Introduction to vertex operator algebras

Shinji Koshida*

September 10, 2023

Abstract

This is the lecture note for the course on vertex operator algebras I taught from October 2022 to February 2023 at Aalto University.

Update history

- 17.2.2023. The last lecture. All previous lecture notes are combined.
- 10.9.2023. Cleaned up lecture notes. Links to exercises.

Contents

*shinji.koshida@aalto.fi

1 Introduction

1.1 Reference guide

There are many textbooks of vertex operator algebras. The ones I did not necessarily follow to prepare this note, but often helpful include [LL04, Kac98, FBZ04].

1.2 Notation/convention

Here will be collected possibly uncommon notations.

- $\mathbb{Z} = \{\dots, -2, -1, -, 1, 2, \dots\}$: set of integers
- $\mathbb{N} = \{0, 1, 2, \dots\}$: set of non-negative integers

Throughout the note, a vector space is a complex vector space unless otherwise specified.

2 Lie algebras and representation

2.1 Enveloping algebra

Definition 2.1. A Lie algebra is a vector space $\mathfrak{g}$ together with a bilinear map $[-,-]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that satisfies

- (skew-symmetry) $[x, y] = -[y, x]$, $x, y \in \mathfrak{g}$,
- (Jacobi identity) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$, $x, y, z \in \mathfrak{g}$

Remark 2.2. The bilinear map $[-,-]$ above is often called the Lie bracket of the Lie algebra $\mathfrak{g}$.

Remark 2.3. We can phrase the Jacobi identity in slightly a different way. By the skew-symmetry, we get

$$[x, [y, z]] - [y, [x, z]] = [[x, y], z], \quad x, y, z \in \mathfrak{g}.$$

We write $\mathrm{ad}(x) := [x, -] \in \mathrm{End}(\mathfrak{g})$, $x \in \mathfrak{g}$ for the adjoint action. As $z \in \mathfrak{g}$ is arbitrary in the above equation, the Jacobi identity can be equivalently stated (under the skew-symmetry) as

$$\mathrm{ad}(x)\mathrm{ad}(y) - \mathrm{ad}(y)\mathrm{ad}(x) = \mathrm{ad}(\mathrm{ad}(x)(y)), \quad x, y \in \mathfrak{g}.$$

Example 2.4. For an associative algebra A, if we define a bilinear map by

$$[a, b] := ab - ba, \quad a, b \in A,$$

then $(A, [-, -])$ forms a Lie algebra.

Example 2.5. The vector space $\mathfrak{sl}_n = \{X \in M_n(\mathbb{C}) | \mathrm{Tr} X = 0\}$ along with the bilinear map

$$[X, Y] := XY - YX, \quad X, Y \in \mathfrak{sl}_n$$

is a Lie algebra. Notice that, in this example, $\mathfrak{sl}_n$ is **not** an associative algebra under the matrix product.

Definition 2.6. Let $(\mathfrak{g}, [-, -]_\mathfrak{g})$, $(\mathfrak{h}, [-, -]_\mathfrak{h})$ be Lie algebras. A linear map $f\colon \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras if the Lie bracket is preserved, i.e.,

$$[f(x), f(y)]_\mathfrak{h} = f([x, y]_\mathfrak{g}), \quad x, y \in \mathfrak{g}$$

is satisfied.

For a vector space V, we can always construct the **tensor algebra** over V as follows. As a vector space, it is given by

$$T(V) := \bigoplus_{n=0}^{\infty} V^{\otimes n},$$

where, as convention, we set $V^{\otimes 0} = \mathbb{C}$. The product structure is induced by

$$V^{\otimes m} \times V^{\otimes n} \to V^{\otimes (m+n)}; \quad (v, w) \mapsto v \otimes w$$

Then, the product is associative and $1 \in V^{\otimes 0}$ is the unit.

Since a Lie algebra $\mathfrak{g}$ is a vector space, we can apply this construction to $\mathfrak{g}$ and obtain a tensor algebra $T(\mathfrak{g})$, but so far, the algebra structure is nothing to do with the Lie algebra structure of $\mathfrak{g}$. We define $\mathcal{I}_\mathfrak{g}$ as the two-sided ideal of $T(\mathfrak{g})$ generated by

$$x \otimes y - y \otimes x - [x, y], \quad x, y \in \mathfrak{g}$$

and set

$$\mathcal{U}(\mathfrak{g}) := T(\mathfrak{g}) / \mathcal{I}_\mathfrak{g}.$$

This algebra is called the (universal) enveloping algebra of $\mathfrak{g}$.

Proposition 2.7. *Let $\mathfrak{g}$ be a Lie algebra and $\mathcal{U}(\mathfrak{g})$ be its enveloping algebra. The linear map defined by*

$$\sigma\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g}); \, x \mapsto x + \mathcal{I}_\mathfrak{g}$$

is a homomorphism of Lie algebras.

Proof. The goal is to show

$$[\sigma(x), \sigma(y)] = \sigma([x, y])$$

for all $x, y \in \mathfrak{g}$. The left hand side is

$$\sigma(x)\sigma(y) - \sigma(y)\sigma(x) = x \otimes y - y \otimes x + \mathcal{I}_{\mathfrak{g}},$$

while the right hand side is

$$[x, y] + \mathcal{I}_{\mathfrak{g}}.$$

They are the same element in $\mathcal{U}(\mathfrak{g})$ by definition of the ideal $\mathcal{I}_{\mathfrak{g}}$. □

2.2 Poincaré–Birkhoff–Witt theorem

Let $\mathfrak{g}$ be a Lie algebra and $\{x_i\}_{i \in I}$ be a basis of $\mathfrak{g}$. We fix a linear order $\leq$ of the index set I. We introduce a set

$$\Omega = \{\boldsymbol{i} = (i_1 \leq i_2 \leq \cdots \leq i_k) \in I^k | k \in \mathbb{N}\},$$

where $I^0 = \{\emptyset\}$.

Theorem 2.8 (Poincaré–Birkhoff–Witt (PBW))**.** *The enveloping algebra* $\mathcal{U}(\mathfrak{g})$ *has a basis*

$$\sigma(x_{i_1})\sigma(x_{i_2}) \cdots \sigma(x_{i_k}), \quad \boldsymbol{i} = (i_1 \leq i_2 \leq \cdots \leq i_k) \in \Omega,$$

where we agree that $\emptyset \in \Omega$ *corresponds to* $1 \in \mathcal{U}(\mathfrak{g})$*.*

Proof. See e.g. [Hum12]. □

In particular, $\sigma\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ is injective, and we can consider $\mathfrak{g}$ as a subspace of $\mathcal{U}(\mathfrak{g})$. Because of that, we often simply write $x \in \mathcal{U}(\mathfrak{g})$ in the sense of $\sigma(x)$.

It is often convenient to have the PBW theorem in another format. A typical situation where we use the PBW theorem is that the Lie algebra $\mathfrak{g}$ decomposes into Lie subalgebras:

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2. \tag{2.1}$$

Here, $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are supposed to be Lie subalgebras of $\mathfrak{g}$, but notice that the above decomposition is merely **as a vector space**. Nonetheless, each Lie subalgebra has its own enveloping algebra, and we have the following version of the PBW theorem:

Corollary 2.9. *Assume the decomposition (2.1) into Lie subalgebras as above. Then, have an isomorphism*

$$\mathcal{U}(\mathfrak{g}) = \mathcal{U}(\mathfrak{g}_1) \otimes \mathcal{U}(\mathfrak{g}_2)$$

of $(\mathcal{U}(\mathfrak{g}_1), \mathcal{U}(\mathfrak{g}_2))$*-bimodules.*

2.3 Representation of $\mathfrak{g}$ and $\mathcal{U}(\mathfrak{g})$-module

Let $\mathfrak{g}$ is a Lie algebra. Recall that a representation of $\mathfrak{g}$ is a vector space V along with a Lie algebra homomorphism

$$\rho\colon \mathfrak{g} \to \mathrm{End}(V).$$

Notice that $\mathrm{End}(V)$ is an associative algebra, so is a Lie algebra as was seen in Example 2.4.

By the universality of the enveloping algebra (Exercise 1.1.1), there exists a unique homomorphism $\rho'\colon \mathcal{U}(\mathfrak{g}) \to \mathrm{End}(V)$ such that $\rho = \rho' \circ \sigma$. Using this ρ', we define

$$\mathcal{U}(\mathfrak{g}) \times V \to V; \quad (A, v) \mapsto \rho'(A)v,$$

which equips the vector space V with the structure of a $\mathcal{U}(\mathfrak{g})$-module.

Conversely, suppose that a $\mathcal{U}(\mathfrak{g})$-module V is given. Since $\mathcal{U}(\mathfrak{g})$ is a $\mathbb{C}$-algebra, V is already a vector space. If we define a linear map

$$\rho\colon \mathfrak{g} \to \mathrm{End}(V); \quad x \mapsto (v \mapsto \sigma(x)v),$$

we can check that ρ is a homomorphism of Lie algebras because so is σ. Therefore, the pair (V, ρ) is a representation of $\mathfrak{g}$.

In the sequel of this text, we will not necessarily distinguish representations of $\mathfrak{g}$ from $\mathcal{U}(\mathfrak{g})$-modules.

2.4 Example: infinite dimensional Heisenberg algebra

Let us study one particular example, namely the infinite dimensional Heisenberg algebra. It is the infinite dimensional vector space

$$\widehat{\mathfrak{h}} = \Big(\bigoplus_{n\in\mathbb{Z}} \mathbb{C}\alpha_n\Big) \oplus \mathbb{C}K$$

equipped with the Lie bracket

$$\begin{aligned} [\alpha_m, \alpha_n] &= m\delta_{m+n,0}K, \quad m, n \in \mathbb{Z}, \\ [K, \widehat{\mathfrak{h}}] &= \{0\}. \end{aligned}$$

The pair $(\widehat{\mathfrak{h}}, [-,-])$ is indeed a Lie algebra (Exercise 1.2.1).

We are in particular interested in representations of $\widehat{\mathfrak{h}}$ of the following form: let $\lambda \in \mathbb{C}$ and set

$$\mathcal{F}_\lambda := \mathcal{U}(\widehat{\mathfrak{h}})/\mathcal{I}_\lambda,$$

where $\mathcal{I}_\lambda$ is the left $\mathcal{U}(\widehat{\mathfrak{h}})$-ideal defined by

$$\mathcal{I}_\lambda = \sum_{n>0} \mathcal{U}(\widehat{\mathfrak{h}})\alpha_n + \mathcal{U}(\widehat{\mathfrak{h}})(\alpha_0 - \lambda) + \mathcal{U}(\widehat{\mathfrak{h}})(K - 1).$$

Notice that $\mathcal{F}_\lambda$ is the quotient by a **left** ideal, so the product of $\mathcal{U}(\widehat{\mathfrak{h}})$ induces a natural structure of a **left** $\mathcal{U}(\widehat{\mathfrak{h}})$-module:

$$\mathcal{U}(\widehat{\mathfrak{h}}) \times \mathcal{F}_\lambda \to \mathcal{F}_\lambda; \quad (A, B + \mathcal{I}_\lambda) \mapsto AB + \mathcal{I}_\lambda.$$

Next we study a basis of $\mathcal{F}_\lambda$. Since we need a basis of the quotient space $\mathcal{U}(\widehat{\mathfrak{h}})/\mathcal{I}_\lambda$, once we can decompose the vector space $\mathcal{U}(\widehat{\mathfrak{h}})$ in the form

$$\mathcal{U}(\widehat{\mathfrak{h}}) = U \oplus \mathcal{I}_\lambda$$

with some vector space U, then a basis of U induces a basis of the quotient space. To find such a decomposition, the PBW theorem in the form of Corollary 2.9 is useful.

We define three subspaces $\widehat{\mathfrak{h}}_+$, $\widehat{\mathfrak{h}}_0$, $\widehat{\mathfrak{h}}_-$ of $\widehat{\mathfrak{h}}$ by

$$\widehat{\mathfrak{h}}_\pm = \bigoplus_{\pm n > 0} \mathbb{C}\alpha_n, \quad \widehat{\mathfrak{h}}_0 = \mathbb{C}\alpha_0 \oplus \mathbb{C}K,$$

but it is clear that they are Lie subalgebras of $\widehat{\mathfrak{h}}$ and give a decomposition

$$\widehat{\mathfrak{h}} = \widehat{\mathfrak{h}}_- \oplus \widehat{\mathfrak{h}}_0 \oplus \widehat{\mathfrak{h}}_+$$

of a vector space. Thus, Corollary 2.9 gives a decomposition of the enveloping algebra

$$\mathcal{U}(\widehat{\mathfrak{h}}) = \mathcal{U}(\widehat{\mathfrak{h}}_-) \otimes \mathcal{U}(\widehat{\mathfrak{h}}_0) \otimes \mathcal{U}(\widehat{\mathfrak{h}}_+).$$

We can first notice that the Lie algebra $\widehat{\mathfrak{h}}_0$ is commutative, so its enveloping algebra is isomorphic to the ring of polynomials $\mathbb{C}[\alpha_0, K]$. In particular, it decomposes into

$$\mathcal{U}(\widehat{\mathfrak{h}}_0) = \mathbb{C} \cdot 1 \oplus (\mathcal{U}(\widehat{\mathfrak{h}}_0)(\alpha_0 - \lambda) + \mathcal{U}(\widehat{\mathfrak{h}}_0)(K - 1)).$$

It is also the case that the enveloping algebra $\mathcal{U}(\widehat{\mathfrak{h}}_+)$ admits the following decomposition (Exercise 1.1.3):

$$\mathcal{U}(\widehat{\mathfrak{h}}_+) = \mathbb{C} \cdot 1 \oplus \mathcal{U}(\widehat{\mathfrak{h}}_+)\widehat{\mathfrak{h}}_+.$$

Combining these two decompositions, we obtain

$$\mathcal{U}(\widehat{\mathfrak{h}}) = \mathcal{U}(\widehat{\mathfrak{h}}_-) \oplus \mathcal{U}(\widehat{\mathfrak{h}}_-)(\mathcal{U}(\widehat{\mathfrak{h}}_0)(\alpha_0 - \lambda) + \mathcal{U}(\widehat{\mathfrak{h}}_0)(K - 1)) \oplus \mathcal{U}(\widehat{\mathfrak{h}})\widehat{\mathfrak{h}}_+,$$

but the latter two direct sum components precisely form the left ideal $\mathcal{I}_\lambda$.

