 5.2 *Expand $(1 - z)^{-1}$ in a formal series in z to get $\sum_{n \geq 0} z^n$, and $(1 - z)^{-1} = -z^{-1}(1 - z^{-1})^{-1}$ in a formal series in z^{-1} to get $-\sum_{n<0} z^n$. Subtract these equal expressions; we presumably should get 0, but we actually get $\delta(z)$. Similarly, applying (5.1.4a) to $(1 + z)^{-1} = (z + 1)^{-1}$ again gives us the contradiction $0 = \delta(z)$.*

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

The analytic explanation is that the left expansions in Paradox 5.2 converge only for $|z| < 1$, while the second converges for $|z| > 1$, so it would be naive to expect their formal difference to be 0. We see from this that *it really matters in which variable we expand rational functions*. The seemingly harmless (5.1.4a) is actually a *convention* saying that we'll expand in positive powers of the second variable. For instance, at first glance

$$z_0^{-1}\delta\left(\frac{z_1 - z_2}{z_0}\right) - z_0^{-1}\delta\left(\frac{z_2 - z_1}{-z_0}\right) = z_2^{-1}\delta\left(\frac{z_1 - z_0}{z_2}\right) \tag{5.1.5}$$

is nonsense; it only holds if you expand the terms in positive powers of z_2, z_1 and z_0 respectively. A rational function by itself does not define a unique formal power series. When we need to be explicit, we write $\iota_z(f)$ to expand a rational function f in positive powers of z (i.e. for expanding it about $z = 0$). For example,

$$\iota_z\left(\frac{1}{w - z}\right) - \iota_{z^{-1}}\left(\frac{1}{w - z}\right) = \delta(z/w).$$

Recall the operator product expansion (OPE) of quantum fields (4.3.2), introduced to interpret pointwise products. Here we can study this more explicitly. For most pairs $a(z), b(z) \in (\mathrm{End}\,\mathcal{V})[[z^{\pm 1}]]$, the naive product $a(z)\,b(z)$ will not be algebraically defined. It is easy to prove directly from Proposition 5.1.2 (see theorem 2.3 of [**330**]) that if $(z_1 - z_2)^N [a(z_1), b(z_2)] = 0$, then

$$a(z_1)\,b(z_2) = \sum_{j=0}^{N-1} \frac{c^j(z_2)}{(z_1 - z_2)^{j+1}} + {:}a(z_1)\,b(z_2){:} \tag{5.1.6a}$$

separates $a(z_1)b(z_2)$ into its singular and regular parts, where

$${:}a(z_1)\,b(z_2){:} := \left(\sum_{n\geq 0} a_n z_1^n\right) b(z_2) + b(z_2)\left(\sum_{n<0} a_n z_1^n\right), \tag{5.1.6b}$$

$$c^j_{(k)}(z_2) = \sum_{\ell=0}^{N-1} \frac{j!}{\ell!\,(j-\ell)!} a_{(j-\ell-k)}\, b_{(\ell)}. \tag{5.1.6c}$$

By $1/(z_1 - z_2)^{j+1}$ in (5.1.6a) we mean to expand z_2 in powers from $-j$ to ∞. The point of (5.1.6a) is that the *normal-ordered product* (5.1.6b) is algebraically defined even at $z_1 = z_2$ (Question 5.1.6) so any singular behaviour of $a(z_1)\,b(z_2)$ as $z_1 \to z_2$ is captured by the finitely many series c^j. Equations (5.1.6) are the desired relation in CFT between the singular part of the OPE of quantum fields and the commutators of modes, mentioned in Section 4.3.2. The clarity that vertex algebras bring to quantum field theory (especially CFT) alone makes its definition worth all the pain.

5.1.3 Axioms

We are now prepared to introduce the important new structure called vertex operator algebras (VOAs). Although VOAs are natural from the CFT perspective and appear to be an important and rapidly developing area in mathematics, their definition is difficult and nontrivial examples are not easy to find.

A VOA is an infinite-dimensional graded vector space $\mathcal{V} = \oplus_{n \geq 0} \mathcal{V}_n$ with infinitely many bilinear products $u *_n v$ respecting the grading (in particular $\mathcal{V}_k *_n \mathcal{V}_\ell \subseteq \mathcal{V}_{k+\ell-n-1}$), obeying infinitely many constraints. We can collect all these products into one generating function: to each $u \in \mathcal{V}$ associate the formal power series (a *vertex operator*)

$$Y(u, z) := \sum_{n \in \mathbb{Z}} u_{(n)} z^{-n-1} \in (\mathrm{End}\,\mathcal{V})[[z^{\pm 1}]].$$

For each $u \in \mathcal{V}$, the coefficients $u_{(n)}$ (called modes (5.1.3c)) are maps from $\mathcal{V}$ to $\mathcal{V}$. The product $u *_n v$ is now written $u_{(n)} v := u_{(n)}(v)$. The bilinearity of $*_n$ translates into two things: that $u \mapsto Y(u, z)$ is linear, and that each function $v \mapsto u_{(n)} v$ is itself linear (i.e. $u_{(n)}$ is an endomorphism of $\mathcal{V}$).

Definition 5.1.3 (a) *Let $\mathcal{V}$ be a graded vector space $\mathcal{V} = \oplus_{n=-\infty}^{\infty} \mathcal{V}_n$ such that each subspace $\mathcal{V}_n$ is finite-dimensional. Suppose we have a linear assignment $u \mapsto Y(u, z) = \sum_{n \in \mathbb{Z}} u_{(n)} z^{-n-1}$ from $\mathcal{V}$ into* $(\mathrm{End}\,\mathcal{V})[[z^{\pm 1}]]$ *and a distinguished vector $\mathbf{1} \in \mathcal{V}$ in $\mathcal{V}_0$, obeying the following properties $\forall u, v \in \mathcal{V}$:*

VA1. (grading) *For $u \in \mathcal{V}_k$, $u_{(n)}$ is a linear map from $\mathcal{V}_\ell$ into $\mathcal{V}_{k+\ell-n-1}$;*
VA2. (vacuum) *$Y(\mathbf{1}, z)$ is the identity (i.e. $\mathbf{1}_{(n)} v = \delta_{n,-1} v$);*
VA3. (state-field correspondence) *$Y(u, 0)\mathbf{1}$ exists and equals u;*
VA4. (locality) *$(z_1 - z_2)^M [Y(u, z_1), Y(v, z_2)] = 0$ for some integer $M = M(u, v)$;*
VA5. (regularity) *there is an $N = N(u, v)$ such that $u_{(n)} v = 0$ for all $n \geq N$.*

Any such triple $(\mathcal{V}, Y, \mathbf{1})$ is called a vertex algebra. *The distributions $Y(u, z)$ are called* vertex operators, *and the vector $\mathbf{1}$ is called the* vacuum.
(b) *A vertex algebra $(\mathcal{V}, Y, \mathbf{1})$ is called a* vertex operator algebra (VOA) *if there is a distinguished vector $\omega \in \mathcal{V}_2$ such that*

VOA1. (conformal symmetry) *$L_n := \omega_{(n+1)}$ forms a $\mathfrak{Vir}$-module, whose central term C in* (3.1.5) *acts as $c\, id_{\mathcal{V}}$ for some $c \in \mathbb{C}$;*
VOA2. (conformal weight) *$L_0 v = nv$ whenever $v \in \mathcal{V}_n$;*
VOA3. (translation generator) *$Y(L_{-1} v, z) = \partial_z Y(v, z)$;*
VOA4. (CFT type) *$\mathcal{V}_0 = \mathbb{C}\mathbf{1}$ and $\mathcal{V}_n = \{0\}$ for all $n < 0$.*

The vector ω is called the conformal vector, *and c is called the* central charge, conformal anomaly *or* rank. *The grading n of $u \in \mathcal{V}_n$ is called its* conformal weight.
(c) *A quadruple $(\mathcal{V}, Y, \mathbf{1}, \omega)$ is called a* near-VOA *if all axioms of a VOA are satisfied, except for* VOA4, *and in addition the homogeneous subspaces $\mathcal{V}_n$ are allowed to be infinite-dimensional.*

We prefer the more descriptive name 'conformal vertex algebra' to the historical 'vertex operator algebra', although it is probably too late to dislodge the latter name. We study the Virasoro algebra in Section 3.1.2, where we discuss its relation to conformal transformations. We are more interested in VOAs than vertex algebras, since the Virasoro algebra is essential for the relation of $\mathcal{V}$ to higher genus and in particular to modular functions. The central charge c is an important invariant of $\mathcal{V}$. The original axioms

[**68**] by Borcherds didn't involve $\mathfrak{Vir}$ nor require $\dim(\mathcal{V}_n) < \infty$. The conformal axioms VOA1–VOA3 were introduced in [**201**] along with the name 'vertex operator algebra'. Although VOA4 holds for most important VOAs and yields the richest theory, it is not standard and is included here for simplicity. Note though that with it, VA5 becomes redundant and can be dropped. The name 'near-VOA' is not standard; we need the notion in Section 7.2.2.

In the physics literature, the vacuum $\mathbf{1}$ is often denoted $|0\rangle$, and in place of the expansion $Y(u, z) = \sum_n u_{(n)} z^{-n-1}$ for $u \in \mathcal{V}_k$ appears the expression $\sum_n u_{\{n\}} z^{-n-k}$ (so $L_n = \omega_{\text{ins}}$). This new expansion cleans up some formulae a little; it has the disadvantage though of artificially favouring the 'homogeneous' vectors $u \in \mathcal{V}_k$.

By Proposition 5.1.2, the peculiar-looking VA4 simply says that the commutator $[Y(u, z_1), Y(v, z_2)]$ of two vertex operators is a finite linear combination of derivatives of various orders of the Dirac delta centred at $z_1 = z_2$; this powerful locality axiom is at the heart of a vertex algebra. A recommended exercise is to show that in a VOA, $M = 4$ works in VA4 for $u = v = \omega$; more generally see Question 5.1.4.

By $\mathcal{V} = \oplus \mathcal{V}_n$ here, we mean that any vector $u \in \mathcal{V}$ can be expressed as a *finite* sum $\sum_n u(n)$ of homogeneous vectors $u(n) \in \mathcal{V}_n$. To emphasise this finiteness, the notation

$$\mathcal{V} = \coprod_{n \in \mathbb{N}} \mathcal{V}_n$$

is often used. Note that in a vertex algebra, any series $Y(u, z)v$ will be a finite sum – that is, the infinite sum $Y(u, z)$ is algebraically defined (Section 5.1.2).

An immediate consequence of VA1, VA2 and VOA2 is that $\mathbf{1} \in \mathcal{V}_0$ and $\omega \in \mathcal{V}_2$ – we needn't assume these.

Let $\mathcal{V}$ be a vector space with a linear map $Y : \mathcal{V} \to \operatorname{End} \mathcal{V}$, such that $Y(u)\,Y(v) = Y(v)\,Y(u)$. Also, assume that there exists a distinguished vector $\mathbf{1} \in \mathcal{V}$ such that $Y(\mathbf{1})$ is the identity, and such that $Y(u)\,\mathbf{1} = u$ for all $u \in \mathcal{V}$. It isn't hard to identify such a structure. Given any $u, v \in \mathcal{V}$, define the 'product' $u * v$ to be the value $Y(u)\,v$. The linearity of $Y : \mathcal{V} \to \operatorname{End} \mathcal{V}$, as well as the linearity of each map $Y(u)$, yields the distributivity laws. Also, $\mathbf{1} * u = Y(\mathbf{1})\,u = I\,u = u$ and $u * \mathbf{1} = Y(u)\,\mathbf{1} = u$, so $\mathbf{1}$ is a unit. Evaluating $Y(u)\,Y(v) = Y(v)\,Y(u)$ on the right by w, gives

$$u * (v * w) = Y(u)\,(Y(v)\,w) = Y(v)\,(Y(u)\,w) = v * (u * w).$$

Substituting $w = \mathbf{1}$ gives $u * v = v * u$, that is the product is commutative. Likewise, $u * (v * w) = u * (w * v) = w * (u * v) = (u * v) * w$, so the product is associative. Thus *a vertex algebra is an analogue of a commutative associative algebra with unit, where there is a product* $u *_z v = Y(u, z)\,v$ *at each point* z *in a punctured disc*. A vertex algebra isn't as obscure as it may first look.

Theorem 5.1.4 *The following are equivalent:*

(i) *$\mathcal{V}$ is a commutative vertex algebra, i.e.* $Y(u, z_1)\,Y(v, z_2) = Y(v, z_2)\,Y(u, z_1)$ *for all* $u, v \in \mathcal{V}$;

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

(ii) $\mathcal{V} = \oplus_{n=0}^{\infty}\mathcal{V}_n$ *is a* $\mathbb{Z}$*-graded commutative associative algebra with unit and derivation, with each* $\dim(\mathcal{V}_n) < \infty$*;*
(iii) $\mathcal{V}$ *is a vertex algebra where each vertex operator* $Y(u, z)$ *involves only nonnegative powers of* z*, i.e.* $u_{(n)} = 0$ *for all* $n \geq 0$.

Proof: The equivalence (i) $\Leftrightarrow$ (ii) was essentially established in the previous paragraph. (i) $\Rightarrow$ (iii): Consider the equality

$$\sum_{n\in\mathbb{Z}} u_{(-n-1)}v\, z_1^n = Y(u, z_1)v = Y(u, z_1)Y(v, z_2)\mathbf{1}|_{z_2=0}$$
$$= Y(v, z_2)Y(u, z_1)\mathbf{1}|_{z_2=0} = \sum_{n\geq 0} v_{(-1)}u_{(-n-1)}\mathbf{1}\, z_1^n.$$

Since the expression on the right side involves nonnegative powers of z_1 only, the same must hold for the left side.

(i) $\Leftarrow$ (iii): For any power series $f(z_1, z_2) = \sum_{m,n=0}^{\infty} a_{mn} z_1^m z_2^n \in \mathcal{W}[[z_1, z_2]]$, Proposition 5.1.2 implies that $(z_1 - z_2)^M f(z_1, z_2) = 0 \Rightarrow f(z_1, z_2) = 0$, since each residue of $f(z_1, z_2)$ will be 0. Applying this to $f(z_1, z_2) = [Y(u, z_1), Y(v, z_2)]$ gives the desired result. ∎

Locality VA4 can be rewritten in the form (see Section 3.2 of [**376**])

$$z_0^{-1}\delta\left(\frac{z_1 - z_2}{z_0}\right) Y(u, z_1)Y(v, z_2) - z_0^{-1}\delta\left(\frac{z_2 - z_1}{-z_0}\right) Y(v, z_2)Y(u, z_1)$$
$$= z_2^{-1}\delta\left(\frac{z_1 - z_0}{z_2}\right) Y(Y(u, z_0)v, z_2), \tag{5.1.7a}$$

where the formal series are expanded appropriately. This embodiment of commutativity and associativity in the vertex algebra is called the *Jacobi identity* since it plays an analogous role in VOAs as the Jacobi identity plays in Lie algebras. It corresponds directly to the duality of the sphere with four points removed (namely Figure 6.3(a)). Expanding it out, the coefficient in front of $z_0^{\ell} z_1^m z_2^n$ gives *Borcherds' identity*:

$$\sum_{i\geq 0}(-1)^i \binom{\ell}{i} \left(u_{(\ell+m-i)} \circ v_{(n+i)} - (-1)^{\ell} v_{(\ell+n-i)} \circ u_{(m+i)}\right)$$
$$= \sum_{i\geq 0} \binom{m}{i} (u_{(\ell+i)}v)_{(m+n-i)}. \tag{5.1.7b}$$

Specialising (5.1.7b) to $\ell = 0$ and $m = 0$, respectively, gives us

$$[u_{(m)}, v_{(n)}] = \sum_{i\geq 0} \binom{m}{i} (u_{(i)}v)_{(m+n-i)}, \tag{5.1.7c}$$

$$(u_{(\ell)}v)_{(n)} = \sum_{i\geq 0}(-1)^i \binom{\ell}{i} \left(u_{(\ell-i)} \circ v_{(n+i)} - (-1)^{\ell} v_{(\ell+n-i)} \circ u_{(i)}\right). \tag{5.1.7d}$$

In any vertex algebra, define an endomorphism $T : \mathcal{V} \to \mathcal{V}$ by

$$Tu = u_{(-2)}\mathbf{1}. \tag{5.1.8a}$$

This is the derivation of Theorem 5.1.4(ii). Indeed, applying (5.1.7d) to it and using VA2, we get $Y(Tu, z) = \partial_z Y(u, z)$. Thus in any VOA, VOA3 says

$$L_{-1}u = u_{(-2)}\mathbf{1}. \tag{5.1.8b}$$

Moreover, (5.1.7c) tells us that any $u \in \mathcal{V}_2$ automatically obeys $[u_{(0)}, Y(v, z)] = Y(u_{(0)}v, z)$. Thus in any VOA

$$[L_{-1}, Y(u, z)] = \partial_z Y(u, z).$$

More generally, a more subtle argument (see e.g. proposition 3.1.19 of [**376**]) shows that in any vertex algebra, we have

$$Y(u, z)v = e^{zT} Y(v, -z)u.$$

These equations also allow us to compute explicitly the grading of $u_{(n)}v$ in a VOA, recovering VA1: let $u \in \mathcal{V}_k$, $v \in \mathcal{V}_\ell$, then

$$L_0(u_{(n)}v) = \omega_{(1)}(u_{(n)}(v)) = u_{(n)}(\omega_{(1)}v) + (\omega_{(1)}u)_{(n)}v + (\omega_{(0)}u)_{(n)}v = (k+\ell-n-1)u_{(n)}v.$$

Duality (5.1.7a) also implies (see section 3.8 of [**376**])

$$Y\left(u_{(-m)}v, z\right) = \frac{1}{(m-1)!} :\left(\partial_z^{m-1} Y(u, z)\right) Y(v, z):, \tag{5.1.9a}$$

$$Y\left(u_{(n)}v, z\right) = \mathrm{Res}_{z_1}(z_1 - z)^n [Y(u, z_1), Y(v, z)], \tag{5.1.9b}$$

where $m \geq 1$ and $n \geq 0$. As we see next section, this is quite useful as a way of obtaining the full VOA from a small number of generators.