Lemma 2.10. *We have*

$$\mathcal{I}_\lambda = \mathcal{U}(\widehat{\mathfrak{h}}_-)(\mathcal{U}(\widehat{\mathfrak{h}}_0)(\alpha_0 - \lambda) + \mathcal{U}(\widehat{\mathfrak{h}}_0)(K - 1)) \oplus \mathcal{U}(\widehat{\mathfrak{h}})\widehat{\mathfrak{h}}_+.$$

Proof. The inclusion of the direction "$\supset$" is clear. As

$$\mathcal{U}(\widehat{\mathfrak{h}})\widehat{\mathfrak{h}} = \sum_{n>0} \mathcal{U}(\widehat{\mathfrak{h}})\alpha_n,$$

it suffices to show

$$\mathcal{U}(\widehat{\mathfrak{h}})(\alpha_0 - \lambda) + \mathcal{U}(\widehat{\mathfrak{h}})(K - 1)$$

lies in the right-hand side. This follows from the fact that α_0 and K are central in $\widehat{\mathfrak{h}}$ and by applying the Corollary 2.9. $\square$

Therefore, a basis of $\mathcal{U}(\widehat{\mathfrak{h}}_-)$ induces a basis of $\mathcal{F}_\lambda$. Of course, $\mathcal{U}(\widehat{\mathfrak{h}}_-)$ has the following basis:

$$\alpha_{-n_1} \cdots \alpha_{-n_l}, \quad n_1 \geq \cdots \geq n_l > 0,\, l \in \mathbb{N},$$

which induces a basis

$$\alpha_{-n_1} \cdots \alpha_{-n_l} \left|\lambda\right\rangle, \quad n_1 \geq \cdots \geq n_l > 0,\, l \in \mathbb{N}$$

of $\mathcal{F}_\lambda$, where $\left|\lambda\right\rangle := 1 + \mathcal{I}_\lambda$.

Let us make a few observations on the structure of $\mathcal{F}_\lambda$:

- K acts as $\mathrm{Id}_{\mathcal{F}_\lambda}$.
- α_0 acts as $\lambda \cdot \mathrm{Id}_{\mathcal{F}_\lambda}$.
- $\alpha_n \left|\lambda\right\rangle = 0$, $n > 0$.

It is also useful to picture the space $\mathcal{F}_\lambda$ in the following way:

$$\begin{array}{ccc} & \left|\lambda\right\rangle & \\ & & \\ & \alpha_{-1}\left|\lambda\right\rangle & \\ & & \\ \alpha_{-1}^2\left|\lambda\right\rangle & & \alpha_{-2}\left|\lambda\right\rangle \\ & & \\ \alpha_{-1}^3\left|\lambda\right\rangle & \alpha_{-2}\alpha_{-1}\left|\lambda\right\rangle & \alpha_{-3}\left|\lambda\right\rangle \\ & & \\ & \vdots & \end{array}$$

In this picture, the vertical direction indicate the eigenvalues of the operator $D \in \mathrm{End}(\mathcal{F}_\lambda)$ defined by

$$D\, \alpha_{-n_1} \cdots \alpha_{-n_k} \left|\lambda\right\rangle := (n_1 + \cdots + n_k)\alpha_{-n_1} \cdots \alpha_{-n_l} \left|\lambda\right\rangle$$

for $n_1 \geq \cdots \geq n_l > 0$, $l \in \mathbb{N}$.

Proposition 2.11. *For $n \in \mathbb{Z}$, $[D, \alpha_n] = -n\alpha_n$.*

Proof. Suppose that $n > 0$. Notice that, since $\widehat{\mathfrak{h}}_-$ is a commutative Lie algebra, a vector of the form

$$\alpha_{-n}\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle , \quad n_1 \geq \cdots \geq n_l > 0$$

is already one of the prescribed basis vectors of $\mathcal{F}_\lambda$, so we can act D on it by definition as

$$\begin{aligned} D\alpha_{-n}\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle =&(n + n_1 + \cdots + n_l)\alpha_{-n}\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle \\ =&n\alpha_{-n}\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle + \alpha_{-n}D\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle , \end{aligned}$$

from which $[D, \alpha_{-n}] = n\alpha_{-n}$ follows.

Assume that in the sequence $n_1 \geq \cdots \geq n_l > 0$, n can be found m times:

$$n_1 \geq \cdots \geq n_i > \underbrace{n = \cdots = n}_{m} > n_{i+m+1} \geq \cdots \geq n_l.$$

In this case, applying the commutation relations of the Heisenberg algebra, we get

$$\alpha_n\alpha_{-n_1} \cdots \alpha_{-n_k} |\lambda\rangle = nm\alpha_{-n_1} \cdots \alpha_{-n_i}(\alpha_{-n})^{m-1}\alpha_{-n_{i+m+1}} \cdots \alpha_{-n_l} |\lambda\rangle .$$

Now, we can act D on the right-hand side to find

$$\begin{aligned} &D\alpha_n\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle \\ =&(n_1 + \cdots + n_l - n)nm\alpha_{-n_1} \cdots \alpha_{-n_i}(\alpha_{-n})^{m-1}\alpha_{-n_{i+m+1}} \cdots \alpha_{-n_l} |\lambda\rangle \\ =&(n_1 + \cdots + n_l - n)\alpha_n\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle \\ =&- n\alpha_n\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle + \alpha_n D\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle \end{aligned}$$

proving the identity $[D, \alpha_n] = -n\alpha_n$. ☐

It is clear from the definition of the operator D that the space $\mathcal{F}_\lambda$ admits the following decomposition:

$$\mathcal{F}_\lambda = \bigoplus_{n=0}^{\infty}(\mathcal{F}_\lambda)_{(n)}, \quad (\mathcal{F}_\lambda)_{(n)} := \mathrm{Ker}(D - n \cdot \mathrm{Id}_{\mathcal{F}_\lambda}),$$

but Proposition 2.11 also tells us that each operator α_n has a clear degree in terms of this decomposition as

$$\alpha_n(\mathcal{F}_\lambda)_{(m)} \subset (\mathcal{F}_\lambda)_{(m-n)}, \quad m, n \in \mathbb{Z}.$$

Of course, we understand $(\mathcal{F}_\lambda)_{(n)} = 0$ when $n < 0$. This gives us the following truncation property.

Corollary 2.12. *For any $v \in \mathcal{F}_\lambda$,*

$$\alpha_n v = 0, \quad n \gg 0,$$

where "$n \gg 0$" reads "for all sufficiently large n".

Let us state Corollary 2.12 in a different way. We consider the following formal series

$$\alpha(x) = \sum_{n\in\mathbb{Z}} \alpha_n x^{-n-1} \in \mathrm{End}(\mathcal{F}_\lambda)[[x^{\pm 1}]] = \left\{ \sum_{n\in\mathbb{Z}} T_n x^n \middle| T_n \in \mathrm{End}(\mathcal{F}_\lambda) \right\}.$$

Here α_n is considered its action on $\mathcal{F}_\lambda$. Then, the formal series $\alpha(x)$ can act on any vector $v \in \mathcal{F}_\lambda$ by

$$\alpha(x)v = \sum_{n\in\mathbb{Z}} \alpha_n v\, x^{-n-1} \in \mathcal{F}_\lambda[[x^{\pm 1}]].$$

Corollary 2.12 is also stated as follows.

Corollary 2.13. *For any $v \in \mathcal{F}_\lambda$,*

$$\alpha(x)v \in \mathcal{F}_\lambda((x)) = \left\{ \sum_{n=N}^{\infty} v_n x^n \middle| v_n \in \mathcal{F}_\lambda,\, N \in \mathbb{Z} \right\}.$$

Let us observe one more property of the formal series. When we introduce two distinct formal variables, the composition

$$\alpha(x_1)\alpha(x_2) = \sum_{m,n\in\mathbb{Z}} \alpha_m \alpha_n x_1^{-m-1} x_2^{-n-1} \in \mathrm{End}(\mathcal{F}_\lambda)[[x_1^{\pm 1}, x_2^{\pm 1}]]$$

naturally makes sense. Then the commutator $[\alpha(x_1), \alpha(x_2)]$ also makes sense and is computed as

$$\begin{aligned} [\alpha(x_1), \alpha(x_2)] &= \sum_{m,n\in\mathbb{Z}} [\alpha_m, \alpha_n] x_1^{-m-1} x_2^{-n-1} \\ &= \sum_{m\in\mathbb{Z}} m x_1^{-m-1} x_2^{m-1} \\ &= \frac{\partial}{\partial x_2} \left(\sum_{m\in\mathbb{Z}} x_1^{-m-1} x_2^{m} \right) \\ &= \frac{\partial}{\partial x_2} \left(x_1^{-1} \delta(x_2/x_1) \right), \end{aligned} \tag{2.2}$$

where we set $\delta(x) = \sum_{m\in\mathbb{Z}} x^m$. This way of expressing the commutation relations of the Heisenberg algebra will be essential in the future.

3 Formal calculus

In the last section 2, we have already encountered the following situation. Let $\widehat{\mathfrak{h}}$ be the infinite dimensional Heisenberg algebra and $\mathcal{F}_\lambda$ be the $\mathcal{U}(\widehat{\mathfrak{h}})$-module associated with $\lambda \in \mathbb{C}$. We saw the formal sum of operators

$$\alpha(x) = \sum_{n\in\mathbb{Z}} \alpha_n x^{-n-1} \in \mathrm{End}(\mathcal{F}_\lambda)[[x^{\pm 1}]],$$

and that when it is acted on a vector $v \in \mathcal{F}_\lambda$,

$$\alpha(x)v = \sum_{n\in\mathbb{Z}} \alpha_n v x^{-n-1} \in \mathcal{F}_\lambda((x)).$$

Furthermore, the formal series is composable in the following way:

$$\alpha(x_1)\alpha(x_2) = \sum_{m,n\in\mathbb{Z}} \alpha_m \alpha_n x_1^{-m-1} x_2^{-n-1} \in \mathrm{End}(\mathcal{F}_\lambda)[[x_1^{\pm 1}, x_2^{\pm 1}]]$$

and the commutator becomes

$$[\alpha(x_1), \alpha(x_2)] = \frac{\partial}{\partial x_2} \left(x_1^{-1} \delta(x_2/x_1) \right).$$

In this section, we are going to see more on this kind of formal calculus, which will be needed in vertex operator algebras.

3.1 Single variable case

Let V be a vector space. The space of formal series with coefficients in V is

$$V[[x^{\pm 1}]] := \left\{ \sum_{n\in\mathbb{Z}} v_n x^n \middle| v_n \in V \; (n \in \mathbb{Z}) \right\}.$$

By defining addition and scalar product coefficient-wise, we can think of $V[[x^{\pm 1}]]$ as a vector space. As a subspace of $V[[x^{\pm 1}]]$, the space of formal Laurent series is

$$V((x)) := \left\{ \sum_{n=-N}^{\infty} v_n x^n \middle| v_n \in V \; (n \geq -N), \, N \in \mathbb{Z} \right\},$$

and the space of formal power series is

$$V[[x]] := \left\{ \sum_{n=0}^{\infty} v_n x^n \middle| v_n \in V \; (n \in \mathbb{N}) \right\}.$$

Occasionally, we might also need the space of polynomials $V[x]$ and the space of Laurent polynomials $V[x^{\pm 1}]$.

Why formal series needed?

The following are all elements of $\mathbb{C}[[x^{\pm 1}]]$:

(1) $\sum_{n=0}^{\infty} x^n$: converging in $\mathbb{D} = \{z \in \mathbb{C} | |z| < 1\}$.

(2) $\sum_{n\in\mathbb{Z}} (-1)^n q^{\binom{n}{2}} x^n$ ($|q| < 1$): converging in $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$.

(3) $\sum_{n=0}^{\infty} n! x^n$: not converging.

(1) and (2) both converge to functions but on different domains, and (3) never converges. Already in this sense, elements of $\mathbb{C}[[x^{\pm 1}]]$ are considered generalization of functions, but let us make this point more precise.

First we notice the following fact.

Lemma 3.1. *The natural product $x^m \cdot x^n = x^{m+n}$ along with the scalar product on V gives rise to a bilinear map*

$$V[[x^{\pm 1}]] \times \mathbb{C}[x^{\pm 1}] \to V[[x^{\pm 1}]]; \quad (f(x), g(x)) \mapsto f(x)g(x).$$

Proof. We expand each as $f(x) = \sum_{n\in\mathbb{Z}} f_n x^n$ and $g(x) = \sum_{n=-N}^{N} g_n x^n$ with some $N \geq 0$, and see that

$$f(x)g(x) = \sum_{m\in\mathbb{Z}} \sum_{n=-N}^{N} f_m g_n x^{m+n} = \sum_{m\in\mathbb{Z}} \left(\sum_{n=-N}^{N} f_{m-n} g_n \right) x^m. \tag{3.1}$$

Each coefficient of x^m is indeed a finite sum. $\square$

Let us introduce the following notation for the functional taking the coefficient of a certain power:

Definition 3.2. For $f(x) \in V[[x^{\pm 1}]]$, the coefficient of x^n in $f(x)$ will be denoted by $[x^n]f(x)$.

Now, we can understand the space $V[[x^{\pm}]]$ as the space of functionals on $\mathbb{C}[x^{\pm}]$ with values in V.

Proposition 3.3. *For each $f \in V[[x^{\pm 1}]]$, we define the linear functional $\varphi_f \colon \mathbb{C}[x^{\pm 1}] \to V$*

$$\varphi_f(g(x)) = [x^{-1}](f(x)g(x)) \quad (g(x) \in \mathbb{C}[x^{\pm 1}]).$$

Then, for $f_1(x), f_2(x) \in V[[x^{\pm 1}]]$, $f_1(x) = f_2(x)$ if and only if $\varphi_{f_1} = \varphi_{f_2}$.

Proof. The desired result follows from the fact that

$$\varphi_f(x^{-n-1}) = [x^{-1}](f(x)x^{-n-1}) = [x^n]f(x)$$

holds for any $f(x) \in V[[x^{\pm 1}]]$ and $n \in \mathbb{Z}$. $\square$

Proposition 3.3 suggests that, in particular, we can regard $\mathrm{End}(V)[[x^{\pm 1}]]$ as a space of operator-valued distributions with test functions in $\mathbb{C}[x^{\pm 1}]$. Therefore, a 'quantum field' might live there.

Square of distribution?

It is always difficult to make sense of the square of a distribution, which is also the case in the formal setting.

Let us take as an example the infinite dimensional Heisenberg algebra $\widehat{\mathfrak{h}}$ and its representation $\mathcal{F}_\lambda$. The naive square of $\alpha(x)$ might be

$$\alpha(x)^2 \stackrel{?}{=} \sum_{m,n\in\mathbb{Z}} \alpha_m\alpha_n x^{-m-n-2} = \sum_{m\in\mathbb{Z}} \left(\sum_{n\in\mathbb{Z}} \alpha_{m-n}\alpha_n\right) x^{-m-2}.$$

The coefficient of x^{-2} is $\sum_{n\in\mathbb{Z}} \alpha_{-n}\alpha_n$, but if each of the summand is acted on $|\lambda\rangle$, we get

$$\alpha_{-n}\alpha_n |\lambda\rangle = \begin{cases} 0, & n > 0, \\ \lambda |\lambda\rangle, & n = 0, \\ (\alpha_n\alpha_{-n} - n) |\lambda\rangle = -n |\lambda\rangle, & n < 0. \end{cases}$$

Therefore, $\sum_{n\in\mathbb{Z}} \alpha_{-n}\alpha_n$ is the sum of infinitely many non-zero operators, which does not make sense in $\mathrm{End}(\mathcal{F}_\lambda)$.

There two solutions to this problem:

(1) Normally ordered product (explained later in Section 4).

(2) Use two variables (already used in Section 2).

3.2 Two-variable case

The space of formal series in two variables is

$$V[[x_1^{\pm1}, x_2^{\pm1}]] = \left\{ \sum_{n_1,n_2\in\mathbb{Z}} v_{n_1n_2} x_1^{n_1} x_2^{n_2} \middle| v_{n_1n_2} \in V \; (n_1, n_2 \in \mathbb{Z}) \right\}.$$

Similarly to the single-variable case, we will also encounter the following subspaces:

$$V((x_1^{\pm1}, x_2^{\pm1})) = \left\{ \sum_{n_1,n_2\geq -N} v_{n_1n_2} x_1^{n_1} x_2^{n_2} \middle| v_{n_1n_2} \in V \; (n_1, n_2 \geq -N),\ N \in \mathbb{Z} \right\},$$

$$V[[x_1, x_2]] = \left\{ \sum_{n_1,n_2=0}^{\infty} v_{n_1n_2} x_1^{n_1} x_2^{n_2} \middle| v_{n_1n_2} \in V \; (n_1, n_2 \in \mathbb{N}) \right\}$$

as well as $V[x_1^{\pm1}, x_2^{\pm1}]$, $V[x_1, x_2]$.

Delta series

Although the delta series $\delta(x) = \sum_{n\in\mathbb{Z}} x^n$ can be defined in a single variable, we rarely see it in that form. We rather define

$$\delta(x_1/x_2) = \sum_{n\in\mathbb{Z}} x_1^n x_2^{-n} \in \mathbb{C}[[x_1^{\pm 1}, x_2^{\pm 1}]].$$

The following property explains why it is "delta":

Proposition 3.4. *For $f(x) \in V[[x^{\pm 1}]]$, we have the identity*

$$f(x_1)\delta(x_1/x_2) = f(x_2)\delta(x_1/x_2)$$

in $V[[x_1^{\pm 1}, x_2^{\pm 1}]]$.

Proof. Let us expand $f(x) = \sum_{n\in\mathbb{Z}} f_n x^n$. Then, we can see that

$$\begin{aligned} f(x_1)\delta(x_1/x_2) &= \sum_{n\in\mathbb{Z}} f_n x_1^n \sum_{m\in\mathbb{Z}} x_1^m x_2^{-m} \\ &= \sum_{n,m\in\mathbb{Z}} f_{n-m} x_1^n x_2^{-m} \\ &= \sum_{m\in\mathbb{Z}} f_m x_2^m \sum_{n\in\mathbb{Z}} x_1^n x_2^{-n} \\ &= f(x_2)\delta(x_1/x_2). \end{aligned}$$

Notice that the expression in the second line proves that both sides of the equation lie in $V[[x_1^{\pm 1}, x_2^{\pm 1}]]$. $\square$

Convention of binomial expansion

We need to agree on a convention of expanding rational functions. Essentially, it comes down to how to interpret $(x_1 - x_2)^n$ with some (possibly negative) $n \in \mathbb{Z}$ as a formal series.

Definition 3.5 (Binomial expansion)**.** For $n \in \mathbb{Z}$, we define $(x_1 - x_2)^n$ as an element of $\mathbb{C}[[x_1^{\pm}, x_2^{\pm}]]$ by

$$(x_1 - x_2)^n = \sum_{k=0}^{\infty} \binom{n}{k} x_1^{n-k}(-x_2)^k.$$

Here, the binomial coefficients are given by

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}, \quad k \in \mathbb{Z}_{\geq 0}.$$

In other words, we always expand the rational function $(x_1 - x_2)^n$ in the region $|x_1| > |x_2|$. Let us observe some properties following from the definition:

- If $n \geq 0$, $(x_1 - x_2)^n = (-x_2 + x_1)^n$ is a polynomial.
- For $m, n \in \mathbb{Z}$,

$$(x_1 - x_2)^m (x_1 - x_2)^n = (x_1 - x_2)^{m+n}.$$

  In fact, both sides are expansions of the same rational function in the same region. (So this is a result of complex analysis.)

- For a similar reason, for $n \geq 0$,

$$(x_1 - x_2)^{-n-1} = \partial_{x_2}^{(n)} (x_1 - x_2)^{-1} = (-1)^n \partial_{x_1}^{(n)} (x_1 - x_2)^{-1},$$

  where $\partial_x^{(n)} = \frac{1}{n!} \frac{\partial^n}{\partial x^n}$ is the divided power of a differential operator.

- By direct computation,

$$\begin{aligned}
&(x_1 - x_2)^{-1} - (-x_2 + x_1)^{-1} \\
=& \sum_{k=0}^{\infty} \binom{-1}{k} x_1^{-1-k} (-x_2)^k - \sum_{k=0}^{\infty} \binom{-1}{k} (-x_2)^{-1-k} x_1^k \\
=& \sum_{k=0}^{\infty} x_1^{-k-1} x_2^k + \sum_{k=0}^{\infty} x_2^{-k-1} x_1^k \\
=& \sum_{n \in \mathbb{Z}} x_1^{-n-1} x_2^n \\
=& x_1^{-1} \delta(x_2/x_1).
\end{aligned}$$

  Here we applied the formula $\binom{n}{k} = (-1)^k \binom{-n+k-1}{k}$ at $n = -1$.

- Combining the last two properties, we get, for $n \geq 0$,

$$\begin{aligned}
(x_1 - x_2)^{-n-1} - (-x_2 + x_1)^{-n-1} =& \partial_{x_2}^{(n)} \left(x_1^{-1} \delta(x_2/x_1) \right) \\
=& (-1)^n \partial_{x_1}^{(n)} \left(x_1^{-1} \delta(x_2/x_1) \right).
\end{aligned}$$

  In particular, $(x_1 - x_2)^{-n-1} \neq (-x_2 + x_1)^{-n-1}$.

- However, if we multiply sufficiently large power of $x_1 - x_2$, the derivative of the delta distribution can be cleared: if $m \geq n + 1$,

$$\begin{aligned}
&(x_1 - x_2)^m \partial_{x_2}^{(n)} \left(x_1^{-1} \delta(x_2/x_1) \right) \\
=& (x_1 - x_2)^m \left((x_1 - x_2)^{-n-1} - (-x_2 + x_1)^{-n-1} \right) \\
=& (x_1 - x_2)^{m-n-1} - (-x_2 + x_1)^{m-n-1} \\
=& 0. \qquad (3.2)
\end{aligned}$$

Taylor's formula

It will be convenient to keep in mind the formal analogue of Taylor's formula.

Proposition 3.6 (Taylor's formula)**.** *For $f(x) \in V[[x^{\pm 1}]]$,*

$$f(x + x_1) = \sum_{k=0}^{\infty} \partial_x^{(k)} f(x) x_1^k$$

in $V[[x^{\pm 1}, x]]$. Here the left-hand-side is understood in term of the binomial expansion in Definition 3.5.

Proof. The result follows from direct computation: if $f(x) = \sum_{n \in \mathbb{Z}} f_n x^n$

$$\begin{aligned} f(x + x_1) &= \sum_{n \in \mathbb{Z}} f_n (x + x_1)^n \\ &= \sum_{n \in \mathbb{Z}} f_n \sum_{k=0}^{\infty} \binom{n}{k} x^{n-k} x_1^k \\ &= \sum_{k=0}^{\infty} \partial_x^{(k)} \left(\sum_{n \in \mathbb{Z}} f_n x^n \right) x_1^k, \end{aligned}$$

which is exactly the right-hand-side of the desired identity. ☐

3.3 More variables

In practice, we will often use the following version of the delta series that involves three variables:

$$\delta\left(\frac{x_1 - x_2}{x_0}\right) = \sum_{k=0}^{\infty} (-x_2)^k \partial_{x_1}^{(k)} \delta(x_1/x_0) = \sum_{n \in \mathbb{Z}} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k x_0^{-n} x_1^{n-k} x_2^k,$$

which lies in $\mathbb{C}[[x_0^{\pm 1}, x_1^{\pm 1}, x_2^{\pm 1}]]$. (Only positive powers of x_2 appear, though.)

One typical usage of this delta distribution is as follows:

Proposition 3.7. *Let $f(x_0, x_1, x_2) \in V((x_0, x_1, x_2))$ be a Laurent series in three variables. Then*

$$f(x_0, x_1, x_2)\delta\left(\frac{x_1 - x_2}{x_0}\right) = f(x_1 - x_2, x_1, x_2)\delta\left(\frac{x_1 - x_2}{x_0}\right) \tag{3.3}$$

in $V[[x_0^{\pm 1}, x_1^{\pm 1}, x_2^{\pm 1}]]$.

Remark 3.8. To interpret Proposition 3.7, it is convenient to assume that $f(x_0, x_1, x_2)$ is a Laurent polynomial. In this case, when we substitute $x_0 = x_1 - x_2$, the result is a rational function of x_1 and x_2 with poles possibly at $x_1 = 0$, $x_2 = 0$, and $x_1 = x_2$. On the right-hand side of (3.3), the same rational function is expanded in the region $|x_1| > |x_2|$. Therefore, we can interpret that the multiplication by $\delta\left(\frac{x_1 - x_2}{x_0}\right)$ results in the substitution $x_0 = x_1 - x_2$ and expansion in the region $|x_1| > |x_2|$ at the same time.

Proof of Proposition 3.7. We first check that the left-hand side of (3.3) actually makes sense. Let us expand $f(x_0, x_1, x_2)$ as

$$f(x_0, x_1, x_2) = \sum_{r,s,t\in\mathbb{Z}} f_{r,s,t} x_0^r x_1^s x_2^t, \quad f_{r,s,t} \in V.$$

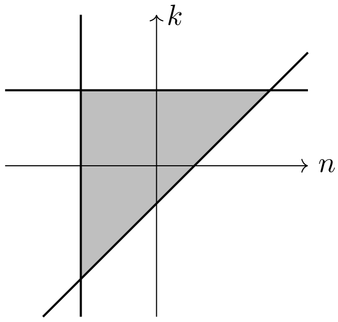
Since $f(x_0, x_1, x_2) \in V((x_0, x_1, x_2))$, there is $N \in \mathbb{Z}$ such that $f_{r,s,t} = 0$ whenever one of r, s, t is less than $-N$. Together with the explicit expansion of $\delta\left(\frac{x_1-x_2}{x_0}\right)$, the left-hand side of (3.3) becomes

$$f(x_0, x_1, x_2)\delta\left(\frac{x_1 - x_2}{x_0}\right) = \sum_{r,s,t\in\mathbb{Z}} \left(\sum_{n\in\mathbb{Z}} \sum_{k=0}^{\infty} \binom{n}{k} (-1)^k f_{r+n,s-n+k,t-k} \right) x_0^r x_1^s x_2^t.$$

The question is if, for each $r, s, t \in \mathbb{Z}$, the coefficient of $x_0^r x_1^s x_2^t$ is a finite sum. Notice that any pair (n, k) that contributes to the sum must satisfy

$$n \geq -r - N, \quad k - n \geq -s - N, \quad k \leq t + N,$$

but there are only finitely many such pairs:



The rest is a straightforward manipulation:

$$\begin{aligned}
&f(x_0, x_1, x_2)\delta\left(\frac{x_1 - x_2}{x_0}\right) \\
&= \sum_{r,s,t\in\mathbb{Z}} \sum_{k=0}^{\infty} f_{r,s,t} x_0^r x_1^s x_2^t (-x_2)^k \partial_{x_1}^{(k)} \delta(x_1/x_0) \\
&= \sum_{r,s,t\in\mathbb{Z}} \sum_{k=0}^{\infty} f_{r,s,t} x_1^s x_2^t (-x_2)^k \partial_{x_1}^{(k)} \left(x_1^r \delta(x_1/x_0)\right) \\
&= \sum_{r,s,t\in\mathbb{Z}} f_{r,s,t} x_1^s x_2^t (x_1 - x_2)^r \delta\left(\frac{x_1 - x_2}{x_0}\right) \\
&= f(x_1 - x_2, x_1, x_2)\delta\left(\frac{x_1 - x_2}{x_0}\right).
\end{aligned}$$

The manipulation only involves Propositions 3.4 and 3.6, both of which were proved either by change of summation indices or change of order of summation. This ensures that the right-hand side also lies in the correct space of formal series. □

4 Heisenberg vertex algebra

Rather than introducing vertex operator algebras, we start with vertex algebras (without "operator") that are defined by fewer axioms. Among other vertex algebras, this section is devoted to constructing the simplest non-trivial one, namely, the Heisenberg vertex algebra. Construction of more complicated vertex algebras will be discussed in the next section 5.

4.1 Definition of a vertex algebra

Let us see the definition of a vertex algebra right away.

Definition 4.1. A vertex algebra is a triple $(V, \mathbf{1}, Y)$ of

- a $\mathbb{C}$-vector space V,
- a distinguished vector $\mathbf{1} \in V$,
- and a linear map

$$Y(-,x)\colon V \to \mathrm{End}(V)[[x^{\pm 1}]]; \quad a \mapsto Y(a,x) = \sum_{n\in\mathbb{Z}} a_{(n)} x^{-n-1}$$

satisfying the following axiom set.

(VA1) Field condition: For any $a, b \in V$, we have

$$Y(a,x)b \in V((x)).$$

In other words, $a_{(n)}b = 0$ for $n \gg 0$.

(VA2) Vacuum axiom: We have $Y(\mathbf{1},x) = \mathrm{Id}_V = \mathrm{Id}_V \cdot x^0$. For any $a \in V$, we have

$$Y(a,x)\mathbf{1} \in a + V[[x]]x.$$

(VA3) Translation axiom: Define $T \in \mathrm{End}(V)$ by $T(a) := a_{(-2)}\mathbf{1}$, $a \in V$. Then,

$$[T, Y(a,x)] = \frac{d}{dx} Y(a,x), \quad a \in V.$$

(VA4) Locality: For any $a, b \in V$,

$$(x_1 - x_2)^N [Y(a,x_1), Y(b,x_2)] = 0, \quad N \gg 0.$$

We call the distinguished vector $\mathbf{1}$ the **vacuum vector** and the linear map $Y(-,x)$ the **state-field correspondence map**.