Unexpectedly, modular functions arise in VOA theory through the generating functions of the dimensions of the homogeneous spaces:

$$\mathrm{tr}_{\mathcal{V}} q^{L_0} = \sum_{n=0}^{\infty} \dim \mathcal{V}_n \, q^n. \tag{5.1.10a}$$

We also see this important theme in, for example, Section 3.1.2. As in (3.1.10), a small refinement should be made: by the *graded dimension* $\chi_{\mathcal{V}}(\tau)$ of $\mathcal{V}$ we mean

$$\chi_{\mathcal{V}}(\tau) := \mathrm{tr}_{\mathcal{V}} e^{2\pi \mathrm{i}\tau\,(L_0 - c/24)} = q^{-c/24} \sum_{n=0}^{\infty} \dim \mathcal{V}_n \, q^n, \tag{5.1.10b}$$

where as always $q = e^{2\pi \mathrm{i}\tau}$. The reason for the $q \mapsto \tau$ change-of-variables here will turn out to be the same as why Gauss and Jacobi introduced it into Euler's generating function $1 + 2x + 2x^4 + 2x^9 + \cdots$: both the graded dimension of $\mathcal{V}$ and Euler's generating function are naturally associated with tori. Explanations for the now-familiar $-c/24$ shift are given in Sections 3.2.3 and 5.3.4. Incidentally, the term *character* is also used in the literature for $\chi_{\mathcal{V}}(\tau)$, but Section 5.3.3 contains our diatribe against this misnomer.

Section 1.5 illustrates the usefulness of the Killing form in Lie theory. Similarly, our VOAs all have a nondegenerate invariant bilinear form [**199**] – a bilinear pairing

$(u|v) \in \mathbb{C}$ for $u, v \in \mathcal{V}$, such that

$$(Y(u,z)v|w) = \left(v|Y(e^{zL_1}(-z^{-2})^{L_0}u, z^{-1})w\right), \qquad \forall u, v, w \in \mathcal{V}. \tag{5.1.11a}$$

That this complicated definition is right is explained in remark 5.3.3 of [**199**] and equation (54) of [**242**]. For such a form, the homogeneous spaces $\mathcal{V}_m$ are mutually orthogonal, symmetry $(u|v) = (v|u)$ holds, and we recover familiar RCFT formulae such as $(L_n u|v) = (u|L_{-n}v)$. It is known (section 3 of [**380**]) that there is a unique invariant bilinear form (up to a scalar factor), provided that $\mathcal{V}$ is *simple* (defined in Section 5.3.1) and

$$L_1\mathcal{V}_1 = 0 \tag{5.1.11b}$$

– both conditions are always satisfied by our VOAs. In this case the bilinear form restricted to each space $\mathcal{V}_n$ will be nondegenerate. The most convenient normalisation is

$$(\mathbf{1}|\mathbf{1}) = -1, \tag{5.1.12a}$$

because for this choice the bilinear form on the homogeneous space $\mathcal{V}_1$ becomes

$$(u|v) = u_1 v, \qquad \forall u, v \in \mathcal{V}_1. \tag{5.1.12b}$$

The invariant bilinear form plays an important role in CFT as well as Moonshine.

By a *vertex operator superalgebra* we mean there is a $\mathbb{Z}_2$-grading of $\mathcal{V} = \mathcal{V}_{\bar{0}} \oplus \mathcal{V}_{\bar{1}}$ into even and odd parity subspaces, and for u, v both odd the commutator in, for example, Axiom VA4 is replaced by an anti-commutator. Their basic theory is very similar to that of VOAs (see e.g. [**330**]). For instance, we write

$$\chi_{\mathcal{V}}(\tau) := \chi_{\mathcal{V}_{\bar{0}}}(\tau) - \chi_{\mathcal{V}_{\bar{1}}}(\tau).$$

Although we occasionally allude to vertex operator superalgebras (e.g. in Sections 5.4.2 and 7.3.5), we won't develop their theory.

In RCFT, $\mathcal{V}$ would be the 'Hilbert space of states' (more carefully, $\mathcal{V}$ is a dense subspace of it), and $z = e^{t+\mathrm{i}x}$ would be a local complex coordinate on a Riemann surface. L_0 generates time translations, and so its eigenvalues (the conformal weights) are energy. For each state u, the vertex operator $Y(u,z)$ is a meromorphic (chiral) quantum field. $Y(\omega, z)$ is the stress–energy tensor T. Physically, the requirement that $\mathcal{V}_n = 0$ for $n < 0$ says that the vacuum $\mathbf{1} = |0\rangle$ is the state with minimal energy. Also, $z = 0$ in VA3 corresponds to the time limit $t \to -\infty$. The most important axiom, VA4, says that quantum fields commute away from $z_1 = z_2$, and so are local. It is equivalent to the duality of chiral blocks in CFT, discussed in Sections 4.3.2, 4.4.1, 6.1.4.

In Segal's language (Section 4.4.1), $Y(u,z)$ appears quite naturally. Consider the virtual event of two strings combining to form a third. To first order (i.e. the tree-level Feynman diagram), this would correspond in Segal's language to a 'pair-of-pants', or a sphere with three punctures, two of which are negatively oriented (corresponding to incoming strings) and the other positively oriented. We can think of this as the Riemann sphere $\mathbb{C} \cup \{\infty\}$; put the punctures at ∞ (outgoing) and z and 0 (incoming). Segal's functor $\mathcal{T}$ associates with this a z-dependent homomorphism $\varphi_z : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$, where

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

$\varphi_z(u, v) = Y(u, z)v \in \mathcal{V}$. Incidentally, the symbol 'Y' *could* have been chosen because of this 'pair-of-pants' picture (time flows from the top of the 'Y' to the bottom).[1]

By VOA1, any VOA is a $\mathfrak{Vir}$-module. For most VOAs, this module is highly reducible. By a *conformal primary* v of conformal weight k we mean $L_n v = 0$ for all $n > 0$ and $L_0 v = kv$ for some k. These states are especially well behaved. Any such primary generates a highest-weight module for $\mathfrak{Vir}$, on the space spanned by the elements $L_{-n_1} \cdots L_{-n_m} v$. The VOAs we are interested in are generated by the conformal primaries together with the operators L_n, in the sense that $\mathcal{V}$ can be decomposed into a direct sum (usually infinite) of highest-weight $\mathfrak{Vir}$-modules.

Question 5.1.1. Theorem 5.1.1 actually provides a realisation for a highest-weight representation of $A_1{}^{(1)}$. Identify that representation.

Question 5.1.2. Using the notion of algebraic summability, write down an algebraic definition of $\lim_{z_1 \to z_2} F(z_1, z_2)$ valid for formal power series $F(z_1, z_2) \in \mathcal{W}[[z_1^{\pm 1}, z_2^{\pm 1}]]$ realising the intuition of substituting in $z_1 = z_2$. Prove that $\lim_{z_1 \to z_2} F(z_1, z_2)$ 'algebraically exists' iff the product $F(z_1, z_2)\,\delta(z_1/z_2)$ does, in which case $F(z_1, z_2)\,\delta(z_1/z_2) = F(z_2, z_2)\,\delta(z_1/z_2)$.

Question 5.1.3. (a) Given any formal power series $F(z) \in \mathcal{W}[[z^{\pm 1}]]$, prove that

$$e^{w \frac{\mathrm{d}}{\mathrm{d}z}} F(z) = F(z + w).$$

(b) Prove (5.1.5).

Question 5.1.4. (a) Let $\mathcal{V}$ be any VOA, and $u, v \in \mathcal{V}$. Then for any $k \in \mathbb{Z}$, prove that

$$(z_1 - z_2)^k \,[Y(u, z_1), Y(v, z_2)] = \sum_{\ell \geq 0} \frac{1}{\ell!} \left(\partial_{z_2}^{\ell} z_1^{-1} \delta(z_2/z_1)\right) Y\left(u_{(k+\ell)} v, z_2\right).$$

(b) Let $u \in \mathcal{V}_m, v \in \mathcal{V}_n$ be homogeneous vectors in any vertex algebra $\mathcal{V}$. Prove that $M(u, v) = m + n$ works in VA4.

Question 5.1.5. (a) Prove that in any vertex algebra, the vacuum $\mathbf{1}$ is translation-invariant, i.e. $T\mathbf{1} = 0$.
(b) In any VOA, verify that the span of L_{-1}, L_0, L_1 is the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$. Verify that the vacuum is invariant under it.

Question 5.1.6. Prove that for any a, b, c in a vertex algebra $\mathcal{V}$, every coefficient z^n of $:a(z)\,b(z):\,u$ involves a finite sum, and for all but finitely many negative n this sum is 0.

5.2 Basic theory

A VOA is a remarkably rich algebraic structure, with infinitely many heavily constrained products. In this section we continue to work out the easy consequences of the axioms.

[1] But it wasn't. Remarkably, the actual historical reason is that Y comes after X, and X was the name arbitrarily chosen in [**201**] for a pre-vertex operator. The symbol Y first appeared in their chapter 8; Borcherds used the symbol Q.

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

The deep role of the Virasoro algebra remains hidden in this section. We also associate VOAs with lattices and affine algebras.

5.2.1 Basic definitions and properties

For any $u \in \mathcal{V}_n$, define $o(u) = u_{(n-1)}$. Then VA1 tells us $o(u)$ preserves each grade, that is it maps each homogeneous space $\mathcal{V}_m$ to itself. In particular, every space $\mathcal{V}_n$ has an algebraic structure defined by $u \times v = o(u)\, v$. In the CFT literature, these are called the *zero-mode algebras* (because $u_{(n-1)} = u_{\{0\}}$).

Typically, the zero-mode algebras $\mathcal{V}_n$ are quite complicated. However, consider $\mathcal{V}_1$. Put $\ell = m = n = 0$ in (5.1.7b) and hit it with any $w \in \mathcal{V}$: we get $u_{(0)}(v_{(0)}w) - v_{(0)}(u_{(0)}w) = (u_{(0)}v)_{(0)}w$. If we now formally write $[xy] := x_{(0)}y$, then this becomes $[u[vw]] - [v[uw]] = [[uv]w]$, which is one of the forms of the Lie algebra Jacobi identity (1.4.1b). Thus our bracket will be anti-associative if it is anti-commutative, in which case $\mathcal{V}_1$ will be a Lie algebra. But is it anti-commutative? From (5.1.9) we get

$$u_{(n)}v = \sum_{i \geq 0} \frac{1}{i!}(-1)^{i+n+1}(L_{-1})^i (v_{(n+i)}u) \tag{5.2.1}$$

so $u_{(0)}v \equiv -v_{(0)}u \pmod{L_{-1}\mathcal{V}}$. However, from VA1, VOA4 and Question 5.1.5, we get

$$(L_{-1}\mathcal{V})_1 = L_{-1}(\mathcal{V}_0) = L_{-1}(\mathbb{C}\mathbf{1}) = \{0\}.$$

Thus, in any VOA, $\mathcal{V}_1$ is a finite-dimensional Lie algebra. Each homogeneous space $\mathcal{V}_n$ is a module for $\mathcal{V}_1$.

Given any $u, v \in \mathcal{V}_1$, $u_{(1)}v \in \mathcal{V}_0 = \mathbb{C}\mathbf{1}$, and so define $(u|v) \in \mathbb{C}$ by $(u|v)\,\mathbf{1} = u_{(1)}v$. From (5.2.1), $(u|v) = (v|u)$, so $(\star|\star)$ defines a symmetric bilinear form on $\mathcal{V}_1$. We would like $(\star|\star)$ to respect the Lie algebra structure, that is be $[\star\star]$-invariant. We compute from (5.1.7) and VA2

$$([uv]|t)\,\mathbf{1} = -v_{(0)}((u|t)\mathbf{1}) + (u|[vt])\,\mathbf{1} = (u|[vt])\,\mathbf{1}, \tag{5.2.2}$$

that is $([uv]|t) = (u|[vt])$ and $(\star|\star)$ is indeed $[\star\star]$-invariant. Of course, this bilinear form is identical with that of (5.1.12b), and so provided (5.1.11b) is satisfied, it will be nondegenerate.

The existence of this bilinear form severely restricts the possibilities for the Lie algebra $\mathcal{V}_1$. Such Lie algebras are called *self-dual* and are precisely those for which the Sugawara construction (3.2.15) works. They are studied, for instance, in [**415**], [**189**], [**384**] – see also example 2.1 in [**156**]. If we also demand that the VOA be *weakly rational* (Definition 5.3.2), then $\mathcal{V}_1$ will be *reductive* (i.e. a direct sum of simple and trivial Lie algebras) [**156**].

The affinisation $\mathcal{V}_1{}^{(1)}$ of the Lie algebra $\mathcal{V}_1$ also appears naturally in the VOA $\mathcal{V}$. In particular, the modes $u_{(n)}$, for all $u \in \mathcal{V}_1$ and $n \in \mathbb{Z}$, have the commutators

$$u_{(m)} \circ v_{(n)} - v_{(n)} \circ u_{(m)} = ([u, v])_{(m+n)} + m\,(u|v)\delta_{m+n,0}\mathbf{1}_{(-1)}.$$

Thus these $u_{(n)}$, together with centre $\mathbb{C}\mathbf{1}_{(-1)}$ and derivation L_{-1}, span a $\mathcal{V}_1{}^{(1)}$-module.

More generally, in Section 7.2.2 we need to obtain a Lie algebra from a near-VOA $\mathcal{V}$. As before, we obtain a Lie algebra structure on $\mathcal{V}/L_{-1}\mathcal{V}$, and it has an invariant bilinear form if we restrict to $\mathcal{V}_1/L_{-1}\mathcal{V}_0$. In the situations we will be interested in, this algebra is too large, but it can be reduced as follows. Define

$$\mathcal{P}V_n := \{u \in \mathcal{V}_n \mid L_m u = 0 \text{ for all } m > 0\}, \tag{5.2.3}$$

i.e. the conformal primaries with conformal weight n. Then a straightforward calculation verifies that $\mathcal{P}V_1/(L_{-1}\mathcal{V}_0 \cap \mathcal{P}V_1)$ is itself a Lie algebra, with the usual bracket. Through the map $u \mapsto u_{(0)}$, this Lie algebra acts on $\mathcal{V}$ and this action commutes with that of L_m. These associations of Lie algebras to (near-)VOAs are due to Borcherds [**68**].

By an *automorphism* (or symmetry) α of a VOA $\mathcal{V}$ we mean an invertible linear map $\alpha : \mathcal{V} \to \mathcal{V}$ obeying

$$\alpha(Y(u, z)\, v) = Y(\alpha(u), z)\, \alpha(v),$$

together with $\alpha(\mathbf{1}) = \mathbf{1}$ and $\alpha(\omega) = \omega$. This is how group theory arises in VOAs. The automorphism group can be finite (e.g. $\mathrm{Aut}(V^\natural) = \mathbb{M}$) or infinite (e.g. $\mathrm{Aut}(\mathcal{V}(\Lambda)) \cong (\mathbb{R}^\times)^{24} \rtimes Co_0$), but it can be finite only if $\mathcal{V}_1 = 0$ (Question 5.2.2). Conjecturally, at least when $\mathcal{V}$ is sufficiently nice, $\mathrm{Aut}(\mathcal{V})$ will be finite if (and only if) $\mathcal{V}_1 = 0$.

Similar arguments (Question 5.2.3) show that when $\mathcal{V}_1 = 0$, $\mathcal{V}_2$ is a commutative non-associative algebra with product $u \times v := u_{(1)}v \in \mathcal{V}_2$ and identity element $\frac{1}{2}\omega$. Moreover, an 'associative' bilinear form can be defined on $\mathcal{V}_2$ (Question 5.2.3). For example, the Moonshine module $V^\natural$ satisfies $V_1^\natural = 0$ (Section 7.2.1), and $V_2^\natural$ is none other than the Griess algebra [**263**] extended by an identity element.

The operators $u_{(0)}$, $u \in \mathcal{V}$, are derivations (i.e. infinitesimal automorphisms) of $\mathcal{V}$, that is

$$[u_{(0)}, Y(v, z)] = Y(u_{(0)}(v), z), \tag{5.2.4}$$

and so $\exp(u_{(0)})$ is an automorphism of $\mathcal{V}$ if it is defined. This is important to the BRST cohomology construction (Question 5.2.4), borrowed from string theory.

5.2.2 Examples

Unlike more classical algebraic structures, VOAs are notorious for having no easy examples. In this section we construct families of them, in the most direct way possible. This explicitness has the drawback of making the constructions seem *ad hoc*. The reader interested in seeing the naturality of these constructions should consult the more sophisticated treatments in, for example, [**330**], [**376**].

Recall from (3.2.12a) the oscillator algebra $\mathfrak{g} = \mathfrak{u}_1{}^{(1)}$, with basis consisting of a_n, $n \in \mathbb{Z}$, together with the central term C. For any nonzero level $k \in \mathbb{C}$, we get a 'vacuum module' $\mathcal{V}(\mathfrak{g}, k)$ defined to have basis consisting of the formal combinations

$$a_{-m_1} \cdots a_{-m_r}\, \mathbf{1} \tag{5.2.5a}$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

for $r \geq 0$, where $m_1 \geq m_2 \geq \cdots \geq m_r \geq 1$. Using the actions $C\mathbf{1} = k\mathbf{1}$, $a_n\mathbf{1} = 0$ for $n \geq 0$, we see $\mathcal{V}(\mathfrak{g}, k)$ has a $\mathfrak{u}_1{}^{(1)}$-module structure. Of course $\mathfrak{u}_1$ embeds into $\mathcal{V}(\mathfrak{g}, k)$ by $x \in \mathfrak{u}_1$ goes to $xa_{-1}\mathbf{1}$.