Example 4.2 (Commutative vertex algebra). Let $V = \mathbb{C}[t^{\pm 1}]$ be the space of Laurent polynomials in t. In this space, we can take $\mathbf{1} := 1$ and the linear map $Y(-,x)\colon V \to \mathrm{End}(V)[[x^{\pm 1}]]$ defined by

$$Y(p(t),x)q(t) = p(t+x)q(t) = \sum_{j=0}^{\infty} \partial^{(j)}p(t) \cdot q(t)\, x^j.$$

Then, the triple $(V, \mathbf{1}, Y)$ forms a vertex algebra (Exercise 2.1.1.)

Unfortunately, there is no other obvious example of a vertex algebra. Nevertheless, we hope to have a vertex algebra associated to the infinite-dimensional Heisenberg algebra $\widehat{\mathfrak{h}}$ as we have already seen a few pieces of evidence. In fact, on any Fock representation $\mathcal{F}_\lambda$, the operator-valued formal series

$$\alpha(x) = \sum_{n\in\mathbb{Z}} \alpha_n x^{-n-1} \in \mathrm{End}(\mathcal{F}_\lambda)[[x^{\pm 1}]]$$

satisfies the property required by **(VA1)**: $\alpha(x)v \in \mathcal{F}_\lambda((x))$ for all $v \in \mathcal{F}_\lambda$ (Corollary 2.13). Furthermore, if we combine the calculus in (2.2) and (3.2), we can see that the commutator $[\alpha(x), \alpha(y)]$ is annihilated by multiplying $(x-y)^2$:

$$(x-y)^2[\alpha(x), \alpha(y)] = 0,$$

which could be part of **(VA4)**.

Therefore, we could expect that there is a vertex algebra $(\mathcal{F}_\lambda, \mathbf{1}, Y)$ and a vector $v_\alpha \in \mathcal{F}_\lambda$ such that

$$Y(v_\alpha, x) = \alpha(x).$$

If we admit that we should take $\mathbf{1} = |\lambda\rangle$, there is an immediate consequence of the observation

$$Y(v_\alpha, x)\mathbf{1} \in \lambda\,|\lambda\rangle\, x^{-1} + \alpha_{-1}\,|\lambda\rangle + \mathcal{F}_\lambda[[x]]x.$$

As **(VA2)** must hold, the only possibility of λ is $\lambda = 0$, and the vector v_α is also determined as $v_\alpha = \alpha_{-1}\,|0\rangle$.

In the sequel of this section, we are going to seek the structure $(\mathcal{F}_0, |0\rangle, Y)$ of a vertex allgebra. An immediate problem is that we do not know how to define $Y(-,x)$ on the whole space $\mathcal{F}_0$.

The strategy is as follows.

(1) The Fock space $\mathcal{F}_0$ has the PBW basis

$$\alpha_{-n_1}\cdots\alpha_{-n_l}\,|0\rangle\,, \quad n_1 \geq \cdots \geq n_l > 0,\ l \in \mathbb{N}.$$

For each $\boldsymbol{n} = (n_1 \geq \cdots \geq n_l > 0)$, we will find $A^{\boldsymbol{n}}(x) \in \mathrm{End}(\mathcal{F}_0)[[x^{\pm 1}]]$ such that

$$A^{\boldsymbol{n}}(x) \left|0\right\rangle \in \alpha_{-n_1} \cdots \alpha_{-n_l} \left|0\right\rangle + \mathcal{F}_0[[x]]x$$

and define $Y(-, x)$ by

$$\begin{aligned} Y(\alpha_{-n_1} \cdots \alpha_{-n_k} \left|0\right\rangle, x) &:= A^{\boldsymbol{n}}(x), \quad n_1 \geq \cdots \geq n_l > 0,\, l \in \mathbb{N}, \\ Y(\left|0\right\rangle, x) &:= \mathrm{Id}_{\mathcal{F}_0} \cdot x^0 \end{aligned}$$

(2) Show that the triple $(\mathcal{F}_0, \left|0\right\rangle, Y)$ satisfies **(VA1)–(VA4)**. (**(VA2)** is clear.)

4.2 Normally ordered product

We decompose the formal series $\alpha(x)$ into two parts:

$$\alpha(x) = \alpha(x)_+ + \alpha(x)_-,$$

where

$$\alpha(x)_+ = \sum_{n \leq -1} \alpha_n x^{-n-1}, \quad \alpha(x)_- = \sum_{n \geq 0} \alpha_n x^{-n-1}.$$

In other words, $\alpha(x)_+$ and $\alpha(x)_-$ are the nonnegative and negative power parts of $\alpha(x)$, respectively. We also define

$$(\partial^{(k)}\alpha(x))_\pm := \partial^{(k)}\alpha(x)_\pm$$

for $k \in \mathbb{N}$.

It will be convenient to consider $\partial^{(k)}\alpha$ just a symbol. If $A \in \{\partial^{(k)}\alpha\}_{k \in \mathbb{N}}$ is one of those symbols, $A(x)$ can be naturally understood as the corresponding formal series.

Definition 4.3 (Normally ordered product (resolved))**.** For $A^1, \ldots, A^P \in \{\partial^{(k)}\alpha\}_{k \in \mathbb{N}}$, the formal series

$$\mathop{\circ}\limits^{\circ} A^1(x_1) \cdots A^P(x_P) \mathop{\circ}\limits^{\circ} \in \mathrm{End}(\mathcal{F}_0)[[x_1^{\pm 1}, \ldots, x_P^{\pm 1}]]$$

is defined recursively by

$$\mathop{\circ}\limits^{\circ} A^P(x_P) \mathop{\circ}\limits^{\circ} := A^P(x_P)$$

and

$$\begin{aligned} &\mathop{\circ}\limits^{\circ} A^1(x_1) A^2(x_2) \cdots A^P(x_P) \mathop{\circ}\limits^{\circ} \\ :=\, &A^1(x_1)_+ \mathop{\circ}\limits^{\circ} A^2(x_2) \cdots A^P(x_P) \mathop{\circ}\limits^{\circ} + \mathop{\circ}\limits^{\circ} A^2(x_2) \cdots A^P(x_P) \mathop{\circ}\limits^{\circ} A^1(x_1)_-. \end{aligned}$$

Proposition 4.4. *Let $A^1, \dots, A^P \in \{\partial^{(k)}\alpha\}_{k\in\mathbb{N}}$. Then,*

$$\,^\circ_\circ A^1(x_1)\cdots A^P(x_P)\,^\circ_\circ v \in \mathcal{F}_0((x_1, \dots, x_P))$$

for all $v \in \mathcal{F}_0$.

Proof. As $\partial^{(k)}\alpha(x)v \in \mathcal{F}_0((x))$ for all $v \in \mathcal{F}_0$, it suffices to show the following: for $A \in \{\partial^{(k)}\alpha\}_{k\in\mathbb{N}}$ and $B(x_2, \dots, x_P) \in \mathrm{End}(\mathcal{F}_0)[[x_2^{\pm 1}, \dots, x_P^{\pm 1}]]$ such that $B(x_2, \dots, x_P)v \in \mathcal{F}_0((x_2, \dots, x_P))$ for all $v \in \mathcal{F}_0$,

$$\,^\circ_\circ A(x_1)B(x_2, \dots, x_P)\,^\circ_\circ := A(x_1)_+ B(x_2, \dots, x_P) + B(x_2, \dots, x_P)A(x_1)_-$$

satisfies

$$\,^\circ_\circ A(x_1)B(x_2, \dots, x_P)\,^\circ_\circ v \in \mathcal{F}_0((x_1, x_2, \dots, x_P))$$

for all $v \in \mathcal{F}_0$.

Let us expand $A(x)$ and $B(x_2, \dots, x_P)$ into

$$\begin{aligned} A(x) &= \sum_{n\in\mathbb{Z}} A_n x^{-n-1}, \\ B(x_2, \dots, x_P) &= \sum_{n_2,\dots,n_P\in\mathbb{Z}} B_{n_2,\dots,n_P}\, x_2^{-n_2-1}\cdots x_P^{-n_P-1}. \end{aligned}$$

For each $v \in \mathcal{F}_0$, we may explicitly get

$$\begin{aligned} &\,^\circ_\circ A(x_1)B(x_2, \dots, x_P)\,^\circ_\circ v \\ &= \sum_{n_1\le -1}\sum_{n_2,\dots,n_P\in\mathbb{Z}} A_{n_1}B_{n_2,\dots,n_P} v\, x_1^{-n_1-1}x_2^{-n_2-1}\cdots x_P^{-n_P-1} \\ &\quad + \sum_{n_1\ge 0}\sum_{n_2,\dots,n_P\in\mathbb{Z}} B_{n_2,\dots,n_P}A_{n_1} v\, x_1^{-n_1-1}x_2^{-n_2-1}\cdots x_P^{-n_P-1}. \end{aligned}$$

The first sum already lies in $\mathcal{F}_0((x_1, \dots, x_P))$, so we can concentrate on the second sum. There exists $K'(v)$ depending on v such that $A_{n_1}v = 0$ if $n_1 > K'(v)$. Hence, the sum over n_1 is truncated to $K'(v)$. For each $n_1 = 0, \dots, K'(v)$, we can choose $K''(n_1, v)$ such that $B_{n_2,\dots,n_P}A_{n_1}v = 0$ if one of $n_2, \dots, n_P$ is greater than $K''(n_1, v)$. If we set K to be the largest among $K'(v), K''(0, v), \dots, K''(K'(v), v)$, we can see that the second sum lies in

$$\mathcal{F}_0[[x_1, \dots, x_N]](x_1\cdots x_N)^{-K-1} \subset \mathcal{F}_0((x_1, \dots, x_N)).$$

Therefore, the desired result is proved. □

For a formal series $f(x_1, \dots, x_N) \in \mathcal{F}_0((x_1, \dots, x_N))$, it makes sense to specialize all the variable as $x_i = x$, $i = 1, \dots, N$. This justfies the following.

Corollary 4.5. *Let $A^1, \ldots, A^P \in \{\partial^{(k)}\alpha\}_{k\in\mathbb{N}}$. In their normally ordered product, if we specialize $x_1, \ldots, x_P$ into a single variable x, the result*

$$\circ\!\!\!\circ\, A^1(x) \cdots A^P(x) \,\circ\!\!\!\circ$$

lies in $\mathrm{End}(\mathcal{F}_0)[[x^{\pm 1}]]$ *and*

$$\circ\!\!\!\circ\, A^1(x) \cdots A^P(x) \,\circ\!\!\!\circ\, v \in \mathcal{F}_0((x))$$

for all $v \in \mathcal{F}_0$.

The formal series $\circ\!\!\!\circ\, A^1(x) \cdots A^P(x) \,\circ\!\!\!\circ$ in Corollary 4.5 is also (even more often) called the normally ordered product of $A^1(x), \ldots, A^P(x)$.

The following proposition accomplishes the first step of our strategy.

Proposition 4.6. *For $n_1 \geq \cdots \geq n_l > 0$,*

$$\circ\!\!\!\circ\, \partial^{(n_1-1)}\alpha(x) \cdots \partial^{(n_l-1)}\alpha(x) \,\circ\!\!\!\circ\, |0\rangle \in \alpha_{-n_1} \cdots \alpha_{-n_l} |0\rangle + \mathcal{F}_0[[x]]x.$$

Proof. Recall that

$$\partial^{(n-1)}\alpha(x) = \sum_{m\in\mathbb{Z}} \binom{-m-1}{n-1} \alpha_m x^{-m-j}.$$

We can show (Exercise 2.2.1) that the normally order product can be expanded into

$$\begin{aligned} &\circ\!\!\!\circ\, \partial^{(n_1-1)}\alpha(x) \cdots \partial^{(n_l-1)}\alpha(x) \,\circ\!\!\!\circ \\ = &\sum_{m_1,\ldots,m_l\in\mathbb{Z}} \binom{-m_1-1}{n_1-1} \cdots \binom{-m_l-1}{n_l-1} \circ\!\!\!\circ\, \alpha_{m_1} \cdots \alpha_{m_l} \,\circ\!\!\!\circ\, x^{-m_1-n_1} \cdots x^{-m_l-n_l}. \end{aligned}$$

The coefficient operators are defined by

$$\circ\!\!\!\circ\, \alpha_{m_1} \cdots \alpha_{m_l} \,\circ\!\!\!\circ = \alpha_{m_{\sigma(1)}} \cdots \alpha_{m_{\sigma(l)}},$$

where σ is a permutation of $\{1, \ldots, l\}$, such that

$$m_{\sigma(1)}, \ldots, m_{\sigma(i)} < 0 \leq m_{\sigma(i+1)}, \ldots, m_{\sigma(l)}$$

with some $i \in \{1, \ldots, l\}$. When acted on the vacuum $|0\rangle$, $\circ\!\!\!\circ\, \alpha_{m_1} \cdots \alpha_{m_l} \,\circ\!\!\!\circ$ annihilates it if one of $m_1, \ldots, m_l$ is non-negative. If m_i is negative, we can observe that $\binom{-m_i-1}{n_i-1} = 0$ unless $m_i \leq -n_i$. Therefore, we get

$$\begin{aligned} &\circ\!\!\!\circ\, \partial^{(n_1-1)}\alpha(x) \cdots \partial^{(n_l-1)}\alpha(x) \,\circ\!\!\!\circ\, |0\rangle \\ = &\sum_{m_1\leq -n_1} \cdots \sum_{m_l\leq -n_l} \binom{-m_1-1}{n_1-1} \cdots \binom{-m_l-1}{n_l-1} \\ &\qquad \cdot \alpha_{m_1} \cdots \alpha_{m_l} |0\rangle \, x^{-m_1-n_1} \cdots x^{-m_l-n_l}. \end{aligned} \tag{4.1}$$

This series has only non-negative powers of x and the constant term is precisely

$$\alpha_{-n_1} \cdots \alpha_{-n_l} |0\rangle ,$$

which was the desired result. $\square$

For later use, let us see the coefficient of x in the series (4.1). Such terms occur when only one of $m_1, \ldots, m_l$, say m_i is $-n_i - 1$, and the others m_j, $j \neq i$ all stay $-n_j$. As the binomial coefficient is $\binom{n_i}{n_i-1} = n_i$, the coefficient of x is

$$\begin{aligned} &[x] \left({}^{\circ}_{\circ} \partial^{(n_1-1)}\alpha(x) \cdots \partial^{(n_l-1)}\alpha(x) {}^{\circ}_{\circ} |0\rangle \right) \\ &= \sum_{i=1}^{l} n_i \, \alpha_{-n_1} \cdots \alpha_{-n_{i-1}} \alpha_{-n_i-1} \alpha_{-n_{i+1}} \cdots \alpha_{-n_l} |0\rangle . \end{aligned} \tag{4.2}$$

Motivated by Proposition 4.6, we define the linear map

$$Y(-,x) \colon \mathcal{F}_0 \to \mathrm{End}(\mathcal{F}_0)[[x^{\pm 1}]]$$

by

$$Y(\alpha_{-n_1} \cdots \alpha_{-n_l} |0\rangle , x) := {}^{\circ}_{\circ} \partial^{(n_1-1)}\alpha(x) \cdots \partial^{(n_l-1)}\alpha(x) {}^{\circ}_{\circ}$$

for each $n_1 \geq \cdots \geq n_l$, $l \in \mathbb{N}$ and $Y(|0\rangle , x) = \mathrm{Id}_{\mathcal{F}_0}$. From its construction, **(VA2)** is clearly satisfied. From Corollary 4.5, **(VA1)** is satisfied too. The rest is to show **(VA3)** and **(VA4)**.