We claim that $\mathcal{V}(\mathfrak{u}_1{}^{(1)}, k)$ has a VOA structure, for $k \neq 0$. For the assignment of vertex operators, it suffices by (5.1.9) to define $Y(x, z)$ for $x \in \mathfrak{u}_1$: we get the 'current'

$$Y(xa_{-1}\mathbf{1}, z) := x \sum_{n \in \mathbb{Z}} a_n z^{-n-1}. \tag{5.2.5b}$$

All other vertex operators follow from (5.1.9). For example, for $m \geq 1$,

$$Y(a_{-m}\mathbf{1}, z) = \frac{1}{(m-1)!} \partial_z^{m-1} \sum_n a_n z^{-n-1}.$$

The unique singular term in the OPE (5.1.6) of the basic current with itself is

$$Y(a_{-1}\mathbf{1}, z_1) Y(a_{-1}\mathbf{1}, z_2) = \frac{C}{(z_1 - z_2)^2} + \cdots \tag{5.2.5c}$$

The Sugawara construction (3.2.14a) says here that the conformal vector is

$$\omega = \frac{1}{2k} a_{-1} a_{-1} \mathbf{1}, \tag{5.2.5d}$$

which makes $\mathcal{V}(\mathfrak{g}, k)$ into a (highly reducible) $\mathfrak{Vir}$-module with central charge $c = 1$. We also get the commutation relations

$$[L_m, a_n] = -n a_{m+n}. \tag{5.2.5e}$$

In particular, the grading, given as we know by L_0, assigns the basis vector (5.2.5a) conformal weight $m_1 + \cdots + m_r$, so the current (5.2.5b) has conformal weight 1.

There is an obvious generalisation to any abelian Lie algebra $\overline{\mathfrak{h}} = \mathbb{C}^d$ with a choice of nondegenerate inner product on the space $\overline{\mathfrak{h}}$ (this defines the central term of the affine bracket (3.2.12a)). Namely, replace a with an orthonormal basis $a^1, \ldots, a^d$ of $\mathbb{C}^d$; the basis of the VOA is built up from all the operators a^i_{-n} as in (5.2.5a). These VOAs $\mathcal{V}(\overline{\mathfrak{h}}^{(1)}, k)$ are often called *Heisenberg VOAs*, because $\overline{\mathfrak{h}}^{(1)}$ is a Heisenberg algebra (i.e. a Lie algebra $\mathfrak{h}$ with $[\mathfrak{h}, \mathfrak{h}]$ equal to the centre of $\mathfrak{h}$). It turns out (Question 5.2.6) that the VOA $\mathcal{V}(\overline{\mathfrak{h}}^{(1)}, k)$ is independent of the choice of level k, provided $k \neq 0$, and also the choice of inner product, provided it is nondegenerate. We will let $\mathcal{V}(\mathbb{C}^n)$ denote the Heisenberg VOA with level $k = 1$ and standard inner product on the abelian Lie algebra $\overline{\mathfrak{h}} = \mathbb{C}^n$.

The generalisation to any affine algebra $\mathfrak{g} = \overline{\mathfrak{g}}^{(1)}$ [**68**], [**201**], [**202**], [**384**] is also straightforward. To any level $k \in \mathbb{C}$, $k \neq -h^\vee$ ($h^\vee$ the dual Coxeter number of $\overline{\mathfrak{g}}$), we get a natural VOA structure $\mathcal{V}(\mathfrak{g}, k)$ on the Verma module $M(k\omega_0)$ associated with highest weight $k\omega_0$, with central charge (3.2.9c). For example, from the Sugawara construction (3.2.15), the conformal vector is

$$\omega = \frac{1}{2(k + h^\vee)} \sum_i a^i_{(-1)} b^i_{(-1)} \mathbf{1}, \tag{5.2.6}$$

where $a^i, b^j \in \overline{\mathfrak{g}}$ are bases for $\overline{\mathfrak{g}}$, dual with respect to the Killing form on $\overline{\mathfrak{g}}$: $(a^i|b^j) = \delta_{ij}$. Any pair of dual bases give the same ω – the element $\frac{1}{2}\sum_i a^i b^i$ in the universal enveloping algebra $U(\overline{\mathfrak{g}})$ is simply the Casimir operator, and lies in the centre of $U(\overline{\mathfrak{g}})$. The only important difference here from the Heisenberg VOA is that sometimes there are 'null vectors', that is the Verma module $M(\lambda)$ may not be irreducible. In fact, maximal numbers of null vectors is the signature of the most interesting levels, namely $k \in \mathbb{N}$. We should quotient out all null vectors: by $\mathcal{V}(\mathfrak{g}, k)$ we mean the VOA structure (5.2.6) on the irreducible $\mathfrak{g}$-module $L(k\omega_0)$ defined in Section 3.2.3. Most interesting (because of its representation theory – Section 5.3) is $\mathcal{V}(\mathfrak{g}, k)$ when $k \in \mathbb{N}$, what we will call *integrable affine VOAs*.

The Lie algebra $\mathcal{V}_1$ associated with these affine algebra VOAs $\mathcal{V} = \mathcal{V}(\mathfrak{g}, k)$ is isomorphic to the reductive Lie algebra $\overline{\mathfrak{g}}$. Its affinisation, defined last subsection, equals $\mathfrak{g}$.

The forbidden level $k = -h^\vee$ is called the *critical level* and is very interesting in its own way. The conformal structure is lost (the conformal vector (5.2.6) won't exist), but the Möbius symmetry remains. The affine algebra vertex algebras at critical level have a highly nontrivial centre, and through it are related to geometric Langlands (see e.g. the discussion in section 17.4 of [**197**]). For this reason, it should be interesting to study it from the context of CFT.

Another relatively simple class of VOAs are associated with lattices [**68**], [**201**]. The simplest possibility is an n-dimensional positive-definite lattice L (Section 1.2.1), all of whose inner products $a \cdot b$ are *even* integers. By $\mathbb{C}\{L\}$ we mean the (infinite-dimensional) group algebra of the additive group L, written using formal exponentials: for each vector $v \in L$, we have a basis vector e^v of $\mathbb{C}\{L\}$, which multiply by $e^u e^v = e^{u+v}$. Let $\overline{\mathfrak{h}} = \mathbb{C} \otimes L \cong \mathbb{C}^n$ be the underlying complex vector space of L, interpreted as an abelian Lie algebra. It inherits the inner product of L. The underlying vector space of the VOA $\mathcal{V}(L)$ is $\mathcal{V}(\overline{\mathfrak{h}}) \otimes \mathbb{C}\{L\}$, where $\mathcal{V}(\overline{\mathfrak{h}})$ is the Heisenberg VOA constructed earlier. The vertex operator $Y(h \otimes 1, z)$, for $h \in \mathcal{V}(\overline{\mathfrak{h}})$, equals the vertex operator $Y(h, z)$ in $\mathcal{V}(\overline{\mathfrak{h}})$. Less clear is how to define the vertex operators $Y(1 \otimes e^\alpha, z)$, but once we know how the affine algebra $\overline{\mathfrak{h}}^{(1)}$ acts on the group algebra $\mathbb{C}\{L\}$, they will be heavily constrained by the OPEs (5.1.6a). Define $ht^m.e^\alpha = (h|\alpha)\,\delta_{m,0}e^\alpha$, for any $h \in \overline{\mathfrak{h}}$ and $\alpha \in L$, where we identify $\alpha \in L$ with the corresponding vector in $\overline{\mathfrak{h}} = \mathbb{C} \otimes L$. Then the OPE (5.1.6a) tells us (as usual displaying only the singular terms)

$$h(z_1)\,Y(1 \otimes e^\alpha, z_2) = \frac{(h|\alpha)}{z_1 - z_2} Y(1 \otimes e^\alpha, z_2) + \cdots$$

From this, and the pairwise locality of these vertex operators, we derive the formula

$$Y(1 \otimes e^\alpha, z) = e^\alpha \exp\left(-\sum_{j<0} \frac{z^{-j}}{j}\alpha_j\right) \exp\left(-\sum_{j>0} \frac{z^{-j}}{j}\alpha_j\right) z^{\alpha_0}.$$

In the usual way, this determines all vertex operators $Y(h \otimes e^\alpha, z)$. The vacuum is $\mathbf{1} \times 1$ and conformal vector ω is $\omega \otimes 1$; the central charge c though now equals the dimension n of L. The vectors $h \otimes 1$ for $h \in \overline{\mathfrak{h}}$ have conformal weight 1, while $\mathbf{1} \otimes e^\alpha$ have conformal weight $(\alpha|\alpha)/2$.

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

The construction is the same for any even positive-definite lattice L (i.e. all norm-squareds are even), except that the group algebra $\mathbb{C}\{L\}$ should be 'twisted' so that $e^\alpha e^\beta = (-1)^{(\alpha|\beta)} e^\beta e^\alpha$. If instead L is an *odd* positive-definite lattice (i.e. an integral lattice with some vectors of odd norm-square), the same construction yields vertex operator superalgebras (i.e. VOAs except the locality axiom VA4 can involve anti-commutators). For example, $L = \mathbb{Z}$ describes two fermions.

Repeating this construction for an *indefinite* even lattice L will yield a near-VOA. To see this, note that the conformal weight of $\mathbf{1} \otimes e^\alpha$ is $(\alpha|\alpha)/2$. If we regard $\mathcal{V}(L)$ as graded by L rather than by $\mathbb{Z}$, we obtain a grading into finite-dimensional subspaces.

There are several ways to construct new VOAs from old ones. For example, one can take the direct sum of VOAs with equal central charge (this doesn't change the central charge), or tensor products of arbitrary VOAs (the central charge adds) – see section 3.12 of [**376**]. The *orbifold construction* mods out by discrete symmetries: for a finite group G of symmetries of a VOA $\mathcal{V}$, let $\mathcal{V}^G$ denote the subspace of $\mathcal{V}$ fixed by G; then $\mathcal{V}^G$ is a vertex operator subalgebra of $\mathcal{V}$ – see Sections 4.3.4 and 5.3.6.

Finally, *Goddard–Kent–Olive (GKO) coset construction* [**250**] mods out by continuous symmetries. In particular, let $(\mathcal{V}, Y, \mathbf{1}, \omega)$ and $(\mathcal{V}', Y, \mathbf{1}, \omega')$ be VOAs with $\mathcal{V}' \subset \mathcal{V}$. So $\mathcal{V}'$ would be a vertex operator subalgebra of $\mathcal{V}$ except the conformal vectors need not be equal. Assume, however, that $\omega' \in \mathcal{V}_2$ and $L_1\omega' = 0$. The coset construction finds a VOA structure on the centraliser

$$\begin{aligned} C_{\mathcal{V}}(\mathcal{V}') &:= \{v \in \mathcal{V} \mid [Y(v, z_1), Y(u, z_2)] = 0 \ \forall u \in \mathcal{V}'\} \\ &= \{v \in \mathcal{V} \mid v_n u = 0 \ \forall u \in \mathcal{V}', n \in \mathbb{Z}\}. \end{aligned} \tag{5.2.7}$$

The equality in (5.2.7) follows from Question 5.2.5. Then $(C_{\mathcal{V}}(\mathcal{V}'), Y, \mathbf{1}, \omega - \omega')$ is a VOA with central charge $c - c'$. In the VOA language, this was developed in [**202**]; see also the lucid treatment in section 3.11 of [**376**].

A conjecture of Moore and Seiberg [**436**], [**437**] states that every RCFT arises from orbifold and coset constructions applied to lattice and affine algebra theories (generously enough interpreted). They speculate that this would be the analogue here of Tannaka–Krein duality (Section 1.6.2). We seem a long way from proving this optimistic guess, even in a more limited context of sufficiently nice VOAs.

The most famous VOA is the *Moonshine module* $V^\natural$, constructed in 1984 in a *tour de force* by Frenkel–Lepowsky–Meurman [**200**]. It has central charge $c = 24$, with $V^\natural = V_0^\natural \oplus V_1^\natural \oplus V_2^\natural \oplus \cdots$, where $V_0^\natural = \mathbb{C}1$ is one-dimensional, $V_1^\natural = \{0\}$ is trivial and $V_2^\natural = (\mathbb{C}\omega) \oplus$ (Griess algebra) is $(1 + 196883)$-dimensional. Its automorphism group is precisely the Monster $\mathbb{M}$. Thus each graded piece $V_n^\natural$ is a finite-dimensional $\mathbb{M}$-module. It has graded dimension J, and is the space (0.3.1) lying in the heart of Conway and Norton's Monstrous Moonshine (see Sections 4.3.4 and 7.2.1).

A formal parallel exists between integral lattices L and VOAs $\mathcal{V}$ [**201**], [**248**]. The dimension n of L corresponds to the central charge c of $\mathcal{V}$. An even lattice corresponds to a VOA while an odd lattice corresponds to a vertex operator superalgebra. As we see in the next section, the determinant $|L|$ relates to a measure of how many irreducible modules the VOA has. The norm-$\sqrt{2}$ vectors in L correspond to the vectors in

$\mathcal{V}_1$ – indeed, the norm-$\sqrt{2}$ vectors in a lattice L are special because they generate a Coxeter subgroup in Aut(L); the vectors in $\mathcal{V}_1$ are special because they generate a continuous subgroup (a Lie group) of Aut($\mathcal{V}$). In particular, the Leech lattice Λ and the Moonshine module $V^\natural$ play analogous roles (Section 7.2.1). Analogies of these kinds are always useful in their easy role as squirrels. The battle-cry 'Why invent when one can profitably copy?' is heard not only in Hollywood.

Question 5.2.1. Let $\mathcal{V}$ be a VOA, and let a finite group G act as automorphisms on $\mathcal{V}$, so each space $\mathcal{V}_n$ is a (finite-dimensional) G-module. Prove that for each n, $\mathcal{V}_n$ is a G-submodule of $\mathcal{V}_{n+1}$. (*Hint*: Consider the map L_{-1}.)

Question 5.2.2. In any VOA, define a map $e^{o(v)} : \mathcal{V} \to \mathcal{V}$ for each $v \in \mathcal{V}_1$, and show that for $v \neq 0$ it defines a nontrivial automorphism of $\mathcal{V}$. Verify that $e^{\mathcal{V}_1}$ generates a normal subgroup of Aut($\mathcal{V}$), and hence that Aut($\mathcal{V}$) will be uncountable if $\mathcal{V}_1 \neq 0$.

Question 5.2.3. Suppose a VOA $\mathcal{V}$ has $\mathcal{V}_1 = 0$. For $u, v \in \mathcal{V}_2$, define $u \times v = u_1 v$. Verify that $\mathcal{V}_2$ is commutative with this product, with identity $\omega/2$. Define a $\mathbb{C}$-valued bilinear form on $\mathcal{V}_2$ and discover how it is compatible with $\times$.

Question 5.2.4. Let $\mathcal{V}$ be a vertex algebra, and suppose $u \in \mathcal{V}_k$ satisfies $(u_{(0)})^2 = 0$. Prove that $\mathcal{V}^{(u)} = \ker u_{(0)}/\operatorname{im} u_{(0)}$ is itself a vertex algebra.

Question 5.2.5. Prove that $[Y(u, z_1), Y(v, z_2)] = 0$ iff $u_n v = 0$ for all $n \geq 0$.

Question 5.2.6. (a) Suppose both V, V' are complex n-dimensional vector spaces together with choices of nondegenerate inner-products. Verify that the Heisenberg VOAs $\mathcal{V}(V^{(1)}, k)$ and $\mathcal{V}(V'^{(1)}, k')$ are isomorphic as VOAs, provided only that k, k' are both nonzero.
(b) Let $\mathfrak{g} = \overline{\mathfrak{g}}^{(1)}$ be the nontwisted affine algebra associated with a simple finite-dimensional Lie algebra $\overline{\mathfrak{g}}$, and let $k \neq k'$ be two complex numbers, both distinct from the critical level $-h^\vee$. When are the affine algebra VOAs $\mathcal{V}(\mathfrak{g}, k)$ and $\mathcal{V}(\mathfrak{g}, k')$ isomorphic as VOAs?
(c) Let L, L' be two positive-definite lattices, all of whose inner-products $u \cdot v$ are even integers. When are the lattice VOAs $\mathcal{V}(L)$ and $\mathcal{V}(L')$ isomorphic as VOAs?

Question 5.2.7. Find an even indefinite lattice L such that the near-VOA $\mathcal{V}(L)$ has finite-dimensional homogeneous spaces $\mathcal{V}(L)_n$ for all $n \in \mathbb{Z}$.

5.3 Representation theory: the algebraic meaning of Moonshine

We know affine algebras have modules (namely the integrable ones) with interesting characters. However they have many other modules that are far less interesting, even if we restrict to highest weight ones with positive integer level. What general principle distinguishes the interesting ones from the generic? Of the uncountably many level $k \in \mathbb{N}$ highest-weight $X_r{}^{(1)}$-modules, the integrable ones are precisely those that are unitary. It is tempting then to guess that unitarity is the key principle. However, the reason to doubt its fundamental role is that there are RCFTs (e.g. the Yang–Lee model with central charge $c = -22/5$, see section 7.4.1 of [**131**]) whose graded dimensions obey all of the properties the affine characters do, but whose modules are not unitary.

The key feature possessed by the integrable affine modules is that they are unexpectedly small – that is, the null vectors in the associated Verma module, all of which are quotiented away, are maximally numerous. In other words, they are also modules of a sufficiently nice ('rational') VOA. The appearance of an affine algebra here is not directly significant, rather it is the appearance of that rational VOA. Modules of those VOAs may or may not be unitary. *VOAs serve as the unifying mathematics underlying the modules singled out by Moonshine.*[2]

The *raison d'être* of VOAs are their modules, and in Moonshine we are primarily interested in their graded dimensions and characters. It is to this important topic – *the algebraic meaning of Moonshine* – that we finally turn. See also [**199**], [**376**].

5.3.1 Fundamentals

A module of a VOA $\mathcal{V}$ is a vector space on which $\mathcal{V}$ acts, in such a way that this action preserves all possible structure. More precisely:

Definition 5.3.1 [199] *Let $\mathcal{V}$ be a VOA. A* weak $\mathcal{V}$-module (M, Y_M) *is an $\mathbb{N}$-graded vector space $M = \oplus_{n\in\mathbb{N}} M_{[n]}$, and a linear map $Y_M : \mathcal{V} \to \mathrm{End}\, M[[z^{\pm 1}]]$, written $Y_M(u, z) = \sum_{n\in\mathbb{Z}} u_{(n)} z^{-n-1}$, such that for any $u \in \mathcal{V}_k$, the mode $u_{(n)}$ is a linear map from $M_{[\ell]}$ into $M_{[k+\ell-n-1]}$,*

$$Y_M(\mathbf{1}, z) = id_M, \tag{5.3.1a}$$

$$z_0^{-1}\delta\left(\frac{z_1 - z_2}{z_0}\right) Y_M(u, z_1) Y_M(v, z_2) - z_0^{-1}\delta\left(\frac{z_2 - z_1}{-z_0}\right) Y_M(v, z_2) Y_M(u, z_1)$$
$$= z_2^{-1}\delta\left(\frac{z_1 - z_0}{z_2}\right) Y_M(Y(u, z_0)v, z_2), \tag{5.3.1b}$$

where each mode $u_{(n)}$ operates on M. The $Y_M(u, z)$ are also called vertex operators. *A weak $\mathcal{V}$-module (M, Y_M) is called a* $\mathcal{V}$-module *if in addition it comes with a grading $M = \oplus_{\alpha\in\mathbb{C}} M_\alpha$, with $M_\alpha = 0$ for* $\mathrm{Re}(\alpha)$ *sufficiently negative, obeying*

$$M_\alpha = \{x \in M \mid L_0 x = \alpha x\} \tag{5.3.1c}$$

(the eigenvalue α is again called the conformal weight *of $y \in M_\alpha$), and all homogeneous spaces M_α are finite-dimensional.*

We are interested in $\mathcal{V}$-modules. For the VOAs of interest to us (see Definition 5.3.2), the conformal weights are always rational (hence the name). Definition 5.3.1 uses the Jacobi identity (5.1.7a) rather than the simpler locality VA4 because, although locality and the Jacobi identity are equivalent for VOAs, for modules the Jacobi identity is stronger (see chapter 4 of [**376**]).

[2] Victor Kac expresses a related position by isolating locality as the key principle [**329**].

As before, the modes $L_n = \omega_{(n+1)}$ of the conformal vector $\omega \in \mathcal{V}$ yield on M a representation of the Virasoro algebra $\mathfrak{Vir}$, with the same central charge c as $\mathcal{V}$. In analogy with (5.1.10b), the *graded dimension* of a $\mathcal{V}$-module M is defined to be

$$\chi_M(\tau) := \mathrm{tr}_M e^{2\pi \mathrm{i}\tau\,(L_0 - c/24)} = q^{-c/24} \sum_{\alpha \in \mathbb{C}} \dim M_\alpha\, q^\alpha . \tag{5.3.2}$$

It is fundamental to the whole theory that these χ_M are modular, at least for 'nice' $\mathcal{V}$ and M (see Theorem 5.3.8 below). The automorphism group of $\mathcal{V}$ acts on each homogeneous space M_α – that is, each M_α carries a representation of $\mathrm{Aut}(\mathcal{V})$, and so the q-coefficients of $\chi_M(\tau)$ are dimensions of $\mathrm{Aut}(\mathcal{V})$-representations (famous examples being (0.2.1)).

It is straightforward [**199**], [**376**] to write down the definitions of $\mathcal{V}$-module homomorphism, direct sum of $\mathcal{V}$-modules, submodule, irreducible module (no nontrivial submodule), completely reducible module (i.e. M can be written as a direct sum of irreducible $\mathcal{V}$-modules), etc. Invariant bilinear forms can be defined for modules as in (5.1.11a), and have analogous properties [**199**], [**380**].

The easiest example of a $\mathcal{V}$-module, of course, is $\mathcal{V}$ itself, called the *adjoint* module. If $\mathcal{V}$ is irreducible as a $\mathcal{V}$-module, it is called *simple* (see Definition 6.2.3). All VOAs of interest in this book are simple. An example of a nonsimple vertex algebra is the affine algebra vertex algebra at critical level $k = -h^\vee$.

The notion of tensor product – called *fusion* $M \boxtimes N$ – for VOA modules is unexpectedly subtle. For example, the infinite-dimensional adjoint module $\mathcal{V}$ should have trivial fusions, just like the one-dimensional Lie algebra module $\mathbb{C}$ has trivial tensor products. See, for example, [**298**], [**222**], [**382**] for various approaches. Fusion products in a weakly rational VOA can be decomposed into irreducible modules as usual:

$$M \boxtimes N \cong \oplus_{P \in \Phi(\mathcal{V})} \mathcal{N}_{MN}^{P}\, P , \tag{5.3.3}$$

where the multiplicities $\mathcal{N}_{MN}^{P}$ are called *fusion coefficients*. These numbers are most easily defined (via Schur's Lemma) as the dimension of the space of intertwiners [**199**] (Definition 6.1.9). For semi-simple Lie algebras, the tensor product of modules defines a symmetric monoidal category (Section 1.6.2); for nice VOAs, the fusion of modules defines a braided monoidal category and the structure constants $\mathcal{N}_{MN}^{P}$ a fusion ring (Section 6.2.2).

Definition 5.3.2 [574] *A VOA $\mathcal{V}$ is called* weakly rational *if every $\mathcal{V}$-module is completely reducible, $\mathcal{V}$ has only a finite number of irreducible modules, and every irreducible weak $\mathcal{V}$-module is a $\mathcal{V}$-module.*

Let $\Phi(\mathcal{V})$ denote the set of irreducible $\mathcal{V}$-modules. Most of our VOAs will be weakly rational. The term 'weakly rational' is not standard; *rational* is sometimes used. However, a rational VOA should enjoy all properties of the chiral algebra of a RCFT, which is why we reserve the term 'rational' for the stronger notion presented in Definition 6.2.3.

Lemma 5.3.3 [**574**] *Let $\mathcal{V}$ be a weakly rational VOA, and let M be any irreducible $\mathcal{V}$-module. Then there is a number $h \in \mathbb{Q}$ such that the homogeneous subspace M_h is nonzero, and such that if $M_\alpha \neq 0$ for some $\alpha \in \mathbb{C}$, then $\alpha - h \in \mathbb{N}$.*

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

The proof isn't difficult – see page 244 of [**574**] for a more general argument. We call $h = h(M)$ the *conformal weight* of M, and the space $M_h = M_{[0]}$ the *lowest-weight space* of M. For example, the conformal weight $h(\mathcal{V})$ of the adjoint module is 0. The lowest-weight space M_h generates the whole module, in the sense (5.1.9a) that M is spanned by vectors of the form $(u_1)_{(n_1)} \cdots (u_k)_{(n_k)} y$ for $u_i \in \mathcal{V}$ and $y \in M_h$. The lemma implies that for such a module M, we have $\chi_M(\tau + 1) = e^{2\pi \mathrm{i}\, h(M)} \chi_M(\tau)$ as formal power series.

In both finite group theory and Lie theory, given any module M, a module structure can also be found on the vector space dual M^* of M in a straightforward way. This module is called the *dual* or *contragredient* of M. Something similar happens for VOAs. However, the naive dual of an infinite-dimensional space tends to be too large (recall that in infinite dimensions, the double-dual $(V^*)^*$ properly contains V), so here we take instead the *restricted dual* $M^\star$ of M, defined by

$$M^\star = \oplus_\alpha (M_\alpha)^*. \tag{5.3.4a}$$

The explicit $\mathcal{V}$-module structure on $M^\star$ (see section 5.2 of [**199**]) is quite complicated and closely related to the definition of invariant bilinear form in (5.1.11a). Note that

$$\chi_{M^\star}(\tau) = \chi_M(\tau) \tag{5.3.4b}$$

even though $M^\star$ and M are usually non-isomorphic as $\mathcal{V}$-modules. Thus our graded dimensions (5.3.2) won't always distinguish modules, something that was independently observed in the context of Monstrous Moonshine, as we'll see. We return to this bothersome but not unexpected fact in Section 5.3.3. The more obscure term 'contragredient' is usually used for $M^\star$, as 'dual' has too many unfortunately independent meanings. The notion of contragredient module plays a large role in RCFT: roughly, $M^\star$ is the anti-particle of M, and they are related by charge-conjugation C.

All VOAs $\mathcal{V}$ of interest to us have an anti-linear involution $u \mapsto u^*$ such that the invariant bilinear form $(u|v)$ of (5.1.11a) satisfies

$$(u|v^*) = \overline{(v|u^*)}, \qquad \forall u, v \in \mathcal{V}. \tag{5.3.5a}$$

The notion of *unitary module* M is important in physics: it is a $\mathcal{V}$-module in which the bilinear form on M satisfies

$$(ux|y)_M = (x|u^* y)_M, \qquad \forall u \in \mathcal{V}, x, y \in M. \tag{5.3.5b}$$

Consider first the lattice VOA $\mathcal{V}(L)$ constructed in Section 5.2.2, where L is a positive-definite even lattice (recall the definitions in Section 1.2.1). It is weakly rational, and its irreducible modules are parametrised naturally by the cosets L^*/L, where $L^* \supseteq L$ is the dual lattice to L [**144**]. The explicit construction of these modules $M[t]$, for $[t] \in L^*/L$, is very similar to that of the VOA $\mathcal{V}_L$ itself – see section 6.5 of [**376**]. Thus the number $\|\Phi(\mathcal{V}(L))\|$ of its irreducible modules is given by the determinant $|L|$ of the lattice. The adjoint module is $M[0]$. The module $M[t]$ has contragredient $M[-t]$ and graded dimension

$$\chi_{M[t]}(\tau) = \frac{\Theta_{t+L}(\tau)}{\eta(\tau)^n}, \tag{5.3.6}$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

where n is the dimension of L, η is the Dedekind eta function (2.2.6b) and Θ_{t+L} is the theta series of (2.2.11a). The fusion product here is $M[t] \boxtimes M[t'] = M[t+t']$.

The Heisenberg VOAs are not weakly rational. For example, $\mathcal{V}(\mathbb{C})$ has a distinct irreducible module $M(\lambda)$ (namely the Verma module $V(\lambda)$ of (3.2.12b)) for every $\lambda \in \mathbb{C}$. The adjoint module is $M(0)$, and the contragredient of $M(\lambda)$ is $M(-\lambda)$. Only the modules with $\lambda \in \mathbb{R}$ are unitary. The graded dimension of $M(\lambda)$ is given in (3.2.12c).

However, if $\overline{\mathfrak{g}}$ is a simple Lie algebra and $\mathfrak{g} = \overline{\mathfrak{g}}^{(1)}$ is the associated nontwisted affine algebra, then the VOA $\mathcal{V}(\mathfrak{g}, k)$ will be weakly rational iff the level k lies in $\mathbb{N}$. Just as the VOA $\mathcal{V}(\mathfrak{g}, k)$ is the $\mathfrak{g}$-module $L(k\omega_0)$ with additional structure, the irreducible $\mathcal{V}(\mathfrak{g}, k)$-modules can be identified with the $\mathfrak{g}$-modules $L(\lambda)$, for level-k integrable highest weights $\lambda \in P_+^k(\mathfrak{g})$ [**202**]. In particular, the VOA graded dimension will equal the corresponding *specialised* affine algebra characters $\chi_\lambda(2\pi \mathrm{i} \tau \ell_0) = \chi_\lambda(\tau, 0, 0)$ of (3.2.11c). The usual tensor product $L(\lambda) \otimes L(\mu)$ of affine algebra modules is less interesting than the fusion product $L(\lambda) \boxtimes L(\mu)$ – in the former, levels add and the tensor product coefficients $T_{\lambda\mu}^{\nu}$ can be infinite, while the latter is studied in Section 6.2.1.

A weakly rational VOA is called *holomorphic* if it has a unique irreducible module. As usual this terminology comes from RCFT: a holomorphic VOA can be the left-moving chiral algebra of a CFT with trivial right-moving chiral algebra, so the physical correlation functions (4.3.1a) of such a CFT would be holomorphic (at least locally, when all insertion points z_i are distinct). Thus the lattice VOA $\mathcal{V}(L)$ is holomorphic iff the lattice L is self-dual. The most famous example of a holomorphic VOA though is the Moonshine module $V^\natural$ [**145**]. In fact, its holomorphicity is one of the keys to Monstrous Moonshine (see Question 5.3.4).

5.3.2 Zhu's algebra

In many ways a VOA resembles a Lie algebra, and this analogy has often been exploited to flesh out the theory of VOAs. However, the representation theory of the *weakly rational* VOAs resembles that of a finite group.

Consider for concreteness the symmetric group $G = \mathcal{S}_3$. Its representation theory is captured by its group algebra $\mathbb{C}G$ (Section 1.1.3), that is the formal span of the elements $\sigma \in G = \{(1), (12), (23), (13), (123), (132)\}$, where G acts by left multiplication. The associative algebra $\mathbb{C}G$ is semi-simple, and so is a direct sum of matrix algebras: here,

$$\mathbb{C}G \cong M_{1\times 1} \oplus M_{1\times 1} \oplus M_{2\times 2}, \tag{5.3.7a}$$

where the first summand $M_{1\times 1}$ contains one copy of the trivial one-dimensional irreducible representation $\rho_1(\sigma) = 1$, the second summand $M_{1\times 1}$ contains one copy of the 'sign' one-dimensional irreducible representation $\rho_s(\sigma) = (-1)^\sigma$, and the four-dimensional algebra $M_{2\times 2}$ contains a continuum of copies of the two-dimensional irreducible representation ρ_2. More precisely, the three subspaces of the group algebra $\mathbb{C}G$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

specified by (5.3.7a) are

$$V_1 = \mathbb{C}\{(1)+(12)+(23)+(13)+(123)+(132)\} \cong \rho_1, \tag{5.3.7b}$$

$$V_s = \mathbb{C}\{(1)-(12)-(23)-(13)+(123)+(132)\} \cong \rho_s, \tag{5.3.7c}$$

$$V_2 = \mathbb{C}\{(1)-(123), (1)-(132), (12)-(23), (12)-(13)\} \cong \rho_2 \oplus \rho_2. \tag{5.3.7d}$$

Incidentally, the different copies of the irreducible module ρ_2 in the subspace V_2 are parametrised by the projective line $\mathbb{P}^1(\mathbb{R}) \cong S^1$: choosing a nonzero point x in

$$\mathbb{C}\{(1)-(12)+(23)-(132), (23)-(13)+(123)-(132)\}, \tag{5.3.7e}$$

and hitting with arbitrary $\sigma \in G$, spans a copy $V_2(x)$ of the two-dimensional module ρ_2, and $V_2(x) \cap V_2(x') = \{0\}$ unless x and x' are complex multiples of each other, in which case $V_2(x)$ and $V_2(x')$ are equal as sets. On the other hand, choosing a generic element of V_2 (respectively $\mathbb{C}G$) will span all of V_2 (respectively $\mathbb{C}G$).

The representation theory of a finite group G is equivalent to that of the associative algebra $\mathbb{C}G$. Likewise, for semi-simple Lie algebras $\mathfrak{g}$ there is also an associative algebra, generated by $\mathfrak{g}$, which classifies all irreducible $\mathfrak{g}$-modules: the universal enveloping algebra $U(\mathfrak{g})$ (Section 1.5.3). However, it is infinite-dimensional, reflecting the fact that $\mathfrak{g}$ has infinitely many inequivalent irreducible modules.

Remarkably, weakly rational VOAs $\mathcal{V}$ have (like finite G), a finite-dimensional associative semi-simple algebra, denoted $A(\mathcal{V})$, which classifies the finitely many irreducible $\mathcal{V}$-modules. As we know, the full module M can be generated from its lowest-weight space M_h, by repeatedly acting by modes of $\mathcal{V}$, and so it suffices to study M_h. Now, the zero-modes $o(u)$, defined at the beginning of Section 5.2.1, act on each homogeneous space M_α; Zhu's algebra $A(\mathcal{V})$ is the algebra of zero-modes, as seen by the lowest-weight spaces M_h. A more formal construction, which will begin next paragraph, is due to Zhu [**574**], although it was anticipated in physics [**429**], [**87**]. Similar to the above, each irreducible $\mathcal{V}$-module M corresponds to a linear functional f_M on $\mathcal{V}$ (Section 4.4.4); a certain large subspace $O(\mathcal{V})$ of $\mathcal{V}$ lies in the kernel of all functionals $f_M \circ o(v)\ \forall v \in \mathcal{V}$, so each of these defines a well-defined functional on the quotient $A(\mathcal{V}) := \mathcal{V}/O(\mathcal{V})$. The quotient $A(\mathcal{V})$ has a product $u * v$ making it into an associative algebra; the space of functionals $f_M \circ o(v)$ carries a module action of $A(\mathcal{V})$, and as such can be identified with the dual M_h^* of the lowest-weight space of M. Conversely, any (irreducible) right-module for $A(\mathcal{V})$ is the lowest-weight space of an (irreducible) $\mathcal{V}$-module M. This physically motivated treatment of Zhu's algebra is fleshed out in [**227**].