4.3 Translation axiom (VA3)

Recall the definition of the operator T: when it is acted on $a \in V$, the result is $T(a) = a_{(-2)}\mathbf{1} = [x]\,(Y(a,x)\mathbf{1})$. In our case, the coefficient of x in $Y(a,x)\mathbf{1}$ has been studied in (4.2) for basis vectors of $\mathcal{F}_0$. Therefore, the operator T is given by

$$T\left(\alpha_{-n_1} \cdots \alpha_{-n_l} |0\rangle\right) = \sum_{i=1}^{l} n_i \, \alpha_{-n_1} \cdots \alpha_{-n_{i-1}} \alpha_{-n_i-1} \alpha_{-n_{i+1}} \cdots \alpha_{-n_l} |0\rangle$$

for $n_1 \geq \cdots \geq n_l > 0$, $l \in \mathbb{N}$.

Lemma 4.7. *We have the commutation relations $[T, \alpha_n] = -n\alpha_{n-1}$, $n \in \mathbb{Z}$ of operators on $\mathcal{F}_0$.*

Proof. As an explicit definition of T is given, it is a straightforward (but possibly tedious) computation (Exercise 2.3.1). $\square$

Proposition 4.8. *For any $a \in \mathcal{F}_0$,*

$$[T, Y(a,x)] = \frac{d}{dx} Y(a,x).$$

Proof. Lemma 4.7 justifies

$$[T, \alpha(x)] = \frac{d}{dx}\alpha(x).$$

We can decompose this equation into the non-negative and negative power parts and differentiate both sides by x arbitrarily many times. Therefore, we get

$$[T, \partial^{(k)}\alpha(x)_{\pm}] = \frac{d}{dx}\partial^{(k)}\alpha(x)_{\pm}.$$

Now the result follows from the definition of $Y(-, x)$ on the PBW basis and the normally ordered product. $\square$

4.4 Wick's theorem

In order to prove the locality **(VA4)**, we need to prepare Wick's theorem. Here, we are concerned with the product of operator-valued formal series

$${}^{\circ}_{\circ} A^1(x) \cdots A^P(x) {}^{\circ}_{\circ} {}^{\circ}_{\circ} B^1(y) \cdots B^Q(y) {}^{\circ}_{\circ} \in \mathrm{End}(\mathcal{F}_0)[[x^{\pm 1}, y^{\pm 1}]],$$

where $A^1, \ldots, A^P, B^1, \ldots, B^Q \in \{\partial^{(k)}\alpha\}_{k \in \mathbb{N}}$.

Under those fixed P and Q, we introduce the set of vertices

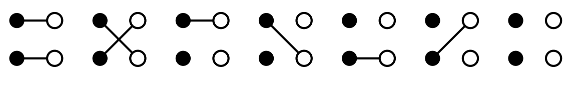
$$\mathsf{V}_{(P,Q)} = \mathsf{V}^A_{(P)} \cup \mathsf{V}^B_{(Q)},$$

where

$$\mathsf{V}^A_{(P)} = \left\{ \overset{1}{\bullet} \ \overset{2}{\bullet} \cdots \overset{P}{\bullet} \right\}, \quad \mathsf{V}^B_{(Q)} = \left\{ \overset{1}{\circ} \ \overset{2}{\circ} \cdots \overset{Q}{\circ} \right\}$$

are the sets of P vertices of type A and Q vertices of type B, respectively. A Feynman graph over $\mathsf{V}_{(P,Q)}$ is a bipartite graph with the vertex set $\mathsf{V}_{(P,Q)}$ with respect to the prescribed types of vertices such that each vertex is connected to at most one edge. For a Feynman graph Γ over $\mathsf{V}_{(P,Q)}$, each edge is denoted by (i, j) with $i \in \mathsf{V}^A_{(P)}$ and $j \in \mathsf{V}^B_{(Q)}$. We also write $\Phi^A(\Gamma)$ for the set of vertices of type A that are not connected to an edge in Γ, and similarly $\Phi^B(\Gamma)$. The set of Feynman graphs over $\mathsf{V}_{(P,Q)}$ will be $\mathsf{FG}_{(P,Q)}$.

Example 4.9. When $P = Q = 2$, $\mathsf{FG}_{(2,2)}$ consists of



Theorem 4.10 (Wick's theorem)**.** *For* $A^1, \ldots, A^P, B^1, \ldots, B^Q \in \{\partial^{(k)}\alpha\}_{k \in \mathbb{N}}$,

$$\begin{aligned} &{}^{\circ}_{\circ} A^1(x) \cdots A^P(x) {}^{\circ}_{\circ} {}^{\circ}_{\circ} B^1(y) \cdots B^Q(y) {}^{\circ}_{\circ} \\ &= \sum_{\Gamma \in \mathsf{FG}_{(P,Q)}} \prod_{(i,j) \in \Gamma} [A^i(x)_-, B^j(y)_+] \, {}^{\circ}_{\circ} \prod_{i \in \Phi^A(\Gamma)} A^i(x) \prod_{j \in \Phi^B(\Gamma)} B^j(y) {}^{\circ}_{\circ}. \end{aligned}$$

Proof (sketch). The product of normally ordered products

$$\overset{\circ}{\circ} A^1(x) \cdots A^P(x) \overset{\circ}{\circ} \overset{\circ}{\circ} B^1(y) \cdots B^Q(y) \overset{\circ}{\circ}$$

is a sum of terms of the form

$$\begin{aligned} & A^{i_1}(x)_+ \cdots A^{i_s}(x)_+ \\ & \cdot A^{i_{s+1}}(x)_- \cdots A^{i_P}(x)_- B^{j_1}(y)_+ \cdots B^{j_t}(y)_+ \\ & \cdot B^{j_{t+1}}(y)_- \cdots B^{j_Q}(y)_-, \end{aligned}$$

where $\{i_1, \dots, i_P\} = \{1, \dots, P\}$ and $\{j_1, \dots, j_Q\} = \{1, \dots, Q\}$. Each term of the above form is not normally ordered because $A^{i_{s+1}}(x)_-, \dots, A^{i_P}(x)_-$ are located on the left of $B^{j_1}(y)_+, \dots, B^{j_t}(y)_+$. While rearranging the product of them into the normal order, we apply the commutation relation

$$A^i(x)_- B^j(y)_+ = B^j(y)_+ A^i(x)_- + [A^i(x)_-, B^j(y)_+]$$

to every possible pair, which is why we have a choice between taking the commutator $[A^i(x)_-, B^j(y)_+]$ and leaving them in the normally ordered product. By taking care of the combinatorics (which is left to Exercise 2.4.1) carefully, we get the result. □

Note that as an operator-valued formal series, all coefficients of the commutator

$$[A^i(x)_-, B^j(y)_+]$$

are the identity operator. More specifically, we can observe the following:

$$\begin{aligned} [\alpha(x)_-, \alpha(y)_+] &= \sum_{m \geq 0} \sum_{n \leq -1} [\alpha_m, \alpha_n] x^{-m-1} y^{-n-1} \\ &= \sum_{m=0}^{\infty} m x^{-m-1} y^{m-1} \\ &= \frac{\partial}{\partial y} \sum_{m=0}^{\infty} x^{-m-1} y^m \\ &= \frac{\partial}{\partial y} \frac{1}{x-y} \\ &= \frac{1}{(x-y)^2}. \end{aligned}$$

(Remember the convention of the binomial expansion in Definition 3.5.) By differentiating both sides, we get

$$
\begin{aligned}
[\partial^{(m-1)}\alpha(x)_-, \partial^{(n-1)}\alpha(y)_+] &= \partial_x^{(m-1)}\partial_y^{(n-1)}\frac{1}{(x-y)^2} \\
&= \frac{(-1)^{m-1}(m+n-2)!}{(m-1)!(n-1)!}\partial_y^{(m+n-2)}\frac{1}{(x-y)^2} \\
&= \frac{(-1)^{m-1}(m+n-2)!}{(m-1)!(n-1)!}\binom{-2}{m+n-2}\frac{1}{(x-y)^{m+n}} \\
&= (-1)^{n-1}\frac{(m+n-1)!}{(m-1)!(n-1)!}\frac{1}{(x-y)^{m+n}}.
\end{aligned}
$$

When we switch the role of (x, j) and (y, k),

$$
\begin{aligned}
[\partial^{(n-1)}\alpha(y)_-, \partial^{(m-1)}\alpha(x)_+] &= (-1)^{m-1}\frac{(m+n-1)!}{(m-1)!(n-1)!}\frac{1}{(y-x)^{m+n}} \\
&= (-1)^{n-1}\frac{(m+n-1)!}{(m-1)!(n-1)!}\frac{1}{(-y+x)^{m+n}}.
\end{aligned}
$$

We are now ready to prove the locality **(VA4)**.

Theorem 4.11. *For $A^1, \ldots, A^P, B^1, \ldots, B^Q \in \{\partial^{(k)}\alpha\}_{k\in\mathbb{N}}$,*

$$
(x-y)^N \left[{}^{\circ}_{\circ} A^1(x) \cdots A^P(x) {}^{\circ}_{\circ}, {}^{\circ}_{\circ} B^1(y) \cdots B^Q(y) {}^{\circ}_{\circ} \right] = 0, \quad N \gg 0.
$$

Proof. From Theorem 4.10, the commutator admits the following expression:

$$
\begin{aligned}
&\left[{}^{\circ}_{\circ} A^1(x) \cdots A^P(x) {}^{\circ}_{\circ}, {}^{\circ}_{\circ} B^1(y) \cdots B^Q(y) {}^{\circ}_{\circ} \right] \\
&= \sum_{\Gamma \in \mathsf{FG}_{(P,Q)}} \left(\prod_{(i,j)\in\Gamma} [A^i(x)_-, B^j(y)_+] - \prod_{(i,j)\in\Gamma} [B^j(y)_-, A^i(x)_+] \right) \\
&\qquad \times {}^{\circ}_{\circ} \prod_{i\in\Phi^A(\Gamma)} A^i(x) \prod_{j\in\Phi^B(\Gamma)} B^j(y) {}^{\circ}_{\circ}.
\end{aligned}
$$

For each $\Gamma \in \mathsf{FG}_{(P,Q)}$, if we set

$$
C_\Gamma(x,y) = \prod_{(i,j)\in\Gamma} [A^i(x)_-, B^j(y)_+] - \prod_{(i,j)\in\Gamma} [B^j(y)_-, A^i(x)_+],
$$

there exists a constant c_Γ and $N_\Gamma \in \mathbb{N}$ such that

$$
\begin{aligned}
C_\Gamma(x,y) &= c_\Gamma \left(\frac{1}{(x-y)^{N_\Gamma+1}} - \frac{1}{(-y+x)^{N_\Gamma+1}} \right) \\
&= c_\Gamma \partial_y^{(N_\Gamma)}(x\delta(y/x)).
\end{aligned}
$$

Essentially because of Proposition 4.4, the triple product

$$
(x-y)^{N_\Gamma+1} C_\Gamma(x,y) {}^{\circ}_{\circ} \prod_{i\in\Phi^A(\Gamma)} A^i(x) \prod_{j\in\Phi^B(\Gamma)} B^j(y) {}^{\circ}_{\circ}
$$

exists in $\mathrm{End}(\mathcal{F}_0)[[x^{\pm1}, y^{\pm1}]]$ and is 0. If we take N as the largest among N_Γ, $\Gamma \in \mathsf{FG}_{(P,Q)}$, the desired result follows. □

4.5 Conclusion

In this section, we have constructed the first non-trivial vertex algebra:

Theorem 4.12. *The triple* $(\mathcal{F}_0, |0\rangle, Y)$ *of*

- $\mathcal{F}_0$*: the Fock representation of* $\widehat{\mathfrak{h}}$ *corresponding to* $\lambda = 0$*,*
- $|0\rangle \in \mathcal{F}_0$*, and*
- *the linear map* $Y(-, x)\colon \mathcal{F}_0 \to \mathrm{End}(\mathcal{F}_0)[[x^{\pm 1}]]$ *defined by*

$$Y(\alpha_{-n_1} \cdots \alpha_{-n_l} |0\rangle, x) := {}^{\circ}_{\circ}\, \partial^{(n_1-1)}\alpha(x) \cdots \partial^{(n_l-1)}\alpha(x) \,{}^{\circ}_{\circ},$$

  for $n_1 \geq \cdots \geq n_l > 0$*,* $l \in \mathbb{N}$ *and* $Y(|0\rangle, x) := \mathrm{Id}_{\mathcal{F}_0} \cdot x^0$

forms a vertex algebra.

This vertex algebra is called the Heisenberg vertex algebra.

5 Construction of vertex algebras

In the previous section 4, we constructed the first non-trivial example of vertex algebra, the Heisenberg vertex algebra. There were two highlights in the construction:

(1) Normally ordered product,

(2) Wick's theorem.

The first one, normally ordered product, was used to define the linear map $Y(-, x)$, and it worked (surprisingly) well. However, we could raise the following question.

Question. How did we come up with the normally ordered product?

The second one, Wick's theorem, was crucial in proving the locality, but that relied on the fact that the Heisenberg algebra is *almost commutative*: α_m and α_n commute when m and n are of the same sign. We cannot expect that a similar reasoning works in the case where the underlying algebra is more *noncommutative* (e.g. Virasoro algebra). Hence, here is another question:

Question. How do we prove the locality in a more general setting?

5.1 Jacobi identity

We first address the first question. It will turn out that the normally ordered product is simply hidden in the set of axioms of vertex algebra.

There are, in fact, several different sets of axioms that define a vertex algebra. The axioms **(VA1)**–**(VA4)** are easiest to check, but for the present purpose, we would like to replace the axioms **(VA3)** and **(VA4)** with the so-called "Jacobi identity", from which we know how we should define $Y(-,x)$.

From now on, we assume that $(V, \mathbf{1}, Y)$ is a vertex algebra defined in terms of the axioms **(VA1)**–**(VA4)**. Let us see various consequences of the axioms.