Zhu's treatment is similar. For $u, v \in \mathcal{V}$, where $u \in \mathcal{V}_k$, define a product

$$u * v = \mathrm{Res}_z \left(Y(u,z)\, v \frac{(z+1)^k}{z} \right), \tag{5.3.8a}$$

or equivalently, in terms of the modes,

$$(u * v)_{(n)} = \sum_{m \geq k} u_{(-1-m)} \circ v_{(m+n)} + \sum_{m \leq k-1} v_{(m+n)} \circ u_{(-1-m)}. \tag{5.3.8b}$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

Extend $*$ linearly to all $u \in \mathcal{V}$. Let $O(\mathcal{V})$ be the subspace of $\mathcal{V}$ spanned by elements

$$(L_{-1}u + L_0 u) * v, \qquad \forall u, v \in \mathcal{V}. \tag{5.3.8c}$$

By *Zhu's algebra* $A(\mathcal{V})$ we mean the quotient $\mathcal{V}/O(\mathcal{V})$.

The point of these definitions is that, on the lowest-weight space M_h of any irreducible $\mathcal{V}$-module M, a straightforward calculation (see page 250 of [**574**]) verifies that

$$o(u * v) = o(u) \circ o(v). \tag{5.3.9a}$$

Using (5.1.8b), (5.1.7d) and VA2, we see that

$$o(L_{-1}u + L_0 u) = 0 \tag{5.3.9b}$$

identically on $\mathcal{V}$. Together, (5.3.9) tell us $o(u) = 0$ on each lowest-weight space M_h, for any $u \in O(\mathcal{V})$. Thus for any class $[u] \in A(\mathcal{V})$, the zero-mode $o(u)$ is a well-defined operator on each M_h.

Theorem 5.3.4 [**574**] *Let $\mathcal{V}$ be a weakly rational VOA (recall Definition 5.3.2) and let $A(\mathcal{V}) = \mathcal{V}/O(\mathcal{V})$ be Zhu's algebra. Then $A(\mathcal{V})$ is a finite-dimensional, associative and semi-simple algebra, isomorphic as an algebra to the matrix algebra*

$$A(\mathcal{V}) \cong \oplus_{M \in \Phi(\mathcal{V})} M_{n(M) \times n(M)},$$

where $\Phi(\mathcal{V})$ is the set of all irreducible $\mathcal{V}$-modules, and $n(M)$ is the dimension of the lowest-weight space M_h.

In other words, there is a one-to-one correspondence between the irreducible modules of $A(\mathcal{V})$ and $\mathcal{V}$; the irreducible $A(\mathcal{V})$-modules can in fact be naturally identified with the lowest-weight spaces M_h of the irreducible $\mathcal{V}$-modules. It is almost identical to what happens with the group algebra of a finite group. Note that the dimension $n(M)$ is the coefficient of the first nontrivial term $n(M)\, q^{h-c/24}$ of the graded dimension χ_M. The hard part of the proof of Theorem 5.3.4 is establishing that an irreducible $A(\mathcal{V})$-module lifts to an irreducible $\mathcal{V}$-module (the basic idea is sketched above). Incidentally, there are non-weakly rational VOAs (coming from 'logarithmic' CFTs) with Zhu's algebra $A(\mathcal{V})$ finite-dimensional but not semi-simple.

For example, Zhu's algebra $A(V^\natural)$ for the Moonshine module $V^\natural$ is one-dimensional, while the integrable affine VOA $\mathcal{V}(\mathfrak{g}, k)$ at level $k \in \mathbb{N}$ has Zhu's algebra

$$A(\mathcal{V}(\mathfrak{g}, k)) \cong \oplus_{\lambda \in P_+^k(\mathfrak{g})} M_{\dim L(\overline{\lambda}) \times \dim L(\overline{\lambda})},$$

where $L(\overline{\lambda})$ is a highest-weight $\overline{\mathfrak{g}}$-module (to get $\overline{\lambda}$, drop λ_0 from λ). In general though, it is hard to compute $A(\mathcal{V})$ (unless the $\mathcal{V}$-modules are already known!) because we lose the grading – expressions like $L_{-1}u + L_0 u$ are not homogeneous.

The definition (5.3.8a) of the product '$*$' in Zhu's algebra can be modified to give the more familiar 'normal-ordered product' (recall (5.1.6))

$$u \cdot v = \mathrm{Res}_z(Y(u, z)\, v z^{k-1}) = u_{(-1)}v \tag{5.3.10a}$$

for $u \in \mathcal{V}_k$, or equivalently in terms of modes

$$(u \cdot v)_{(n)} = \sum_{m \geq 0} u_{(-1-m)} \circ v_{(m+n)} + \sum_{m \leq -1} v_{(m+n)} \circ u_{(-1-m)}. \tag{5.3.10b}$$

Let $O_2(\mathcal{V})$ be the span of all elements of the form $u_{(-2)}v$, and $A_2(\mathcal{V})$ the quotient $\mathcal{V}/O_2(\mathcal{V})$. Then $A_2(\mathcal{V})$ is a graded commutative associative algebra with product '$\cdot$'. It also has a Lie algebra structure, with bracket given by $[uv] = u_{(0)}v$; together, the Lie and associative products define a commutative Poisson algebra. Its main role in VOA theory is in a finiteness condition:

Definition 5.3.5 [574] *A VOA $\mathcal{V}$ is said to be C_2*-cofinite *if the $A_2(\mathcal{V}) = \mathcal{V}/O_2(\mathcal{V})$ is finite-dimensional.*

Most of the important weakly rational VOAs (e.g. the Moonshine module, the lattice VOAs, the affine algebra VOAs at positive integer level) satisfy this condition. The term 'C_2-cofinite' comes from Zhu's name for what we call $O_2(\mathcal{V})$. It has several consequences. Most importantly, the graded dimensions $\chi_M(\tau)$ of a C_2-cofinite VOA converge to functions holomorphic in the upper half-plane $\mathbb{H}$ (theorem 4.4.2 of [**574**]). A C_2-cofinite VOA will have well-defined finite fusion coefficients (5.3.3) (see theorem 13 in [**229**]).

It is conjectured that a VOA is weakly rational if and only if it is C_2-cofinite, but although this would significantly simplify the definition of weakly rational, it seems difficult to prove. Weakly rational VOAs satisfy $\dim A_2(\mathcal{V}) \geq \dim A(\mathcal{V})$ (generalised in lemma 3 of [**229**]), but inequality can occur – for example, the integrable affine algebra VOA $\mathcal{V}(E_8{}^{(1)}, 1)$ has a one-dimensional Zhu's algebra but $A_2(\mathcal{V})$ is at least 249-dimensional [**224**].

A C_2-cofinite VOA is finitely generated in the sense that there will be finitely many vectors $u^1, \ldots, u^n \in \mathcal{V}$ (namely, choose u^i to be the lifts to $\mathcal{V}$ of a basis of $A_2(\mathcal{V})$) such that $\mathcal{V}$ is spanned by all vectors of the form

$$u^{i_1}_{(-m_1)} \cdots u^{i_k}_{(-m_k)} \mathbf{1}, \tag{5.3.11a}$$

where $m_1 > \cdots > m_k > 0$ [**229**]. Something similar (but weaker) holds for $\mathcal{V}$-modules. Using this we quickly obtain a growth estimate: given any C_2-cofinite VOA $\mathcal{V}$, there is a constant $C > 0$ such that, for any irreducible $\mathcal{V}$-module M, the dimension of the homogeneous space M_α is bounded above by

$$\dim M_\alpha < C_M e^{C\sqrt{\alpha - h}}, \tag{5.3.11b}$$

for some constant C_M, where as always $h = h(M)$ is the conformal weight of M. The constant C depends only on $\dim A_2(\mathcal{V})$, while C_M is essentially $\dim M_h$, adjusted slightly to ensure (5.3.11b) also holds for small α.

Various interesting generalisations of Zhu's algebras have appeared in the literature [**149**], [**150**], [**229**], [**410**]. From our point of view, these algebras play a crucial technical role in the statement and proof of the modularity of VOA characters.

5.3.3 The characters of VOAs

The next four subsections mark a climax for the book, as we discuss the modularity of the graded dimensions (5.1.10b), (5.3.2). We also explain why this was anticipated by physicists. But first let's reflect on the notion of character.

Calling the quantities $\chi_{\mathcal{V}}(\tau)$ and $\chi_M(\tau)$ 'characters', as is common in the literature, is a misnomer – they are merely graded dimensions. Defining characters for an algebraic object is as much art as science. The beautiful success of the character theory of semi-simple and Borcherds–Kac–Moody Lie algebras hides the nontrivial intuition that went into the original definitions. Presumably the starting point was that the characters of finite groups are given by the trace. Also, exponentiation associates a Lie group with a Lie algebra. Putting this together leads to the character of (1.5.9a). The characters of (Borcherds–)Kac–Moody algebras then follow by analogy. Unfortunately, the situation for VOAs isn't nearly as clear.

The main properties we may hope a character χ_M to obey are: it specialises to dimension (or graded dimension); it distinguishes inequivalent modules; and it respects direct sum and tensor product (fusion for us), in the sense that $\chi_{M\oplus N} = \chi_M + \chi_N$ and $\chi_{M \boxtimes N} = \chi_M \chi_N$. We would also expect the VOA characters in the special case of the integrable affine VOA $\mathcal{V}(\mathfrak{g}, k)$ to equal the corresponding affine algebra characters χ_λ in (3.2.9a) (recall that the $\mathcal{V}(\mathfrak{g}, k)$-modules can be identified with the integrable $\mathfrak{g}$-modules).

This wish-list is hopelessly optimistic for even the nicest VOAs. The graded dimensions $\chi_{M(\lambda)}(\tau)$ for the integrable affine VOA $\mathcal{V}(\mathfrak{g}, k)$ will not respect the fusion product:

$$\chi_{M(\lambda)\boxtimes M(\mu)}(\tau) \neq \chi_{L(\lambda)\otimes L(\mu)}(\tau) = \chi_{L(\lambda)}(\tau)\,\chi_{L(\mu)}(\tau) = \chi_{M(\lambda)}(\tau)\,\chi_{M(\mu)}(\tau),$$

where $L(\lambda)\otimes L(\mu)$ denotes the tensor product of $\mathfrak{g}$-modules. On the other hand, fusion respects the asymptotic dimensions: for all sufficiently nice VOAs $\mathcal{V}$, the limit

$$\mathcal{D}(M) = \lim_{\tau\to 0}\frac{\chi_M(\tau)}{\chi_{\mathcal{V}}(\tau)}, \tag{5.3.12}$$

called the *quantum dimension* of $M \in \Phi(\mathcal{V})$, satisfies $\mathcal{D}(M \boxtimes N) = \mathcal{D}(M)\,\mathcal{D}(N)$. 'Sufficiently nice' here means any C_2-cofinite weakly rational VOA $\mathcal{V}$ obeying the additional very common property that of all irreducible $\mathcal{V}$-modules $M \in \Phi(\mathcal{V})$, a unique one realises the smallest conformal weight $\min_{M\in\Phi(\mathcal{V})} h(M)$ (in the most familiar examples the unique minimal conformal weight belongs to the adjoint module $M = \mathcal{V}$).

Recall from (5.3.4b) that the graded dimensions $\chi_M(\tau)$ of inequivalent $\mathcal{V}$-modules can be equal. A further example occurs whenever an even positive-definite lattice L has an automorphism α; then any pair $M[t]$, $M[\alpha t]$ of $\mathcal{V}(L)$-modules will have identical graded dimension. However, such equalities need not always have an easy algebraic explanation: for example, in Monstrous Moonshine two McKay–Thompson series (namely, $T_{27A}(\tau) = T_{27B}(\tau)$, corresponding to unrelated elements of order 27) accidentally coincide for no obvious reason. None of this is surprising, since dimensions certainly don't uniquely specify Lie algebra or finite group modules.

We certainly would like VOA characters to distinguish inequivalent $\mathcal{V}$-modules, and in fact be linearly independent. How to do this is clear from the study of lattice theta

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

functions or affine algebra characters: in order to retain more information of the homogeneous spaces M_α than merely their dimensions, we must include more variables in χ_M.

Definition 5.3.6 *The* character *of a $\mathcal{V}$-module M is the* one-point function $\chi_M(\tau, v)$

$$\chi_M(\tau, v) := \mathrm{tr}_M\, o(v)\, q^{L_0 - c/24} = q^{-c/24} \sum_{n=0}^{\infty} \mathrm{tr}_{M_{h+n}} o(v) q^{h+n}. \tag{5.3.13}$$

$h = h(M)$ is the conformal weight of M, and $o(v)$ is the zero-mode (Section 5.2.1) of $v \in \mathcal{V}$, which is an endomorphism on each homogeneous space M_{h+n} (so its trace can be computed by choosing bases and writing $o(v)$ as a matrix for each n). This function χ_M arises naturally in CFT, as the one-point chiral block (Section 4.3.2) on the torus. We explain shortly why it is associated with a torus – this is the source of its modularity.

Note that $\chi_M(\tau, \mathbf{1})$ equals the graded dimension $\chi_M(\tau)$. By definition, the dependence of $\chi_M(\tau, v)$ on $v \in \mathcal{V}$ is linear. Provided $\mathcal{V}$ is C_2-cofinite, theorem 4.4.1 of [**574**] tells us that, for each $v \in \mathcal{V}$, $\chi_M(\tau, v)$ is holomorphic for $\tau \in \mathbb{H}$. This is proved by finding and studying a differential equation satisfied by $\chi_M(\tau, v)$. Their modularity is established in Section 5.3.5.

When $\mathcal{V}$ is weakly rational and C_2-cofinite, the one-point functions are linearly independent and thus distinguish inequivalent $\mathcal{V}$-modules. In fact, we see from the proof of theorem 5.3.1 in [**574**] that if $\mathcal{V}_A$ is any lift from Zhu's algebra $A(\mathcal{V})$ to $\mathcal{V}$, then the one-point functions $\chi_M(\tau, v)$ will remain linearly independent even if v is restricted to the finite-dimensional subspace $\mathcal{V}_A$. For example, the graded dimensions $\chi_M(\tau)$ and $\chi_{M^\star}(\tau)$ are equal, but for $v \in \mathcal{V}_n$ the one-point functions obey

$$\chi_{M^\star}(\tau, v) = (-1)^n \chi_M(\tau, v).$$

Although one-point functions (5.3.13) don't directly respect the fusion product (but recall (5.3.12)), they deserve the title 'character' as they are the simplest linearly independent extension of graded dimension. However, since they depend *linearly* and not *exponentially* on v, how can we reconcile them with the Jacobi theta functions (2.3.7) and the affine algebra characters (3.2.9a)? Mindlessly defining a function

$$\exp[2\pi \mathrm{i} w]\, \mathrm{tr}_M \exp[2\pi \mathrm{i}\, o(v)]\, q^{L_0 - c/24} \tag{5.3.14}$$

for $v \in \mathcal{V}$ and $w \in \mathbb{C}$ will lose modularity.

The key is to realise that, although the exponential $q = e^{2\pi \mathrm{i}\tau}$ is *topological* in origin, the exponential $e^{2\pi \mathrm{i} z}$ in (2.3.7) and (3.2.9a) is *Lie theoretic* in origin. In particular:

Definition 5.3.7 *Let $\mathcal{V}$ be a weakly rational C_2-cofinite VOA. For any $\mathcal{V}$-module $M \in \Phi(\mathcal{V})$, define the* Jacobi character *to be the quantity $\chi^J_M(\tau, v, w)$ given by (5.3.14), except we restrict v to the Lie algebra $\mathcal{V}_1$.*

Of course $v = 0$ and $w = 0$ recovers the graded dimensions. As we know, $e^{o(v)}$ is an automorphism of M for $v \in \mathcal{V}_1$, and as we recall from the McKay–Thompson series the graded trace of automorphisms is worthy of study. Question 5.3.5 asks the reader to

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

verify that χ^J_M recovers affine algebra characters. Of course the complex variable 'w' is merely included for book-keeping. We return to Jacobi characters in Theorem 5.3.9.

If we hadn't restricted v in Definition 5.3.7 to $\mathcal{V}_1$, then linear independence would have been assured by that of the one-point functions $\chi_M(\tau, v)$ (why?). In the familiar examples (e.g. lattice or affine algebra VOAs) we still have linear independence of the Jacobi characters, but it won't hold for all other VOAs.

5.3.4 Braided #5: the physics of modularity

Let's turn next to one of the central questions in the book: why should the VOA characters χ_M have anything to do with modularity? In short, it is because they are toroidal chiral blocks of RCFT, and the mapping class group $\Gamma_{1,1}$ (which must act on those chiral blocks) is $\mathrm{SL}_2(\mathbb{Z})$. While filling in this explanation we'll finally explain the shift '$c/24$' appearing in the definition of the affine algebra characters and more generally the VOA characters χ_M.

Lurking in the background of the following argument is the closed string, with period-1 arc-parameter σ and time-parameter t (recall Section 4.3.2). For the left-moving (holomorphic) sector it is convenient to introduce complex parameters $\sigma - \mathrm{i}t$ and $e^{2\pi\mathrm{i}(\sigma-\mathrm{i}t)}$, which we now call z and w, respectively.