5.1.1 Commutativity

In order to state the commutativity, we need a new type of space

$$\mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-1}, (x_1 - x_2)^{-1}].$$

This could be literally read as the space of polynomials in x_1^{-1}, x_2^{-1} and $(x_1 - x_2)^{-1}$ with coefficients in $\mathbb{C}[[x_1, x_2]]$. However, this is **not** the way we understand it. Instead, we think of it as the localization the ring $\mathbb{C}[[x_1, x_2]]$ by the multiplicatively closed set

$$\{x_1^k x_2^l (x_1 - x_2)^m\}_{k,l,m \geq 0}.$$

Every element of the localized ring is (non-uniquely) expressed as

$$\frac{g(x_1, x_2)}{x_1^k x_2^l (x_1 - x_2)^m}$$

by means of $g(x_1, x_2) \in \mathbb{C}[[x_1, x_2]]$ and $k, l, m \geq 0$. The expansion maps

$$\begin{aligned} \iota_{12} &\colon \mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-1}, (x_1 - x_2)^{-1}] \to \mathbb{C}((x_1))((x_2)), \\ \iota_{21} &\colon \mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-1}, (x_1 - x_2)^{-1}] \to \mathbb{C}((x_2))((x_1)) \end{aligned}$$

are homomorphisms of rings that are defined by

$$\begin{aligned} \iota_{12} &\colon \frac{g(x_1, x_2)}{x_1^k x_2^l (x_1 - x_2)^m} \mapsto g(x_1, x_2) x_1^{-k} x_2^{-l} \sum_{k=0}^{\infty} \binom{-m}{k} x_1^{-m-k} (-x_2)^k, \\ \iota_{21} &\colon \frac{g(x_1, x_2)}{x_1^k x_2^l (x_1 - x_2)^m} \mapsto g(x_1, x_2) x_1^{-k} x_2^{-l} \sum_{k=0}^{\infty} \binom{-m}{k} x_1^k (-x_2)^{-m-k}. \end{aligned}$$

One can check (Exercise 3.1.1) that ι_{12} and ι_{21} are well-defined and injective.

Geometrically, ι_{12} expands rational functions in the region $|x_1| > |x_2|$ and ι_{21} does the same in the opposite region $|x_2| > |x_1|$. The indices for ι refer to the variables under consideration. In the sequel, we will encounter more general ι_{ij}, but it should not be too confusing what it means; it expands rational functions in $|x_i| > |x_j|$.

Proposition 5.1 (Commutativity)**.** *For any $a, b, c \in V$ and $\varphi \in V^*$, there exists a unique element*

$$f(x_1, x_2) \in \mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-2}, (x_1 - x_2)^{-1}]$$

such that

$$\langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle = \iota_{12} f(x_1, x_2),$$
$$\langle \varphi, Y(b, x_2)Y(a, x_1)c \rangle = \iota_{21} f(x_1, x_2).$$

Proof. Due to **(VA4)**, we can choose $N \gg 0$ so that

$$(x_1 - x_2)^N \langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle = (x_1 - x_2)^N \langle \varphi, Y(b, x_2)Y(a, x_1)c \rangle .$$

Due to **(VA1)**, the left-hand side has finitely many negative powers of x_2 and the right-hand side contains finitely many negative powers of x_1. Therefore, there exists $h(x_1, x_2) \in \mathbb{C}((x_1, x_2))$ that coincides with both sides. Let us take a close look at the identity:

$$h(x_1, x_2) = (x_1 - x_2)^N \langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle .$$

We can notice that

$$\langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle \in \mathbb{C}((x_1))((x_2)).$$

In this case, the triple product

$$(x_1 - x_2)^{-N} \cdot (x_1 - x_2)^N \cdot \langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle$$

exists in $\mathbb{C}[[x_1^{\pm 1}, x_2^{\pm 1}]]$. Hence, we get

$$\langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle = (x_1 - x_2)^{-N} h(x_1, x_2).$$

Similarly, we have

$$\langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle = (-x_2 + x_1)^{-N} h(x_1, x_2).$$

Therefore, we can take

$$f(x_1, x_2) = \frac{h(x_1, x_2)}{(x_1 - x_2)^N} \in \mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-1}, (x_1 - x_2)^{-1}]$$

with the desired property. $\square$

5.1.2 Few translation properties

Here, let us observe some properties of the translation operator T.

- The vacuum vector $\mathbf{1}$ is "translation invariant":
$$T\mathbf{1} = \mathbf{1}_{(-2)}\mathbf{1} = 0.$$
- From **(VA3)** and the translation invariance of $\mathbf{1}$, for any $a \in V$, we have (Exercise 2.6.1)
$$Y(a, x)\mathbf{1} = e^{xT}a. \tag{5.1}$$
- By Taylor's formula and **(VA3)**, for any $a \in V$, we have (Exercise 2.6.2)
$$Y(a, x + y) = e^{yT}Y(a, x)e^{-yT}. \tag{5.2}$$

5.1.3 Skew-symmetry

Proposition 5.2. *For $a, b \in V$, we have the identity*
$$Y(a, x)b = e^{xT}Y(b, -x)a$$
in $\mathrm{End}(V)[[x^{\pm 1}]]$.

Proof. We start from the locality **(VA4)**: there is $N \gg 0$ such that
$$(x - y)^N Y(a, x)Y(b, y)\mathbf{1} = (x - y)^N Y(b, y)Y(a, x)\mathbf{1}. \tag{5.3}$$
By (5.1) and (5.2), we can continue the right-hand side as
$$\begin{aligned}(x - y)^N Y(b, y)Y(a, x)\mathbf{1} &= (x - y)^N Y(b, y)e^{xT}a \\ &= (x - y)^N e^{xT}Y(b, y - x)a. \end{aligned} \tag{5.4}$$
From the field condition **(VA1)**, there is $M > 0$ such that the right-hand side truncates as
$$(x - y)^N e^{xT}Y(b, y - x)a = (-1)^N e^{xT} \sum_{n=-\infty}^{M} b_{(n)}a\,(y - x)^{N-n-1}.$$
If we take $N \geq M+1$, the right-hand side only contains non-negative powers of y, so we can take $y = 0$, which results in
$$\begin{aligned}x^N Y(a, x)b &= (-1)^N e^{xT} \sum_{n=-\infty}^{M} b_{(n)}a\,(-x)^{N-n-1} \\ &= x^N e^{xT}Y(b, -x)a.\end{aligned}$$
By multiplying x^{-N} on both sides, we get the desired result. □

Remark 5.3. It is clear that the left-hand side of (5.3) only contains non-negative powers of y, so we already know at this point that we can plug in $y = 0$ on both sides. Thus, one might be tempted to do the substitution on (5.4) and conclude the result. This does not necessarily work essentially because of the following reason: if we set $f(x, y) = (y - x)^{-1}$, we have the equality

$$1 = (y - x)f(x, y)$$

in $\mathbb{C}[[x^{\pm 1}, y^{\pm 1}]]$. In particular, both sides only contain non-negative powers of y. However, as $f(x, y)$ is precisely defined by

$$f(x, y) = 1 + \frac{x}{y} + \left(\frac{x}{y}\right)^2 + \cdots ,$$

the substitution $f(x, 0)$ does not make sense.

5.1.4 (Weak) associativity

Proposition 5.4. *For any $a, c \in V$, we can choose $N \gg 0$ such that the identity*

$$(x_0 + x_2)^N Y(a, x_0 + x_2)Y(b, x_2)c = (x_0 + x_2)^N Y(Y(a, x_0)b, x_2)c$$

in $\mathrm{End}(V)[[x_0^{\pm 1}, x_2^{\pm 1}]]$ *holds for all $b \in V$.*

Before proving the proposition, we need to explain what the right-hand side means; the map $Y(-, x)$ is supposed to be a linear map from V to $\mathrm{End}(V)[[x^{\pm 1}]]$, but $Y(a, x_0)b$ lives in $V((x_0))$, but not in V. We interpret the right-hand side by considering that $Y(-, x)$ is also linear with respect to the multiplication by x_0. Concretely, we expand $Y(a, x_0)b$ into

$$Y(a, x_0)b = \sum_{n \in \mathbb{Z}} a_{(n)}b\, x_0^{-n-1}$$

and understand

$$Y(Y(a, x_0)b, x_2) = \sum_{n \in \mathbb{Z}} Y(a_{(n)}b, x_2)x_0^{-n-1}.$$

Proof. We start with the left-hand side. We apply the skew-symmetry Proposition 5.2, the translation rule (5.2) and the locality **(VA4)** to get

$$\begin{aligned}
&(x_0 + x_2)^N Y(a, x_0 + x_2)Y(b, x_2)c \\
&= (x_0 + x_2)^N Y(a, x_0 + x_2)e^{x_2 T} Y(c, -x_2)b \\
&= (x_0 + x_2)^N e^{x_2 T} Y(a, x_0)Y(c, -x_2)b \\
&= (x_0 + x_2)^N e^{x_2 T} Y(c, -x_2)Y(a, x_0)b.
\end{aligned}$$

Notice that we can apply the skew-symmetry Proposition 5.2 again to the right-hand side by considering $Y(a, x_0)b$ to be a single vector to get

$$(x_0 + x_2)^N Y(Y(a, x_0)b, x_2)c,$$

which is the desired right-hand side. $\square$

The space

$$\mathbb{C}[[x_0, x_2]][x_0^{-1}, x_2^{-1}, (x_0 + x_2)^{-1}]$$

in the following proposition is understood in a similar way as before in Proposition 5.1. This is the localization of the ring $\mathbb{C}[[x_0, x_2]]$ by the multiplicatively closet set $\{x_0^k x_2^l (x_0 + x_2)^m\}_{k,l,m \geq 0}$.

Proposition 5.5 (Associativity)**.** *For any $a, b, c \in V$ and $\varphi \in V^*$, there exists a unique element*

$$q(x_0, x_2) \in \mathbb{C}[[x_0, x_2]][x_0^{-1}, x_2^{-1}, (x_0 + x_2)^{-1}]$$

such that

$$\begin{aligned} \langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle &= \iota_{02} q(x_0, x_2), \\ \langle \varphi, Y(Y(a, x_0)b, x_2)c \rangle &= \iota_{20} q(x_0, x_2). \end{aligned}$$

Proof. Proposition 5.5 is to Proposition 5.4 as Proposition 5.1 is to the locality **(VA4)**. From Proposition 5.4, we can choose $N \gg 0$ so that

$$(x_0 + x_2)^N \langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle = (x_0 + x_2)^N \langle \varphi, Y(Y(a, x_0)b, x_2)c \rangle .$$

The left-hand side has finitely many negative powers of x_2 while the right-hand side has finitely many negative powers of x_0. Therefore, both sides are equal to a certain formal series of the form

$$r(x_0, x_2) \in \mathbb{C}((x_0, x_2)).$$

Now, as

$$\langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle \in \mathbb{C}((x_0))((x_2)),$$

the triple product

$$(x_0 + x_2)^{-N} \cdot (x_0 + x_2)^N \cdot \langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle$$

exists in $\mathbb{C}[[x_0^{\pm 1}, x_2^{\pm 1}]]$. Hence, we are allowed to cancel $(x_0 + x_2)^{-N}$ and $(x_0 + x_2)^N$ to get

$$\langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle = (x_0 + x_2)^{-N} r(x_0, x_2).$$

Similarly, we get

$$\langle \varphi, Y(Y(a, x_0)b, x_2)c \rangle = (x_2 + x_0)^{-N} r(x_0, x_2).$$

Therefore, we can conclude the desired result. $\square$

5.1.5 Combining commutativity and associativity

Let us fix $a, b, c \in V$ and $\varphi \in V^*$. From Proposition 5.1, we have a unique element

$$f(x_1, x_2) \in \mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-1}, (x_1 - x_2)^{-1}]$$

such that

$$\langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle = \iota_{12} f(x_1, x_2).$$

At the same time, from Proposition 5.5, there is a unique element

$$q(x_0, x_2) \in \mathbb{C}[[x_0, x_2]][x_0^{-1}, x_2^{-1}, (x_0 + x_2)^{-1}]$$

such that

$$\langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle = \iota_{02} q(x_0, x_2).$$

We can show (Exercise 3.1.2) that

$$(\iota_{02} q(x_0, x_2))|_{x_0 = x_1 - x_2} = \iota_{12} q(x_1 - x_2, x_2),$$

i.e., expansion and substitution commute as long as we choose the expansion region correctly. It is also true that

$$\langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle = \langle \varphi, Y(a, x_0 + x_2)Y(b, x_2)c \rangle |_{x_0 = x_1 - x_2}.$$

Therefore, we can conclude that

$$f(x_1, x_2) = q(x_1 - x_2, x_2) \in \mathbb{C}[[x_1, x_2]][x_1^{-1}, x_2^{-1}, (x_1 - x_2)^{-1}],$$

or equivalently,

$$f(x_0 + x_2, x_2) = q(x_0, x_2) \in \mathbb{C}[[x_0, x_2]][x_0^{-1}, x_2^{-1}, (x_0 + x_2)^{-1}].$$

Suppose that $f(x_1, x_2)$ is (non-uniquely) expressed as

$$f(x_1, x_2) = \frac{g(x_1, x_2)}{x_1^k x_2^l (x_1 - x_2)^m}, \quad \text{with} \quad g(x_1, x_2) \in \mathbb{C}[[x_1, x_2]].$$

Then, we introduce

$$F(x_0, x_1, x_2) := \frac{g(x_1, x_2)}{x_0^m x_1^k x_2^l} \in \mathbb{C}((x_0, x_1, x_2)).$$

By construction, $F(x_0, x_1, x_2)$ satisfies

$$\begin{aligned} F(x_1 - x_2, x_1, x_2) &= \iota_{12} f(x_1, x_2) = \langle \varphi, Y(a, x_1)Y(b, x_2)c \rangle, \\ F(-x_2 + x_1, x_1, x_2) &= \iota_{21} f(x_1, x_2) = \langle \varphi, Y(b, x_2)Y(a, x_1)c \rangle, \\ F(x_0, x_2 + x_0, x_2) &= \iota_{20} q(x_0, x_2) = \langle \varphi, Y(Y(a, x_0)b, x_2)c \rangle, \end{aligned}$$

where the left-hand sides are understood by the convention of binomial expansion (Definition 3.5).