From the perspective of VOAs and CFT, the easiest way to realise the torus $\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau)$ for $\tau \in \mathbb{H}$, starting with the space $\mathbb{C}$, is by first considering the map $z \mapsto e^{2\pi\mathrm{i}z}$ (the '$2\pi\mathrm{i}$' is merely a convenient normalisation). This is a holomorphic map sending neighbourhoods of 0 to neighbourhoods of 1. It changes the global topology, however, sending the plane $\mathbb{C}$ to the annulus $\mathbb{C}\backslash\{0\}$. Now it is simple to obtain our torus: we simply identify z and qz, where as always $q = e^{2\pi\mathrm{i}\tau}$. This is equivalent to taking the finite annulus $\{z \in \mathbb{C} \mid |q| < |z| < 1\}$ and sewing together its two boundary circles by identifying z on the outer circle with qz on the inner. The resulting torus is conformally equivalent to $\mathbb{C}/(\mathbb{Z}+\mathbb{Z}\tau)$ (why?). The point is that the chiral blocks on the torus can be obtained from those of the plane, through this construction of the torus from $\mathbb{C}$. Let us now give the details.

Let $\mathcal{V}$ be any VOA. For any coordinate transformation $z \mapsto w = f(z)$ sending 0 to 0, and holomorphic in a neighbourhood of 0, the Virasoro algebra lets us calculate its effect on any vertex operator: we can write

$$Y(v, z) \mapsto T_f \circ Y(v, z) \circ T_f^{-1} \tag{5.3.15a}$$

for some invertible linear map $T_f : \mathcal{V} \to \mathcal{V}$ (see [**223**], [**295**] for the explicit and general calculation). More precisely, there are $a_i \in \mathbb{C}$ such that (see proposition 2.1.1 in [**295**])

$$f(z) = \exp\left[\sum_{n=0}^{\infty} a_n z^{n+1} \frac{\mathrm{d}}{\mathrm{d}z}\right] z \tag{5.3.15b}$$

as formal power series, where 'exp' is defined by its Taylor series. Then we obtain

$$T_f v = \exp\left[\sum_{n=0}^{\infty} a_n L_n\right] v \tag{5.3.15c}$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

(regularity VA5 implies this map $T_f : \mathcal{V} \to \mathcal{V}$ is always defined). When v is a conformal primary of conformal weight k (recall (5.2.3)), the transformation is particularly nice:

$$T_f \circ Y(v, z) \circ T_f^{-1} = Y(v, w)\,(f'(z))^k. \tag{5.3.15d}$$

The other important special case is the stress–energy tensor $T(z) = Y(\omega, z)$:

$$T_f \circ Y(\omega, z) \circ T_f^{-1} = Y(\omega, w)\,(f'(z))^2 + \frac{c}{12}\,\{f(z), z\}, \tag{5.3.15e}$$

where $\{f, z\}$ is the *Schwarzian derivative*

$$\{f, z\} := \frac{f'''(z)}{f'(z)} - \frac{3}{2}\left(\frac{f''(z)}{f'(z)}\right)^2. \tag{5.3.15f}$$

The factor '$c/12$' in (5.3.15e) is the same as in (3.1.5a). The Schwarzian derivative vanishes if and only if f is a Möbius transformation (i.e. if and only if f conformally maps the Riemann sphere to itself), and so is a measure of how f changes the global topology.

Provided $f(z)$ is holomorphic near 0 and obeys $f(0) = 0$, a second VOA structure can be defined on the vector space $\mathcal{V}$ as follows. The vertex operators are $Y_f(v, z) = Y(T_f v, f(z))$, the vacuum is $\mathbf{1}_f = T_f(\mathbf{1}) = \mathbf{1}$, and conformal vector is $\omega_f = T_f(\omega)$. Let $\mathcal{V}_f$ denote this second VOA. Then $\mathcal{V}$ and $\mathcal{V}_f$ are isomorphic. (See [**293**] for a generalisation dropping the $f(0) = 0$ condition.)

We are interested in the transformation $w = f(z) = e^{2\pi\mathrm{i}z} - 1$. Then everything simplifies and we get

$$\omega_f = 4\pi^2\,(\omega - c/24), \tag{5.3.16a}$$

$$Y_f(v, z) = Y(v, w)\,e^{2\pi\mathrm{i}zk}, \qquad \forall v \in \mathcal{V}_k. \tag{5.3.16b}$$

Although $\mathcal{V}_f$ is a VOA isomorphic to $\mathcal{V}$ sharing the same underlying space, modes and conformal weights are quite different. We will use square brackets to indicate the modes of $\mathcal{V}_f$, and denote its Virasoro generators by $L[n] = (\omega_f)_{[n+1]}$. We find for instance that

$$L[-1] = 2\pi\mathrm{i}\,(L_{-1} + L_0), \tag{5.3.16c}$$

$$L[0] = L_0 + \sum_{i \geq 1} \frac{(-1)^{n-1}}{n(n+1)} L_n. \tag{5.3.16d}$$

Although by the isomorphism of $\mathcal{V}$ and $\mathcal{V}_f$ the homogeneous spaces $\mathcal{V}_n$ and $\mathcal{V}_{[n]}$ must be equal dimension, and in fact carry isomorphic representations of Aut $\mathcal{V}$, we only have $\mathcal{V}_n = \mathcal{V}_{[n]}$ for $n = 0$ or if dim $\mathcal{V}_n = 0$. On the other hand, if $v \in \mathcal{V}$ is a conformal primary of conformal weight k with respect to the operators L_n, then it will be one with respect to the operators $L[n]$ as well (see Question 5.3.2).

For a technical reason, we are also interested in the simple relation between the usual power series modes $L[n]$ of $\mathcal{V}_f$, and the Fourier modes L'_n of $\mathcal{V}$, defined by

$$T(z) = Y(\omega, z) = -\sum_{m=-\infty}^{\infty} L'_m e^{2\pi\mathrm{i}mw}.$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

We get (recall Question 3.1.8)

$$L'_n = L[n] - \delta_{n,0}\frac{c}{24}.$$

The occurrences of '$-c/24$' in, for example, the characters of affine algebras and VOAs can be traced back to its occurrence in (5.3.16a). Mathematically, it is a symptom of the change of global topology, from the plane to an annulus. Physically this is interpreted as the Casimir energy of the cylinder [**3**]; see also the discussion in section 5.4 of [**131**].

Our map f mapped the plane to the annulus $\mathbb{C}\backslash\{-1\}$. To get the torus, we need to identify z on the outer circle $e^{\mathrm{i}\theta} - 1$, with the point $q\,(z+1) - 1 = qe^{i\theta} - 1$ on the inner circle. By the axioms of CFT (e.g. Section 4.4.1), this identification ('sewing') corresponds to taking a trace. For simplicity consider first the vacuum-to-vacuum amplitude ('partition function') on this torus, and write $\tau = s + \mathrm{i}t$. The desired trace will be over the full space of states $\mathcal{H}$, and will be of the 'propagator' for the cylinder, which takes the string and evolves it $2\pi t$ ahead in time and twists it $2\pi s$ arcwise. The infinitesimal generator of twists is the corresponding momentum operator, call it P, and the infinitesimal generator for time evolution is the Hamiltonian H, both in the z-coordinate frame. Thus the partition function will be

$$\mathcal{Z}(\tau) = \mathrm{tr}_{\mathcal{H}} \exp[2\pi \mathrm{i} s P - 2\pi t H].$$

To find, for example, the Hamiltonian, note that changing time by δt changes the w-coordinate by the factor $e^{-2\pi\delta t}$, so the Hamiltonian generates dilations in w (recall the calculation in Section 4.3.2); similarly, the momentum operator generates rotations in w. We obtain

$$\begin{aligned} P &= L'_0 - \overline{L}_0{}' = L[0] - \overline{L}[0] - \frac{c}{24} + \frac{\overline{c}}{24}, \\ H &= L'_0 + \overline{L}'_0 = L[0] + \overline{L}[0] - \frac{c}{24} - \frac{\overline{c}}{24}, \end{aligned}$$

where we use bars to denote the anti-holomorphic quantities. Thus we obtain the familiar expression for the partition function:

$$\mathcal{Z}(\tau) = \mathrm{tr}_{\mathcal{H}} q^{L[0]-c/24}\, \overline{q}^{\overline{L}[0]-\overline{c}/24} = \mathrm{tr}_{\mathcal{H}} q^{L_0 - c/24}\, \overline{q}^{\overline{L}_0 - \overline{c}/24},$$

where the final equality follows from the isomorphism of VOAs $\mathcal{V}$ and $\mathcal{V}_f$. CFT or string theory requires that $\mathcal{Z}(\tau)$ be a function only of the conformal equivalence class of the torus $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$ – in other words, $\mathcal{Z}(\tau)$ must be invariant under the action of the modular group $\mathrm{SL}_2(\mathbb{Z})$.

We are more interested here in the associated chiral quantities, since a VOA is the chiral algebra of the theory. From the previous paragraph, together with the decomposition (4.3.6) of $\mathcal{H}$ into modules of $\mathcal{V} \otimes \mathcal{V}'$, we can now read off the decomposition of $\mathcal{Z}(\tau)$ into chiral blocks (see (4.3.8b)) in a RCFT. Hence the chiral blocks for the torus are

$$\mathrm{tr}_M q^{L_0 - c/24}$$

– that is, they are simply the graded dimensions of the irreducible $\mathcal{V}$-modules, including the strange shift by $c/24$. RCFT requires that this space must carry a projective representation of the mapping class group of the torus $\mathrm{SL}_2(\mathbb{Z})$.

By the same reasoning, we can calculate the n-point chiral blocks on the torus. For $L[0]$-homogeneous vectors $u^i \in \mathcal{V}_{[k_i]}$, they are simply

$$e^{2\pi \mathrm{i} z_1 k_1} \cdots e^{2\pi \mathrm{i} z_n k_n} \mathrm{tr}_M Y_M(u^1, e^{2\pi \mathrm{i} z_1}) \cdots Y_M(u^n, e^{2\pi \mathrm{i} z_n})\, q^{L_0 - c/24}, \tag{5.3.17a}$$

where $u^i \in \mathcal{V}$ are the inserted states and $z_i \in \mathbb{C}$ are the points of insertion. As usual, the definition for nonhomogeneous vectors follows by linearity. By construction these functions automatically have period 1 in each z_i, and it is an easy calculation to verify that they also have period τ in each z_i, and thus the insertion points z_i lie on the torus $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$, as they should. In particular, the reader can verify that the one-point chiral blocks are indeed what we call the one-point functions: for $u \in \mathcal{V}_{[k]}$,

$$e^{2\pi \mathrm{i} z k}\, \mathrm{tr}_M Y_M(u, e^{2\pi \mathrm{i} z}) q^{L_0 - c/24} = \chi_M(\tau, u), \tag{5.3.17b}$$

hence the name of the latter. By the general principles of RCFT, the space of say one-point chiral blocks should carry a projective representation of the mapping class group of the once-punctured torus, i.e. $\mathrm{SL}_2(\mathbb{Z})$ (recall (4.3.9)), called *modular data* (Section 6.1.2). In Section 7.2.4 we find that a much larger group acts naturally on these one-point functions.

In (5.3.17) we inserted states u^i from only the vacuum sector. More generally, however, the states u^i can come from any sector, that is be vectors in any module $M \in \Phi(\mathcal{V})$. In that case the vertex operators Y_M should be replaced by intertwiners $\mathcal{Y}$ (Definition 6.1.9). Although this generalisation is fundamental to VOAs and RCFT, it is less so for Monstrous Moonshine (since $V^\natural$ is holomorphic).

The point of this subsection is to see in some detail how physics (RCFT) anticipates the statement and proof of Zhu's Theorem, to which we now turn.

5.3.5 The modularity of VOA characters

The most important property of the one-point functions is their modularity:

Theorem 5.3.8 (Zhu [574]) *Suppose $\mathcal{V}$ is a C_2-cofinite weakly rational VOA (see Definitions 5.3.2 and 5.3.5), and let $\Phi(\mathcal{V})$ be the finite set of irreducible $\mathcal{V}$-modules. Then there is a representation ρ of $SL_2(\mathbb{Z})$ by complex matrices $\rho(A)$ indexed by $\mathcal{V}$-modules $M, N \in \Phi(\mathcal{V})$, such that the one-point functions (5.3.13) obey*

$$\chi_M\left(\frac{a\tau + b}{c\tau + d}, v\right) = (c\tau + d)^n \sum_{N \in \Phi(\mathcal{V})} \rho\begin{pmatrix} a & b \\ c & d \end{pmatrix}_{MN} \chi_N(\tau, v) \tag{5.3.18a}$$

for any $v \in \mathcal{V}$ obeying $L[0]\, v = nv$ for some $n \in \mathbb{N}$ (see (5.3.16d)).

In particular, the graded dimensions (5.3.2) obey

$$\chi_M\left(\frac{a\tau + b}{c\tau + d}\right) = \sum_{N \in \Phi(\mathcal{V})} \rho\begin{pmatrix} a & b \\ c & d \end{pmatrix}_{MN} \chi_N(\tau), \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}). \tag{5.3.18b}$$

In (5.3.18), the quantity 'c' is an entry of a matrix in $\mathrm{SL}_2(\mathbb{Z})$ and should not be confused with the central charge. As we saw last subsection, $L[0]$ plays the role of L_0 in a Virasoro representation obtained from L_n by a change-of-variables z: $\mathcal{V} = \oplus_n \mathcal{V}_{[n]}$, where $n \in \mathbb{N}$ and $\mathcal{V}_{[n]}$ is the eigenspace of $L[0]$ with eigenvalue n. We can summarise (5.3.18a) by saying that $\chi_M(\tau, v)$ is a vector-valued modular form of weight n and multiplier ρ (recall Definition 2.2.2). We will summarise the proof of Theorem 5.3.8 shortly; see [**442**] for an independent argument.

One-point functions for the Moonshine module $\mathcal{V} = \mathcal{V}^\natural$ are studied in [**155**], where we find that all meromorphic modular forms for $\mathrm{SL}_2(\mathbb{Z})$ appear as some $\chi_{V^\natural}(\tau, v)$, provided the obvious constraints (namely that they be holomorphic in $\mathbb{H}$, have zero constant term in their q-expansion and have at worst a simple pole at $q = 0$) are satisfied – clearly, if the coefficient of q^α in $\chi_M(\tau)$ is zero then it must vanish in all other $\chi_M(\tau, v)$. Thus although we see the Monster in the graded dimension of $V^\natural$, we won't see it in most one-point functions of $V^\natural$.

However, if $v \in \mathcal{V}$ is fixed by some subgroup G_v of the automorphism group of $\mathcal{V}$, then the q^α coefficient of $\chi_M(\tau, v)$ relates to the representations of G_v and the eigenvalues of $o(v)|_{M_\alpha}$ (see Question 5.3.3). Note that in each homogeneous space $\mathcal{V}_n \neq 0$ there will be nonzero vectors invariant under the full automorphism group of $\mathcal{V}$ (why?). For example, we read off from Table 7.3 that in the homogeneous spaces $(V^\natural)_n$ of the Moonshine module for $0 \leq n \leq 7$, the $\mathbb{M}$-invariant subspace has dimension 1, 0, 1, 1, 2, 2, 4, 4, 7, respectively.

The representation ρ in Zhu's Theorem is called modular data (Section 6.1.2). The diagonal matrix $\rho\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ is given in (4.3.10). The matrix $S = \rho\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ relates to the fusion multiplicities $\mathcal{N}_{MN}^P$ via Verlinde's formula (6.1.1b) (at least for nice VOAs – see Section 6.2.2). It is conjectured that, for sufficiently nice VOAs, the representation ρ should be trivial on a congruence subgroup $\Gamma(N)$ (see the Congruence Property 6.1.7). When this is true, each graded dimension $\chi_M(\tau)$ will be a modular function for that $\Gamma(N)$.

If we weaken the hypothesis of weak rationality or C_2-cofiniteness (recall that these are conjectured to be equivalent) in Zhu's Theorem, then we can still recover some kind of modularity. In particular, physicists speak of *quasi-rational* CFTs, which are CFTs with finite fusions; in examples it seems that they still obey some weakened form of Zhu's Theorem (see Section 6.2.2).

Note that Zhu's Theorem is already strong enough to imply that the Moonshine module $V^\natural$ must have graded dimension $J(\tau)$. To see this, note that holomorphicity implies that $\rho(A)$ is a one-dimensional representation of $\mathrm{SL}_2(\mathbb{Z})$. However, $\rho\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ must be trivial and thus

$$\chi_{V^\natural}(A.\tau) = \chi_{V^\natural}(\tau), \qquad \forall A \in \mathrm{SL}_2(\mathbb{Z}).$$

We know $\chi_{V^\natural}(\tau)$ must be holomorphic in $\mathbb{H}$ (all graded dimensions are), has constant term 0 and a simple pole at the cusp. Therefore it equals $J(\tau)$. See also Question 5.3.4.

The proof of the Hauptmodul property for the other McKay–Thompson series T_g is much more subtle, unfortunately.

Zhu's Theorem rigorously generalises RCFT modularity to that of any sufficiently nice VOA. Its proof is long and complicated, but follows closely the intuition of CFT.