Recall the identity

$$x_0^{-1}\delta\left(\frac{x_1 - x_2}{x_0}\right) - x_0^{-1}\delta\left(\frac{x_2 - x_1}{-x_0}\right) = x_2^{-1}\delta\left(\frac{x_1 - x_0}{x_2}\right)$$

and multiply $F(x_0, x_1, x_2)$ on both sides. Combining the identity $x_2^{-1}\delta\left(\frac{x_1-x_0}{x_2}\right) = x_1^{-1}\delta\left(\frac{x_2+x_0}{x_1}\right)$, we obtain

$$\begin{aligned} & x_0^{-1}\delta\left(\frac{x_1 - x_2}{x_0}\right)\langle\varphi, Y(a, x_1)Y(b, x_2)c\rangle \\ & - x_0^{-1}\delta\left(\frac{x_2 - x_1}{-x_0}\right)\langle\varphi, Y(b, x_2)Y(a, x_1)c\rangle \\ & = x_2^{-1}\delta\left(\frac{x_1 - x_0}{x_2}\right)\langle\varphi, Y(Y(a, x_0)b, x_2)c\rangle . \end{aligned}$$

As $a, b, c \in V$ and $\varphi \in V^*$ are arbitrary, we conclude the following.

Proposition 5.6 (Jacobi identity)**.** *For any $a, b \in V$, the identity (Jacobi identity)*

$$\begin{aligned} & x_0^{-1}\delta\left(\frac{x_1 - x_2}{x_0}\right)Y(a, x_1)Y(b, x_2) - x_0^{-1}\delta\left(\frac{x_2 - x_1}{-x_0}\right)Y(b, x_2)Y(a, x_1) \\ & = x_2^{-1}\delta\left(\frac{x_1 - x_0}{x_2}\right)Y(Y(a, x_0)b, x_2) \end{aligned} \tag{5.5}$$

in $\mathrm{End}(V)[[x_0^{\pm1}, x_1^{\pm1}, x_2^{\pm1}]]$ *holds.*

Remark 5.7. We have derived the Jacobi identity from the axioms of vertex algebra. Conversely, under **(VA1)** and **(VA2)**, we can derive **(VA3)** and **(VA4)** from the Jacobi identity (Exercises 3.2). Therefore, we can replace **(VA3)** and **(VA4)** with the Jacobi identity in the definition of vertex algebra.

5.2 n-th product of fields

Let us fix $a, b \in V$ and take the coefficients of x_1^{-1} of the Jacobi identity (5.5). Again, using the identity $x_2^{-1}\delta\left(\frac{x_1-x_0}{x_2}\right) = x_1^{-1}\delta\left(\frac{x_2+x_0}{x_1}\right)$, we obtain

$$\begin{aligned} Y(Y(a, x_0)b, x_2) = [x_1^{-1}]\Bigg(& x_0^{-1}\delta\left(\frac{x_1 - x_2}{x_0}\right)Y(a, x_1)Y(b, x_2) \\ & - x_0^{-1}\delta\left(\frac{x_2 - x_1}{-x_0}\right)Y(b, x_2)Y(a, x_1)\Bigg). \end{aligned}$$

Furthermore, when we take the coefficients of x_0^{-n-1}, $n \in \mathbb{Z}$ we find

$$\begin{aligned} &Y(a_{(n)}b, x_2) \\ &= [x_1^{-1}]\left((x_1 - x_2)^n Y(a, x_1)Y(b, x_2) - (-x_2 + x_1)^n Y(b, x_2)Y(a, x_1)\right). \end{aligned}$$

This means that, if we know $Y(a, x)$ and $Y(b, x)$, we know how to compute $Y(a_{(n)}b, x)$ for $n \in \mathbb{Z}$.

Motivated by this, we make the following definitions. Firstly, for a vector space W, the space of fields on W is defined by

$$\mathcal{E}(W) := \{A(x) \in \mathrm{End}(W)[[x^{\pm 1}]] | A(x)w \in W((x)),\ w \in W\}.$$

Otherwise, $\mathcal{E}(W)$ can be defined by $\mathcal{E}(W) := \mathrm{Hom}\,(W, W((x)))$. For two fields $A(x), B(x) \in \mathcal{E}(W)$, we define their n-th product by

$$A(x)_{(n)}B(x) := [x_1^{-1}]\left((x_1 - x)^n A(x_1)B(x) - (-x + x_1)^n B(x)A(x_1)\right) \quad (5.6)$$

Proposition 5.8. *For $A(x), B(x) \in \mathcal{E}(W)$ and $n \in \mathbb{Z}$, $A(x)_{(n)}B(x) \in \mathcal{E}(W)$.*

Proof. Let us expand $A(x)$ as

$$A(x) = \sum_{n \in \mathbb{Z}} A_n x^{-n-1}.$$

Then, for $w \in W$, we get

$$A(x)_{(n)}B(x)w = \sum_{k=0}^{\infty} \binom{n}{k} (-x)^k A_{n-k} B(x)w - \sum_{k=0}^{\infty} \binom{n}{k} (-x)^{n-k} B(x) A_k w \quad (5.7)$$

from the formula (5.6). The first sum already gives an element of $W((x))$. As $A_k w = 0$, $k \gg 0$, the second sum truncates to a finite sum of elements in $W((x))$. □

In Section 4.2, we defined the normally ordered product of *currents* of the Heisenberg algebra. The definition actually works for fields on a vector space. For $A(x) = \sum_{n \in \mathbb{Z}} A_n x^{-n-1} \in \mathrm{End}(W)[[x^{\pm 1}]]$, we define

$$A(x)_+ = \sum_{n \geq -1} A_n x^{-n-1}, \quad A(x)_- = \sum_{n \geq 0} A_n x^{-n-1}.$$

Then, it **follows**[1] that

$$(\partial A(x))_\pm = \partial(A(x)_\pm).$$

For $A(x), B(x) \in \mathcal{E}(W)$, their normally ordered product is defined by

$$\circ\!\!\circ\, A(x)B(x) \,\circ\!\!\circ = A(x)_+ B(x) + B(x)A(x)_-.$$

[1]Previously for the Heisenberg algebra, we **defined** $(\partial^{(k)}\alpha(x))_\pm$ by $\partial^{(k)}\alpha(x)_\pm$ as we did not define taking $(\cdot)_\pm$ of general fields.

Proposition 5.9. *Let $A(x), B(x) \in \mathcal{E}(W)$ and $n > 0$. Then, we have*

$$A(x)_{(-n)}B(x) = {\circ\atop\circ}(\partial^{(n-1)}A(x))B(x){\circ\atop\circ}.$$

Proof. Notice the identities

$$\begin{aligned}\binom{-n}{k} &= (-1)^k \frac{(n+k-1)\cdots(n)}{k!} = (-1)^k \frac{(n+k-1)!}{k!(n-1)!} \\ &= (-1)^k \binom{n+k-1}{n-1} = (-1)^{n+k+1}\binom{-k-1}{n-1}.\end{aligned}$$

From the explicit formula (5.7), we get

$$\begin{aligned}&A(x)_{(-n)}B(x) \\ &= \sum_{k=0}^{\infty}\binom{n+k-1}{n-1}x^k A_{-n-k}B(x) + \sum_{k=0}^{\infty}\binom{-k-1}{n-1}x^{-n-k}B(x)A_k \\ &= \left(\partial^{(n-1)}\sum_{k=0}^{\infty}A_{-n-k}x^{n+k-1}\right)B(x) + B(x)\left(\partial^{(n-1)}\sum_{k=0}^{\infty}A_k x^{-k-1}\right).\end{aligned}$$

This coincides with ${\circ\atop\circ}(\partial^{(n-1)}A(x))B(x){\circ\atop\circ}$. □

This shows that the normally ordered product has been implemented in the axioms of vertex algebra: for $a, b \in V$, we must have

$$Y(a_{(-n)}b, x) = {\circ\atop\circ}(\partial^{(n-1)}Y(a,x))Y(b,x){\circ\atop\circ}, \quad n > 0.$$

The differential operator ∂ is a derivative in terms of the n-th product as we can see below.

Lemma 5.10. *For $A(x), B(x) \in \mathcal{E}(W)$ and $n \in \mathbb{Z}$,*

$$\partial(A(x)_{(n)}B(x)) = \partial A(x)_{(n)}B(x) + A(x)_{(n)}\partial B(x).$$

Proof. A crucial observation is that

$$[x_1^{-1}]\,(\partial_{x_1}f(x_1)) = 0$$

for any formal series $f(x_1)$ of x_1. This property allows us to perform an *integration by parts* in terms of x_1 under $[x_1^{-1}]$. Also, we should notice that

$$\partial_x(x_1 - x)^n = -\partial_{x_1}(x_1 - x)^n.$$

Then, the result follows from the definition of the n-th product. □

5.3 Dong's Lemma

The next problem is what is a substitute for Wick's theorem in a general setting. It will turn out that, while a combinatorial formula as explicit as Wick's theorem is not possible, the locality is preserved under the n-th products.

Definition 5.11. Let W be a vector space. The operator-valued formal series $A(x), B(x) \in \mathrm{End}(W)[[x^{\pm 1}]]$ are said to be mutually local if

$$(x-y)^N [A(x), B(y)] = 0, \quad N \gg 0.$$

Lemma 5.12 (Dong)**.** *Let $A(x), B(x), C(x) \in \mathcal{E}(W)$ be pairwise mutually local fields on W. Then, for any $n \in \mathbb{Z}$, $A(x)_{(n)}B(x)$ and $C(x)$ are mutually local.*

Proof. We can choose $M \geq -n$ so that

$$\begin{gathered}(x_1-x)^M A(x_1)B(x) = (x_1-x)^M B(x)A(x_1),\\ (x-y)^M B(x)C(y) = (x-y)^M C(y)B(x),\\ (x_1-y)^M A(x_1)C(y) = (x_1-y)^M C(y)A(x_1).\end{gathered}$$

Then, if we take $N = 4M$, we can see that

$$\begin{aligned}&(x-y)^N \left((x_1-x)^n A(x_1)B(x)C(y) - (-x+x_1)^n B(x)A(x_1)C(y)\right)\\ =&\sum_{k=0}^{3M} \binom{3M}{k} (x-x_1)^{3M-k}(x_1-y)^k (x-y)^M \cdot\\ &\qquad \left((x_1-x)^n A(x_1)B(x)C(y) - (-x+x_1)^n B(x)A(x_1)C(y)\right).\end{aligned}$$

If $k \leq M$, we have $3M-k+n \geq M+(M-k)+(M+n) \geq M$, in which case the corresponding summand is simply 0 from the assumption. Hence, we can restrict the summation to $k \geq M+1$. Being multiplied by $(x_1-y)^k(x-y)^M$, $C(y)$ can commute with $A(x_1)$ and $B(x)$. Therefore, we get

$$\begin{aligned}&(x-y)^N \left((x_1-x)^n A(x_1)B(x)C(y) - (-x+x_1)^n B(x)A(x_1)C(y)\right)\\ =&\sum_{k=M+1}^{3M} \binom{3M}{k} (x-x_1)^{3M-k}(x_1-y)^k (x-y)^M \cdot\\ &\qquad \left((x_1-x)^n C(y)A(x_1)B(x) - (-x+x_1)^n C(y)B(x)A(x_1)\right).\end{aligned}$$

At this point, we can extend the summation over k back to starting from 0, and obtain

$$\begin{aligned}&(x-y)^N \left((x_1-x)^n A(x_1)B(x)C(y) - (-x+x_1)^n B(x)A(x_1)C(y)\right)\\ =&(x-y)^N \left((x_1-x)^n C(y)A(x_1)B(x) - (-x+x_1)^n C(y)B(x)A(x_1)\right).\end{aligned}$$

Taking the coefficients of x_1^{-1} of both sides, we get the desired result. □

5.4 Goddard's uniqueness theorem

Theorem 5.13 (Goddard)**.** *Let $(V, \mathbf{1}, Y)$ be a vertex algebra. If a field $A(x) \in \mathcal{E}(V)$ on V satisfies*

$$A(x)\mathbf{1} = Y(a,x)\mathbf{1}$$

for some $a \in V$ and mutually local with all $Y(b,x)$, $b \in V$, we have $A(x) = Y(a,x)$.

Proof. Exercise 3.3.1. □

5.5 Reconstruction theorem

Now we are ready to state the reconstruction theorem; the most fundamental tool we can use to show the existence of a vertex algebra.

Let us first prepare the setting: V is a vector space and $a^i \in V$, $i \in S$ are vectors labelled by a countable set S. We also fix a vector $\mathbf{1} \in V$ and an operator $\widetilde{T} \in \mathrm{End}(V)$. Assume that there exists a family of fields

$$a^i(x) = \sum_{n \in \mathbb{Z}} a^i_{(n)} x^{-n-1} \in \mathcal{E}(V)$$

labelled by $i \in S$, satisfying the following conditions:

(1) $a^i(x)\mathbf{1} \in a^i + V[[x]]x$ for all $i \in S$.

(2) $\widetilde{T}\mathbf{1} = 0$ and $[\widetilde{T}, a^i(x)] = \partial a^i(x)$ for all $i \in S$.

(3) $a^i(x)$ and $a^j(x)$ are mutually local for all $i, j \in S$.

(4) The vector space V is generated by

$$a^{i_1}_{(n_1)} \cdots a^{i_k}_{(n_k)}\mathbf{1}, \quad i_1, \ldots, i_k \in S,\ n_1, \ldots, n_k \in \mathbb{Z},\ k \in \mathbb{N}. \tag{5.8}$$

Theorem 5.14. *Under the assumptions above, there exists a unique vertex algebra $(V, \mathbf{1}, Y)$ such that*

$$Y\left(a^{i_1}_{(n_1)} \cdots a^{i_k}_{(n_k)}\mathbf{1}, x\right) = a^{i_1}(x)_{(n_1)}\left(\cdots\left(a^{i_k}(x)_{(n_k)}\mathrm{Id}_V\right)\cdots\right) \tag{5.9}$$

for all $i_1, \ldots, i_k \in S$, $n_1, \ldots, n_k \in \mathbb{Z}$, $k \in \mathbb{N}$. Furthermore, the translation operator T coincides with $\widetilde{T}$.

Proof. Let us choose a basis from the vectors in (5.8) that contains $\mathbf{1}$ and define the linear map $Y(-,x)\colon V \to \mathrm{End}(V)[[x^{\pm1}]]$ by the formula (5.9). Let us check that the axioms **(VA1)**–**(VA4)** are satisfied. **(VA1)** is clearly true as $Y(-,x)$ is defined in terms of the n-th products and the n-th products preserve fields due to Proposition 5.8. **(VA2)** consists of two parts:

$Y(\mathbf{1}, x) = \mathrm{Id}_V$ is clear from the definition and that $\mathbf{1}$ is one of the basis vectors. Showing the the other part is not that different from the case of the Heisenberg vertex algebra (Exercise 3.4.1). **(VA4)** is a direct consequence of Dong's Lemma 5.12. **(VA3)** is shown as follows. From the definition of the n-th products, we can show

$$[\widetilde{T}, Y(a,x)] = \frac{d}{dx} Y(a,x).$$

When both sides are applied on $\mathbf{1}$, we get

$$\widetilde{T} Y(a,x)\mathbf{1} = \frac{d}{dx} Y(a,x)\mathbf{1}$$

as we have assumed $\widetilde{T}\mathbf{1} = 0$. Now, both sides only contain non-negative powers of x, so we can set $x = 0$ to obtain

$$\widetilde{T}(a) = a_{(-2)}\mathbf{1} = T(a).$$

This proves $\widetilde{T} = T$ and **(VA3)** simultaneously.