Zhu first defines abstractly a space of sequences $(S_1, S_2, \ldots)$ of functions, where each S_n maps n-tuples $(a_1, \ldots, a_n) \in \mathcal{V}^{\otimes_n}$ to meromorphic functions of $(z_1, \ldots, z_n, \tau) \in \mathbb{C}^n \times \mathbb{H}$. They obey several conditions, for example they are doubly-periodic in each variable z_i, with periods 1 and τ. Each function S_n is what we would call a chiral block on the torus $\mathbb{C}/(\mathbb{Z} + \mathbb{Z}\tau)$ with n marked points at z_i; it lies in the space $\mathfrak{B}^{(1,n)}_{\mathcal{V},\ldots,\mathcal{V}}$. Zhu's definition abstracts out the manifest properties of this space. It is immediate from his definition that $\mathrm{SL}_2(\mathbb{Z})$ acts on this space, in exactly the way we would expect from CFT. Verlinde's formula (6.1.2) tells us that the dimensions of these spaces should be independent of the number n of punctures, and in fact CFT tells us that a canonical basis for $\mathfrak{B}^{(1,n)}_{\mathcal{V},\ldots,\mathcal{V}}$ should be

$$(a_1, \ldots, a_n) \mapsto \mathrm{tr}_M\left(Y_M(a_1, e^{2\pi \mathrm{i} z_1}) \cdots Y_M(a_n, e^{2\pi \mathrm{i} z_n})\, q^{L_0}\right) \tag{5.3.19}$$

(appropriately normalised), for each irreducible $\mathcal{V}$-module M. However, showing rigorously that these functions (5.3.19) in fact satisfy his definition, and that they do indeed span his space, are both more difficult. But we see that the modularity in Zhu's Theorem arises through that $\mathrm{SL}_2(\mathbb{Z})$ action on the space of chiral blocks.

The modularity of the Jacobi characters $\chi^J_M(\tau, v, w)$ of Definition 5.3.7 is now easy.

Theorem 5.3.9 *Let $\mathcal{V}$ be a weakly rational C_2-cofinite VOA. Then the Jacobi characters $\chi^J_M(\tau, v, w)$ are holomorphic in $\mathbb{H}$ for any fixed v, w, and obey*

$$\chi^J_M\left(\frac{a\tau + b}{c\tau + d}, \frac{v}{c\tau + d}, w - c\frac{(v|v)}{2(c\tau + d)}\right) = \sum_{N \in \Phi(\mathcal{V})} \rho\begin{pmatrix} a & b \\ c & d \end{pmatrix}_{MN} \chi^J_N(\tau, v, w), \tag{5.3.20}$$

for all $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, $v \in \mathcal{V}_1$, and $w \in \mathbb{C}$, where ρ is as in Theorem 5.3.8 and where the inner-product $(v|v)$ is given by $v_{(1)}v = -(v|v)\mathbf{1}$.

Again, 'c' in (5.3.20) refers to a matrix entry and not the central charge. The transformation on the left side of (5.3.20) is exactly that of, for example, Jacobi theta functions. Theorem 5.3.9 is an easy corollary of the main theorem of [**426**] (which in turn is a corollary of the proof of Theorem 5.3.8 as given in [**574**]). In particular, define

$$Z_M(\tau, u, v) = \mathrm{tr}_M e^{2\pi \mathrm{i}\,(o(v) - (v|u)/2)} q^{L_0 + o(u) - (c + 12(u|u))/24}, \tag{5.3.21a}$$

for any $u, v \in \mathcal{V}_1$, so $\chi^J_M(\tau, v, w) = \exp[2\pi \mathrm{i} w]\, Z_M(\tau, 0, v)$. Then provided $o(v)u = 0$ (i.e. u and v commute in the Lie algebra $\mathcal{V}_1$), [**426**] obtained the transformation law

$$Z_M\left(\frac{a\tau + b}{c\tau + d}, u, v\right) = \sum_{N \in \Phi(\mathcal{V})} \rho\begin{pmatrix} a & b \\ c & d \end{pmatrix}_{MN} Z_N(\tau, cv + du, av + bu), \tag{5.3.21b}$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

for any $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. To prove (5.3.20), it suffices to prove it for the two generators $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and this follows directly from (5.3.21b). Holomorphicity of Z_M follows from Proposition 1.8 of [**151**].

5.3.6 Twisted #5: twisted modules and orbifolds

Last subsection we saw how the modularity of VOA modules permits a one-paragraph proof that the graded dimension of the Moonshine module $V^\natural$ must equal $J(\tau)$. How about the other McKay–Thompson series? In this subsection we find that the notion of $\mathcal{V}$-module must be generalised to the equally fundamental notion of *twisted* $\mathcal{V}$-modules. Twisted modules are vaguely reminiscent of projective representations of groups, but while a projective representation of G is a true representation of some central extension of G, a twisted $\mathcal{V}$-module is a true module of a vertex operator subalgebra of $\mathcal{V}$. Most groups don't have twisted modules, and VOAs don't seem to have a natural notion of a projective module, but Lie algebras have a foot in each camp and as we see in Chapter 3 have both kinds of modules.

Far from being an esoteric development, twisted modules are crucial to Monstrous Moonshine and absolutely central to the whole theory. In CFT and string theory, they arise in the important orbifold construction (Section 4.3.4). Twisted modules of Lie algebras – a baby example of twisted modules of VOAs – are discussed in Sections 1.5.4 and 3.4.1. Moonshine is the relation of VOAs to modular functions; the modular function analogue of this twisting has long been understood and also plays a central role (Section 2.3.3).

Fix a VOA $\mathcal{V}$ and any automorphism $g \in \mathrm{Aut}(\mathcal{V})$ of order N. We can define g-twisted modules [**185**], by blending together the definitions in Sections 3.4.1 and 5.3.1. In particular, decompose $\mathcal{V}$ into eigenspaces of g: $\mathcal{V} = \oplus_{j=0}^{N-1} \mathcal{V}^j$ where $\mathcal{V}^j = \{v \in \mathcal{V} \mid g.v = \xi_N^{-j} v\}$. A *$g$-twisted $\mathcal{V}$-module* (M, Y_M) has a $\mathbb{C}$-grading $M = \oplus_{\alpha \in \mathbb{C}} M_\alpha$, with $\dim M_\alpha < \infty$, as in Definition 5.3.1, as well as a linear map $\mathcal{V} \to \mathrm{End}[[z^{\pm 1/N}]]$, written $Y_M(u, z) = \sum_{r \in \mathbb{Z}/N} u_{(r)} z^{-r-1}$, such that (5.3.1a), (5.3.1c) hold,

$$Y(u, z) = \sum_{r \in -j/N + \mathbb{Z}} u_{(r)} z^{-r-1}, \qquad \forall u \in \mathcal{V}^j, \tag{5.3.22a}$$

and (5.3.1b) becomes

$$\begin{aligned} z_0^{-1} \delta\left(\frac{z_1 - z_2}{z_0}\right) Y_M(u, z_1) Y_M(v, z_2) - z_0^{-1} \delta\left(\frac{z_2 - z_1}{-z_0}\right) Y_M(v, z_2) Y_M(u, z_1) \\ = z_2^{-1} \left(\frac{z_1 - z_0}{z_2}\right)^{-j/N} \delta\left(\frac{z_1 - z_0}{z_2}\right) Y_M(Y(u, z_0)v, z_2), \end{aligned} \tag{5.3.22b}$$

where $u \in \mathcal{V}^j$. We say two g-twisted $\mathcal{V}$-modules M, N are *isomorphic* if there is an isomorphism $\varphi : M \to N$ satisfying $Y_M(\varphi v, z) = Y_N(v, z)\varphi$ for all $v \in N$. Note that an e-twisted $\mathcal{V}$-module (e being the identity of G) is an ordinary $\mathcal{V}$-module.

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

Any $h \in \mathrm{Aut}(\mathcal{V})$ permutes the twisted $\mathcal{V}$-modules as follows. Let M be g-twisted, and for each $v \in \mathcal{V}$ define

$$_hY_M(v, z) := Y_M(h.v, z).$$

Then $(M, {}_hY_M)$ is an $h^{-1}gh$-twisted $\mathcal{V}$-module. When h and g commute, we say the module (M, Y_M) is *h-stable* if (M, Y_M) and $(M, {}_hY_M)$ are isomorphic. We call $h \in \mathrm{Aut}(\mathcal{V})$ an *inner automorphism* of $\mathcal{V}$, and write $h \in \mathrm{Inn}(\mathcal{V})$, if every untwisted $\mathcal{V}$-module is h-stable.

Now let M be an irreducible g-twisted $\mathcal{V}$-module, and G any group of automorphisms $h \in \mathrm{Aut}(\mathcal{V})$ commuting with g such that M is h-stable for all $h \in G$. Then for each $h \in G$, we get an automorphism $\varphi(h) : M \to M$ of M, satisfying $\varphi(h)\, Y_M(v, z)\, \varphi(h)^{-1} = Y_M(h.v, z)$. Hence we can perform Thompson's trick (0.3.3) and write

$$\mathcal{Z}(M, h; \tau) := q^{-c/24}\, \mathrm{tr}_M \varphi(h)\, q^{L_0}. \tag{5.3.23}$$

These $\mathcal{Z}(M, h)$'s are the building blocks of the graded dimensions of various eigenspaces of h in M: for example, if h has order m, then the subspace of M fixed by the automorphism $\varphi(h)$ will have graded dimension $m^{-1} \sum_{i=1}^{m} \mathcal{Z}(M, h^i)$.

This assignment φ does not necessarily define a representation of G in $\mathrm{End}(M)$. However, $\varphi(h_2)^{-1}\varphi(h_1)^{-1}\varphi(h_1h_2)$ clearly commutes with all vertex operators $Y_M(v, z)$ and so by irreducibility of M is a scalar multiple $c_g(h_1, h_2)I$ of the identity. Equivalently, φ is a projective representation of G:

$$\varphi(h_1h_2) = c_g(h_1, h_2)\, \varphi(h_1)\, \varphi(h_2). \tag{5.3.24}$$

For any $h, k \in C_G(g)$ (i.e. commuting with g), $\varphi(khk^{-1}) = \alpha_{k,h}\varphi(k)\varphi(h)\varphi(k)^{-1}$ for some scalar $\alpha_{k,h}$, and thus $\mathcal{Z}(M, khk^{-1}; \tau) = \alpha_{k,h}\mathcal{Z}(M, h; \tau)$ by the cyclic property of trace. This means that, for fixed g, it suffices to restrict to one h from each $C_G(g)$-conjugacy class. By a similar argument (Question 5.3.6), we get that $\mathcal{Z}(M, h; \tau)$ vanishes identically, unless for all $k \in C_G(g)$ commuting with h, the 2-cocycle of (5.3.24) satisfies $c_g(h, k) = c_g(k, h)$. Thus we can further restrict to those h.

Conjecture 5.3.10 [136], [138], [152] *Suppose $\mathcal{V}$ is a weakly rational VOA, with exactly n irreducible $\mathcal{V}$-modules $M_1, \ldots, M_n$. Fix any finite subgroup G of* $\mathrm{Inn}(\mathcal{V})$. *Then:*

(a) *For any $g \in \mathrm{Inn}(\mathcal{V})$, there will be exactly n irreducible g-twisted $\mathcal{V}$-modules $M_1^g, \ldots, M_n^g$. Moreover, each M_i^g has a conformal weight $h_i^g \in \mathbb{Q}$ as in Lemma 5.3.3, and any g-twisted $\mathcal{V}$-module is completely reducible into a direct sum of the M_i^g. Labelling the modules appropriately, we get $(M_i^g, {}_hY_{M_i^g}) \cong (M_i^{h^{-1}gh}, Y_{M_i^{h^{-1}gh}})$. This defines a projective representation $\varphi(h)$ of the centraliser $C_G(g)$ as in (5.3.24).*

(b) *For each commuting pair $g, h \in G$, define $\mathcal{Z}^i_{(g,h)}(\tau) := \mathcal{Z}(M_i^g, h; \tau)$. Then each $\mathcal{Z}^i_{(g,h)}(\tau)$ is holomorphic in $\mathbb{H}$, and is a modular function for (i.e. is fixed by) some congruence subgroup. For any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, there exist scalars $a(A, g, h)_{ij}$*

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

such that

$$\mathcal{Z}^i_{(g,h)}\left(\frac{a\tau+b}{c\tau+d}\right)=\sum_{j=1}^{n}a(A,g,h)_{ij}\,\mathcal{Z}^j_{(g^ah^c,g^bh^d)}(\tau). \tag{5.3.25}$$

(c) *Let $\mathcal{V}^G$ be the vertex operator subalgebra consisting of all $v\in\mathcal{V}$ fixed by all elements of G. Then the $\mathbb{C}$-span of the graded dimensions of all nontwisted $\mathcal{V}^G$-modules will equal that of all $\mathcal{Z}^i_{(g,h)}(\tau)$ for commuting $g,h\in G$, and the total number of irreducible $\mathcal{V}^G$-modules will equal n times the sum, over representatives g of all conjugacy classes in G, of the number of inequivalent irreducible projective representations of $C_G(g)$ with 2-cocycle c_g as in (5.3.24).*

(d) *In the special case that $\mathcal{V}$ is holomorphic (i.e. $n=1$), $\mathrm{Inn}(\mathcal{V})=\mathrm{Aut}(\mathcal{V})$ and the coefficients a_{ij} in (5.3.25) are roots of unity. There is a 3-cocycle $\alpha\in H^3(G,U_1(\mathbb{C}))$ such that the 2-cocycle c_g of (5.3.24) is given by*

$$c_g(h_1,h_2)=\alpha(g,h_1,h_2)\,\alpha(h_1,h_2,g)\,\alpha(h_1,g,h_2)^*.$$

Some progress towards this important conjecture is provided by, for example, [**150**]. Monstrous Moonshine is interested in the holomorphic case (i.e. $n=1$), which is by far the best understood; we return to it in Section 6.2.4. The number of irreducible projective representations in (c) is described in Section 3.1.1. We find in (d) that the cohomology group $H^3(G,U_1(\mathbb{C}))\cong H^4(G,\mathbb{Z})$ (trivial action of G on the coefficients) classifies all the possibilities for the orbifold; the analogous result for nonholomorphic VOAs is much more subtle, being more sensitive to the structure of $\mathcal{V}$, and is still poorly understood.

Part (c) leads us to a Galois theory for $\mathcal{V}^G$. But considering the depth of Jones' Galois theory for subfactors, and the 'Galois theory' for lattices sketched in Section 2.3.5, it is clear that a far more interesting theory is possible for VOAs. It would certainly be interesting to develop this.

The easiest examples of the orbifold construction are of a self-dual lattice VOA $\mathcal{V}(L)$ by a subgroup G of the automorphism group of L (see e.g. [**150**]). We learn in Section 5.2.2 that there is a deep analogy between lattices and VOAs. This orbifold construction of VOAs corresponds directly to the shift construction of lattices outlined in Section 2.3.3.

The most famous VOA, the Moonshine module $V^\natural$, was the original orbifold. Frenkel–Lepowsky–Meurman [**201**] obtained it as the orbifold of the Leech lattice VOA $\mathcal{V}(\Lambda)$ by the ±1-symmetry of Λ. Since Λ is self-dual, $\mathcal{V}(\Lambda)$ is holomorphic. As predicted by Conjecture 5.3.10, there is a unique -1-twisted $\mathcal{V}(\Lambda)$-module. We discuss this orbifold more in Sections 4.3.4 and 7.2.1; see also [**201**] for details.

Question 5.3.1. Let $\mathcal{V}$ be any VOA, and let $\mathcal{W}$ be a vector space and $T:\mathcal{V}\to\mathcal{W}$ be any isomorphism of vector spaces. Use this linear map T to carry the VOA structure on $\mathcal{V}$ to one on $\mathcal{W}$.

Question 5.3.2. Let $\mathcal{V}$ be any VOA and let $\mathcal{V}_{[n]}$ be the grading induced by $L[0]$ in (5.3.16d). Prove for any $N\geq 0$, that

$$\oplus_{n=0}^{N}\mathcal{V}_n=\oplus_{n=0}^{N}\mathcal{V}_{[n]}.$$

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

Question 5.3.3. Find an expression for the coefficient of the q^α term in the one-point function $\chi_M(\tau, v)$, using the representation theory of the stabiliser $G_v < \mathrm{Aut}\,\mathcal{V}$ and the eigenvalues of the zero-mode $o(v)$ restricted to the homogeneous space M_α.

Question 5.3.4. Let $\mathcal{V}$ be any holomorphic weakly rational C_2-cofinite VOA with central charge $c = 24$. Prove that its graded dimension $\chi_{\mathcal{V}}(\tau)$ must equal $J(\tau) + c$, where the constant c is $\dim \mathcal{V}_1$.

Question 5.3.5. (a) Relate the Jacobi character $\chi^J_{\mathcal{V}(L)}$ of a lattice VOA $\mathcal{V}(L)$, for L positive-definite and with even integer inner-products, and the theta series Θ_L of (2.3.7). (b) Relate the Jacobi character $\chi^J_{M(\lambda)}$ of an irreducible module of an integrable affine VOA $\mathcal{V}(\mathfrak{g}, k)$, for $\overline{\mathfrak{g}}$ simple, with the affine algebra character χ_λ of (3.2.9a).

Question 5.3.6. Let M be g-twisted. Show that the series $\mathcal{Z}(M, h)$ of (5.3.23) is identically 0, unless $h \in C_G(g)$ has the property that, for all $k \in C_G(g)$ commuting with h, $c_g(h, k) = c_g(k, h)$. (*Hint*: first show that $\mathcal{Z}(M, hk)$ is identically 0 if $c_g(h, k) \neq c_g(k, h)$; then use the 2-cocycle condition (3.1.1b).)

5.4 Geometric incarnations

Vertex (operator) algebras are a deep construct and, in spite of their complexity, are here to stay. In this section we describe some connections with geometry. Section 5.4.1 describes the programme to rigorously construct CFTs in Segal's sense (Section 4.4.1), from VOAs. Section 5.4.2 reviews the geometric side of vertex operator superalgebras.

5.4.1 Vertex operator algebras and Riemann surfaces

The introductory chapter stated that the physics of Moonshine exploits the duality between Hamilton's and Feynman's pictures of CFT. Manin put it this way back in 1985:

> The quantum theory of (super)strings exists at present in two entirely different mathematical fields. Under canonical quantization it appears to a mathematician as the representation theory of algebras of Heisenberg, Virasoro, and Kac–Moody and their superextensions. Quantization with the help of the Polyakov path integration leads to the analytic theory of algebraic (super)curves and their moduli spaces, to invariants of the type of the analytic curvature, *etc*. Establishment of direct mathematical connections between these two forms of a single theory is a big and important problem. [**402**]

Our best answer to Manin is the theory of geometric vertex operator algebras.

Note that any time we have an algebraic structure with a binary operation (e.g. 'product') ab, we can express multiple products using *binary trees*, which keep track of the brackets. For example, the binary trees in Figure 5.1 correspond to the products XY and $A((BC)D)$, respectively. The external (i.e. valance 1) vertices are assigned vectors,

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

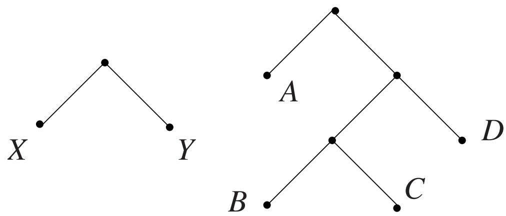

Fig. 5.1 Some binary trees.

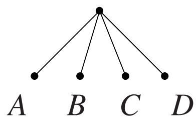

Fig. 5.2 Associativity.

while each internal vertex corresponds to a single product. Different algebraic structures can be axiomatised from this 'geometric' point of view. For instance, if the product is associative (e.g. we have a group), then it doesn't matter where we place the brackets – for example, the above *ABCD*-binary tree can be replaced with the tree in Figure 5.2.

More interesting for us are the geometrical axioms for Lie algebras [**294**]. Let V be any Lie algebra. Then to any binary tree with n legs corresponds a linear map φ from n copies $V \otimes \cdots \otimes V$ of the vector space V, to V. The map corresponding to the *ABCD*-binary tree of Figure 5.1 takes the Lie algebra vectors A, B, C, D to the nested Lie bracket $[A[[BC]D]]$. It is then fairly straightforward to encode all properties of the Lie algebra in the language of trees. For example, anti-commutativity says that if we flip the two descendents of an inner vertex of the tree – for example, in Figure 5.1 flipping D with the 3-vertex tree containing B and C – then the corresponding maps φ differ by a factor of -1. Gluing the root (uppermost vertex) of one tree to an external vertex of another corresponds in the Lie algebra to inserting one nested bracket into the middle of another. The only nontrivial property is anti-associativity (see Question 5.4.2). The result is a formulation of Lie algebra that is completely equivalent to the usual algebraic one [**294**].

Now, if we 'two-dimensionalise' that definition of 'geometric Lie algebra', we get something called a *geometric VOA* [**295**] that is equivalent to the 'algebraic' VOA of Definition 5.1.3. In place of binary trees (Figure 5.1), we have spheres with tubes (Figure 5.3). Equivalently, a sphere with n tubes is the Riemann sphere with n marked points and a choice of local coordinate at each point – an enhanced surface of type $(0,n)$ (Section 2.1.4). The moduli space of binary trees with n legs is a finite set, but the moduli space of spheres with n tubes is an infinite-dimensional complex space. To each such sphere with tubes we get a linear map φ from n copies of our vector space $\mathcal{V}$ (which is our VOA) to $\mathcal{V}$ (or rather a certain completion of $\mathcal{V}$ – a complication caused by the infinite-dimensionality of $\mathcal{V}$). A geometric VOA satisfies meromorphicity requirements, and most importantly the sewing axiom. In fact this map φ is Segal's functor $\mathcal{S}$ described in Section 4.4.1.

https://doi.org/10.1017/9781009401548.006 Published online by Cambridge University Press

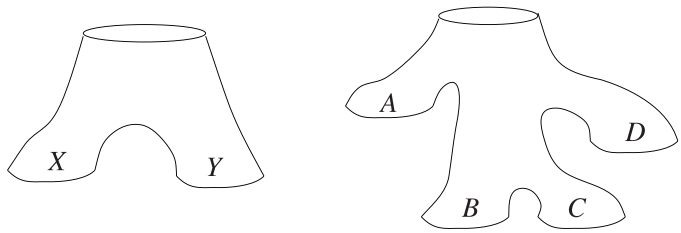

Fig. 5.3 The surfaces corresponding to Figure 5.1.

The point is that the resulting notion of geometric VOA is equivalent to that of algebraic VOA [**295**], though it takes considerable effort to show this. Thus a VOA is an 'algebra' with a two-dimensional analogue of a binary operation. In particular, let P_w be the simplest pair-of-pants, namely the Riemann sphere $\mathbb{PC}^1$ with marked points 0, ∞ and w and local coordinates given by z, $1/z$ and $z - w$ (z being the global coordinate on $\mathbb{C}$). Then the formal series $Y(u, w)v$ corresponds to $\mathcal{S}(P_w)(u \otimes v)$. On the other hand, consider the annulus, that is the Riemann sphere with marked points 0 and ∞, with local coordinates z and $\exp[-\epsilon z^{-1}\mathrm{d}/\mathrm{d}z\, z^{-1}]$. Recalling the realisation $-z^{-1}\mathrm{d}/\mathrm{d}z = \ell_{-2} \in \mathfrak{Witt}$ and the formula $\omega = L_{-2}\mathbf{1}$, we can recover the conformal vector ω by differentiating with respect to ϵ the map obtained from $\mathcal{S}$. The Virasoro algebra is fundamental here, capturing the effect of changing local coordinates (recall (5.3.15c)), and is responsible for the meromorphicity in the geometric VOA. The Jacobi identity (5.1.7a) is obtained from the sewing axiom. This equivalence relates formal power series (algebra) to distribution theory (analysis). It proves that the chiral blocks $\langle v, Y(u_1, z_1) \cdots Y(u_n, z_n)v' \rangle$ will be meromorphic, except for poles at $z_i = z_j$.

As mentioned before, a group corresponds to trees such as Figure 5.2. We can also two-dimensionalise that, and obtain what Huang calls a *vertex group* [**293**]. The easiest examples are $\mathbb{C}^\times$ and the enhanced moduli space $\widehat{\mathfrak{M}}_{0,1}$. Vertex groups should be to VOAs what Lie groups are to Lie algebras.

The motivation for this deep work is to construct examples satisfying Segal's definitions of CFT and modular functor. We know at present no nontrivial examples, although the general belief is that any sufficiently nice VOA will provide one. Huang's work [**295**] establishes this for genus 0, and more recently he has pushed it to genus 1 [**297**].

We end this subsection on a more speculative note [**560**], [**295**]. According to Witten, to understand string theory conceptually, we need a new analogue of Riemannian geometry. In contrast to the more classical 'particle-math', there is a more modern 'string-math'. We have the real numbers (particle physics) versus the complex numbers (string theory); binary trees versus spheres with tubes; Lie algebras versus VOAs; the representation theory of Lie algebras versus RCFT, etc. What are the stringy analogues of calculus, ordinary differential equations, Riemannian manifolds, the Atiyah–Singer Index theorem, . . . ? Huang suggests that just as we could imagine Moonshine as a mystery that is explained in some way by RCFT, perhaps the stringy version of calculus would similarly explain the mystery of two-dimensional gravity, stringy ODEs would explain the mystery of infinite-dimensional integrable systems, stringy Riemannian manifolds would help explain the mystery of mirror symmetry, and the stringy index theorem would help explain the elliptic genus (for this latter possibility, consider the work of Tamanoi reviewed next subsection).

What makes this more subtle is that complexification is not unique. To give a simple example, S^1 can be thought of as the real projective space $\mathbb{P}^1(\mathbb{R})$ and as the Lie group $\mathrm{SO}_2(\mathbb{R})$. The obvious complexification of $\mathbb{P}^n(\mathbb{R})$ is $\mathbb{P}^n(\mathbb{C})$. An obvious complexification of $\mathrm{SO}_n(\mathbb{R})$ is $\mathrm{SO}_n(\mathbb{C})$. But if we think of $\mathrm{O}_n(\mathbb{R})$ more geometrically as the real matrices that preserve the quadratic form $x_1^2 + \cdots + x_n^2$, then its complexification should be those complex matrices that preserve the Hermitian form $|x_1|^2 + \cdots + |x_n|^2$, i.e. $\mathrm{U}_n(\mathbb{C})$. Thus the complexifications of S^1 in these cases would be the 2-sphere $\mathbb{P}^1(\mathbb{C})$, the cylinder $\mathrm{SO}_2(\mathbb{C})$ (i.e. the multiplicative group $\mathbb{C}/\{0\}$) and the 3-sphere $\mathrm{SU}_2(\mathbb{C})$ (as a real Lie group). So the specific complexification obtained depends on the context. In all cases, the way to proceed is to convert the defining relations of the given object into symbols that make sense over $\mathbb{C}$.

What sense can we make of the statement that the complexification of a binary tree is a sphere with punctures? Consider the simplest case: the segment $0 \leq x \leq 1$. This can be thought of geometrically as the locus $(a, b, c) \in \mathbb{R}^3$ satisfying $a + b^2 - 1 = a - c^2 = 0$. Over the complex numbers the parameter 'a' is redundant, and this locus has the obvious complexification $w^2 + z^2 = 1$. We know this is a sphere with two punctures, that is, a cylinder as we would like it to be.

Incidentally, Arnol'd speculates that there is in fact a triality: the reals, the complex numbers and the quaternions. He discusses several examples in [**18**], as well as some applications of this thought. This suggests that there is a third structure, generalising vertex algebras much as vertex algebras generalise Lie algebras.

5.4.2 Vertex operator superalgebras and manifolds

Through the work of Witten and others, we have discovered that much can be learned about a space X, by studying a string theory living in X. Much of this is reviewed in [**291**]. For example, to a Calabi–Yau manifold X [**299**], [**571**] and an element of its complexified Kähler cone, string theory associates two $N = 2$ superconformal field theories, called the A and B models (which focus on respectively the Kähler and complex structures of X). To clarify (and rigorise) these ideas, Malikov–Schechtman–Vaintrob [**401**] suggested how one may construct, given X, the vertex algebra of the $N = 2$ superconformal field theory (the A model) associated with X. This work is clearly fundamental. We can only sketch it here.

To any smooth complex variety X, reference [**401**] associates a sheaf of $(N = 1)$ vertex operator superalgebras, called the *chiral de Rham complex* $\mathcal{MSV}_X$. In other words, to every open set $U \subset X$, there is a near-vertex operator superalgebra $\mathcal{MSV}_X(U)$ (the 'space of sections of $\mathcal{MSV}_X$ over U'). Whenever sets $U \subset V$ are open, there is a surjective restriction map $r_U^V : \mathcal{MSV}_X(V) \to \mathcal{MSV}_X(U)$, which is a homomorphism of near-vertex operator superalgebras. We briefly discuss vertex operator superalgebras in Section 5.1.3. These near-vertex operator superalgebras are bi-graded, by commuting operators L_0 (the Hamiltonian) and J_0 (the fermionic charge) with eigenvalues $\mathbb{N}$ and $\mathbb{Z}$, respectively. They form a complex in the sense that there is a differential Q_{BRST} obeying $Q^2_{BRST} = 0$ and increasing fermionic charge by 1. When the open set U is homeomorphic to an open ball in $\mathbb{C}^n$, then $\mathcal{MSV}_X(U)$ is essentially the tensor product of n copies of what string theory calls a bosonic $(\beta\gamma)$ ghost system (similar to the Heisenberg VOA), with n copies of a (bc) fermionic ghost system. The physics of these ghost systems is described in [**463**].

The prototypical example of 'sheaf' is the structure sheaf $\mathcal{O}_X$, which associates with each open set U the space of functions $f : U \to \mathbb{C}$. The prototypical example of 'complex' is the de Rham complex, given by the space of differential forms on X, with a differential d obeying $\mathrm{d}^2 = 0$ and taking p-forms to $p+1$-forms. Of course the point of a complex is to take the cohomology $H^* = \ker \mathrm{d}/\mathrm{im}\,\mathrm{d}$. The books [**537**] are a readable introduction to algebraic geometry; in particular section 2.2 provides elementary examples of sheaves, and section 6.1 treats sheaf cohomology. For a sheaf $\mathcal{F}$ over X, $H^0(X, \mathcal{F})$ is always the global section $\mathcal{F}(X)$, and it is common for the other $H^i(X, \mathcal{F})$ to all vanish. The name 'chiral de Rham complex' was chosen because the $L_0 = 0$ subspace can be identified with the familiar space of differential forms ('chiral' refers to the chiral algebra of Section 4.3.2 or the chiral ring discussed in [**291**]).

In the case of $\mathcal{MSV}_X$, the sheaf cohomology $H^*(X, \mathcal{MSV}_X)$ yields the global section $\mathcal{MSV}_X(X)$, which is a near-vertex operator superalgebra. The case where X is Calabi–Yau is the most interesting, as $\mathcal{MSV}_X(X)$ has $N = 2$ (rather than merely $N = 1$) supersymmetry, which makes it much richer. $\mathcal{MSV}_X(X)$ is a fundamental invariant associated with X, and much information of X can be recovered from it. For example, the usual de Rham cohomology $H^*_{DR}(X)$ of X is $H^*(\mathcal{MSV}_X(X); Q_{BRST})$. For another example, the elliptic genus (discussed shortly) of X equals $\mathrm{tr}_{\mathcal{MSV}_X(X)} q^{L_0} y^{J_0}$ [**81**].

Elliptic genus appeared in the mid-1980s in both string theory and topology. For details see, for example, [**287**], [**499**], [**523**]. In Thom's cobordism ring Ω, elements are equivalence classes of cobordant manifolds, addition is connected sum and multiplication is Cartesian product. The *universal elliptic genus* $\phi(M)$ is a ring homomorphism from $\mathbb{Q} \otimes \Omega$ to the ring of power series in q, which sends n-dimensional manifolds with spin connections (see [**369**] for the relevant geometry) to a weight $n/2$ modular form of $\Gamma_0(2)$ with integer coefficients. Several variations and generalisations have been introduced, for example, the Witten genus assigns spin manifolds with vanishing first Pontrjagin class a weight $n/2$ modular form of $\mathrm{SL}_2(\mathbb{Z})$ with integer coefficients. On a finite-dimensional manifold M, the index of the Dirac operator (in the heat kernel interpretation) is a path integral in supersymmetric quantum mechanics, that is an integral over the loop space $\mathcal{L}M = \{\gamma : S^1 \to M\}$; the string theory version of this is that the index of the Dirac operator on $\mathcal{L}M$ should be an integral over $\mathcal{L}(\mathcal{L}M)$, that is over smooth maps of tori into M, and this (heuristically) is just the elliptic genus, and explains why it should be modular.

The important rigidity property of the Witten genus with respect to any compact Lie group action on the manifold is a consequence of the modularity of the characters of affine algebras (our Theorem 3.2.3) [**388**]. In physics, elliptic genera arise as partition functions of $N = 2$ superconformal field theories [**561**]. The Witten genus (normalised by η^8) of the Milnor–Kervaire manifold M_0^8, an eight-dimensional manifold built from the E_8 diagram, equals $j^{\frac{1}{3}}$ [**287**]. Also, the elliptic genus of even-dimensional projective spaces $\mathbb{P}^{2n}(\mathbb{C})$ unexpectedly has only nonnegative coefficients and in fact equals the graded dimension of a certain vertex algebra [**400**]; this suggests interesting representation-theoretic questions in the spirit of Monstrous Moonshine. Exciting developments are described in [**517**], including relations with von Neumann (sub)factors.

Related to $\mathcal{MSV}_X$ must be the work of Tamanoi [**521**]. The index of an operator d is $\ker d - \operatorname{coker} d$; we can interpret this geometrically as the superdimension associated with the 'superpair' $(\ker d; \operatorname{coker} d)$ of vector spaces. This is what Tamanoi does with the elliptic genus. In particular, to each closed Riemannian manifold X he associates a vertex operator superalgebra $\mathcal{T}(X)$, determined from its geometry. It has a nonnegative half-integer grading and central charge $N = \dim X/2$. The Riemannian metric of X yields the conformal vector ω. In the special case of a Kähler manifold, the Kähler forms (i.e. the closed real differential forms of type (1,1)) form a level 1 representation of the affine algebra ${D_N}^{(1)}$. Again, the elliptic genus is recovered as the graded dimension of $\mathcal{T}(X)$. It is obviously desirable to relate these invariants $\mathcal{T}(X)$ and $\mathcal{MSV}_X(X)$. We return to elliptic genus in Section 7.3.7.

Question 5.4.1. Find a complexification for the Möbius band.

Question 5.4.2. In a non-associative algebra, the ambiguous product $v_1 \cdots v_n$ can only be evaluated when the $n-1$ pairs of brackets are placed. Let L be any Lie algebra. Prove that for any $n \geq 3$, L has an identity of the form

$$v_1 \cdots v_n = v_1 v_n v_2 v_3 \cdots v_{n-1} + \cdots + v_1 \cdots v_i v_n v_{i+1} \cdots v_{n-1} + \cdots + v_1 v_2 \cdots v_{n-1} v_n.$$

More precisely, for any choice of bracketing on the left, prove that there is a choice of bracketing for each of the $n-1$ terms on the right such that the resulting formula holds for any $v_i \in L$. For example, $[[v_1 v_2] v_3]$ equals $[[v_1 v_3] v_2] + [v_1 [v_2 v_3]]$ and $[[v_1 [v_2 v_3]] v_4]$ equals $[[v_1 v_4][v_2 v_3]] + [v_1 [[v_2 v_4] v_3]] + [v_1 [v_2 [v_3 v_4]]]$.