For another choice of basis vectors, we can get possibly a different linear map $\widetilde{Y}(-,x)\colon V \to \mathrm{End}(V)[[x^{\pm 1}]]$. Now, since $\mathbf{1}$ might not be a basis vector, we cannot show that $\widetilde{Y}(\mathbf{1}, x) = \mathrm{Id}_V$, but otherwise, $(V, \mathbf{1}, \widetilde{Y})$ is almost a vertex algebra. In particular, we can show that, for each $a \in V$,

- $\widetilde{Y}(a,x)\mathbf{1} = e^{x\widetilde{T}} a = Y(a,x)\mathbf{1}$,
- $\widetilde{Y}(a,x)$ and $Y(b,x)$ are mutually local for any $b \in V$.

Therefore, the uniqueness Theorem 5.13 ensures that $\widetilde{Y}(a,x) = Y(a,x)$, and the triple $(V, \mathbf{1}, \widetilde{Y})$ gives the same vertex algebra as $(V, \mathbf{1}, Y)$. $\square$

5.6 Example: Virasoro vertex algebra

As an application of the reconstruction theorem, let us see a family vertex algebras that is associated with the Virasoro algebra. The Virasoro algebra is the infinite dimensional Lie algebra

$$\mathfrak{vir} = \Big(\bigoplus_{n \in \mathbb{Z}} \mathbb{C} L_n \Big) \oplus \mathbb{C} C$$

along with

$$\begin{aligned} [L_m, L_n] &= (m-n)L_{m+n} + \frac{m^3 - m}{12} \delta_{m+n,0} C, \quad m, n \in \mathbb{Z}, \\ [C, \mathfrak{vir}] &= \{0\}. \end{aligned}$$

In order to apply the reconstruction theorem, we first need to set up a vector space V, a distinguished vector $\mathbf{1} \in V$, collections of vectors and fields labeled by a countable set S, and an operator $\widetilde{T}$. We construct the vector space of interest straight away. For $c \in \mathbb{C}$, we set

$$V_c := \mathcal{U}(\mathfrak{vir})/\mathcal{I}_c,$$

where

$$\mathcal{I}_c = \sum_{n \geq -1} \mathcal{U}(\mathfrak{vir}) \cdot L_n + \mathcal{U}(\mathfrak{vir}) \cdot (C - c).$$

We need to argue that V_c is a nonzero vector space. As a vector space, $\mathfrak{vir}$ decomposed into $\mathfrak{vir} = \mathfrak{vir}_+ \oplus \mathbb{C}C \oplus \mathfrak{vir}_-$, where

$$\mathfrak{vir}_+ = \bigoplus_{n \geq -1} \mathbb{C}L_n, \quad \mathfrak{vir}_- = \bigoplus_{n \leq -2} \mathbb{C}L_n.$$

Moreover, $\mathfrak{vir}_\pm$ and $\mathbb{C}C$ are all Lie subalgebras of $\mathfrak{vir}$, so they have their own enveloping algebras $\mathcal{U}(\mathfrak{vir}_\pm)$ and $\mathcal{U}(\mathbb{C}C) \simeq \mathbb{C}[C]$. The PBW theorem implies the isomorphism

$$\mathcal{U}(\mathfrak{vir}) \simeq \mathcal{U}(\mathfrak{vir}_-) \otimes \mathbb{C}[C] \otimes \mathcal{U}(\mathfrak{vir}_+)$$

of vector spaces. When we concentrate on $\mathcal{U}(\mathfrak{vir}_+)$, we notice that

$$\mathcal{U}(\mathfrak{vir}_+) = \mathbb{C} \cdot 1 \oplus \mathcal{U}(\mathfrak{vir}_+)\mathfrak{vir}_+ = \mathbb{C} \cdot 1 \oplus \sum_{n \geq -1} \mathcal{U}(\mathfrak{vir}_+) \cdot L_n$$

from the construction of the enveloping algebra. For the central part, it is clear that

$$\mathbb{C}[C] = \mathbb{C} \cdot 1 \oplus \mathbb{C}[C] \cdot (C - c).$$

Getting back to $\mathcal{U}(\mathfrak{vir})$, we have

$$\mathcal{U}(\mathfrak{vir}) \simeq \mathcal{U}(\mathfrak{vir}_-) \oplus \mathcal{I}_c,$$

and therefore,

$$V_c \simeq \mathcal{U}(\mathfrak{vir}_-).$$

In particular, V_c is a non-zero vector space.

As a distinguished vector, we take $\mathbf{1} = 1 + \mathcal{I}_c \in V_c$. For this example, we take $S = \{*\}$. The corresponding vector and field are $\omega := L_{-2}\mathbf{1}$ and

$$L(x) = \sum_{n \in \mathbb{Z}} L_n x^{-n-2} \in \mathcal{E}(V_c).$$

It will be convenient to write $\omega_{(n)} = L_{n-1}$, $n \in \mathbb{Z}$ in order to match the usual convention. Finally, the operator $\widetilde{T}$ is L_{-1}.

Next, let us check the conditions (1)–(4).

(1) By definition, $\omega_{(n)}\mathbf{1} = L_{n-1}\mathbf{1} = 0$ if $n \geq 0$. Therefore, we have

$$L(x)\mathbf{1} \in \omega + V[[x]]x.$$

(2) Again, $\widetilde{T}\mathbf{1} = L_{-1}\mathbf{1} = 0$ follows from the definition. The other condition is checked as

$$[\widetilde{T}, L(x)] = \sum_{n\in\mathbb{Z}}(-n-1)L_{n-1}x^{-n-2} = \frac{d}{dx}L(x).$$

(3) By direct computation, we get

$$\begin{aligned}&[L(x), L(y)]\\ =&\frac{c}{2}\partial_y^{(3)}(x^{-1}\delta(y/x)) + 2L(y)\partial_y(x^{-1}\delta(y/x)) + \partial L(y)x^{-1}\delta(y/x),\end{aligned}$$

which implies $(x-y)^4[L(x), L(y)] = 0$.

(4) From the PBW theorem, we can see that V_c has the following basis:

$$L_{-n_1}\cdots L_{-n_l}\mathbf{1}, \quad n_1 \geq \cdots \geq n_l \geq 2,\ l \in \mathbb{N}.$$

In particular, V_c is spanned by the following vectors

$$\omega_{(n_1)}\cdots\omega_{(n_l)}\mathbf{1}, \quad n_1, \ldots, n_l \in \mathbb{Z},\ l \in \mathbb{N}.$$

Therefore, we can apply the reconstruction theorem to show that there exists a vertex algebra $(V_c, \mathbf{1}, Y)$ such that

$$Y(\omega_{(n_1)}\cdots\omega_{(n_l)}\mathbf{1}, x) = L(x)_{(n_1)}\left(\cdots(L(x)_{(n_l)}\mathrm{Id}_{V_c})\cdots\right)$$

for $n_1, \ldots, n_l \in \mathbb{Z}$, $l \in \mathbb{N}$. If we concentrate on the PBW basis vectors, the formula of $Y(-, x)$ becomes

$$Y(L_{-n_1}\cdots L_{-n_l}\mathbf{1}, x) = {}^{\circ}_{\circ}\,\partial^{(n_1-2)}L(x)\cdots\partial^{(n_l-2)}L(x)\,{}^{\circ}_{\circ}$$

for $n_1 \geq \cdots \geq n_l \geq 2$, $l \in \mathbb{N}$. This vertex algebra is called the universal Virasoro vertex algebra of central charge c.

6 Vertex operator algebra

In the previous two sections, we have seen the definition and construction of vertex algebras as well as a few examples. A vertex **operator** algebra is a special case of vertex algebra, as we shall see in this section.

6.1 Definition

Before defining a vertex operator algebra, let us first recall the Virasoro algebra. The Virasoro algebra $\mathfrak{vir}$ is the infinite dimensional Lie algebra given by

$$\mathfrak{vir} = \left(\bigoplus_{n\in\mathbb{Z}}\mathbb{C}L_n\right) \oplus \mathbb{C}C$$

with the Lie bracket

$$\begin{aligned}
&[L_m, L_n] = (m-n)L_{m+n} + \frac{m^3-m}{12}\delta_{m+n,0}C, \quad m,n \in \mathbb{Z}, \\
&[C, \mathfrak{vir}] = \{0\}.
\end{aligned}$$

Definition 6.1. A vertex operator algebra is a quadruple $(V, \mathbf{1}, Y, \omega)$ where

- the triple $(V, \mathbf{1}, Y)$ is a vertex algebra.
- $\omega \in V$ is a distinguished vector

satisfying the following axioms.

(VOA1) Virasoro structure: When we apply $Y(-,x)$ to ω:

$$Y(\omega, x) = \sum_{n\in\mathbb{Z}} L_n^V x^{-n-2},$$

there exists $c \in \mathbb{C}$ and the assignments

$$\begin{aligned}
\mathfrak{vir} \to \mathrm{End}(V); \quad L_n &\mapsto L_n^V, \quad n \in \mathbb{Z}, \\
C &\mapsto c \cdot \mathrm{Id}_V
\end{aligned}$$

give a representation of the Virasoro algebra.

(VOA2) $\mathbb{Z}$-gradation: The operator L_0^V acts diagonally on V with integer eigenvalues and the eigenvalues are bounded below:

$$V = \bigoplus_{n\in\mathbb{Z}} V_n,$$

where

$$V_n = \ker(L_0^V - n \cdot \mathrm{Id}_V), \quad n \in \mathbb{Z}$$

are all finite-dimensional, and $V_n = 0$, $n \ll 0$.

(VOA3) L_{-1}-translation: $T = L_{-1}^V$, where T is defined from $Y(-,x)$.

We call the vector ω the conformal vector of the vertex operator algebra and the scalar c the central charge[2].

There are a few properties that are immediate from the definition. For example, due to the vacuum axiom **(VA2)** of a vertex algebra, we know that $\omega = L_{-2}^V\mathbf{1}$. It also follows that $\mathbf{1} \in V_0$ and $\omega \in V_2$ (Exercise 4.2.1).

Less obvious is that, for a vertex operator algebra, each operator of the form $a_{(n)}$, $a \in V$, $n \in \mathbb{Z}$ has a clear degree in terms of the $\mathbb{Z}$-gradation.

[2]Traditionally, c is also called the rank of the vertex operator algebra.

Proposition 6.2. *If $a \in V_m$, then,*

$$a_{(n)}V_d \subset V_{d+m-n-1}, \quad n, d \in \mathbb{Z}.$$

To prove this proposition, it is convenient to first see the following commutator formula that can be stated for a vertex algebra.

Lemma 6.3. *Let $(V, \mathbf{1}, Y)$ be a vertex algebra. For $a, b \in V$ and $m, n \in \mathbb{Z}$, we have*

$$[a_{(m)}, b_{(n)}] = \sum_{k=0}^{\infty} \binom{m}{k} \left(a_{(k)}b\right)_{(m+n-k)}.$$

Proof. This is a consequence of the Jacobi identity (Exercise 4.2.2). □

Proof of Proposition 6.2. Apply Lemma 6.3 appropriately (Exercise 4.2.3). □

6.2 Example: Virasoro VOA

Recall the Virasoro vertex algebra $(V_c, Y, \mathbf{1})$ of central charge $c \in \mathbb{C}$, which is generated by $\omega = L_{-2}\mathbf{1}$.

Proposition 6.4. *The quadruple $(V_c, \mathbf{1}, Y, \omega)$ is a vertex operator algebra of central charge c.*

Proof. As the triple $(V_c, \mathbf{1}, Y)$ is a vertex algebra and $\omega \in V_c$, we are to check the axioms **(VOA1)**–**(VOA3)**.

(VOA1) From the construction of the vertex algebra $(V_c, \mathbf{1}, Y)$, the coefficients of the series

$$Y(\omega, x) = \sum_{n \in \mathbb{Z}} L_n x^{-n-2}$$

form a representation of $\mathfrak{vir}$ of central charge c.

(VOA2) From the PBW theorem and the commutation relation $[L_0, L_n] = -nL_n$, $n \in \mathbb{Z}$, we observe $V_c = \bigoplus_{n=0}^{\infty}(V_c)_n$ with

$$(V_c)_n = \mathrm{Span}\left\{L_{-n_1} \cdots L_{-n_l}\mathbf{1} \,\middle|\, n_1 \geq \cdots \geq n_l \geq 2, \sum_{j=1}^{l} n_j = n\right\}$$

being the eigenspace of L_0 corresponding to the eigenvalue n. Indeed, each $(V_c)_n$ is finite-dimensional and there is no negative eigenvalue of L_0.

(VOA3) The identity $T = L_{-1}$ follows from the construction.

Therefore, the quadruple $(V_c, \mathbf{1}, Y, \omega)$ is a vertex operator algebra of central charge c. □

6.3 Example: Heisenberg VOA

The Heisenberg vertex algebra $(\mathcal{F}_0, |0\rangle, Y)$ is associated with the Heisenberg algebra, but there is no obvious action of the Virasoro algebra. Nevertheless, it can be equipped with the structure of a vertex operator algebra.

Proposition 6.5. *Let* $(\mathcal{F}_0, |0\rangle, Y)$ *be the Heisenberg vertex algebra, and*

$$\omega = \frac{1}{2}\alpha_{-1}^2 |0\rangle .$$

Then, the quadruple $(\mathcal{F}_0, |0\rangle, Y, \omega)$ *is a vertex operator algebra of central charge* 1.

Proof. Let us write the field corresponding to ω as

$$L(x) = \sum_{n\in\mathbb{Z}} L_n x^{-n-2} := Y(\omega, x) = \frac{1}{2} {}^\circ_\circ \alpha(x)^2 {}^\circ_\circ .$$

When we apply $L(x)$ to $|0\rangle$, we can see that $L_n|0\rangle = 0$, $n \geq -1$. In particular, the vacuum vector $|0\rangle$ is an eigenvector of L_0 corresponding to the eigenvalue 0.

(VOA1) The goal is to show

$$[L(x), L(y)] = \frac{1}{2}\partial_y^{(3)} x^{-1}\delta(y/x) + 2L(y)\partial_y x^{-1}\delta(y/x) + \partial L(y) x^{-1}\delta(y/x).$$

By Wick's formula, we get

$$L(x)L(y) = \frac{1/2}{(x-y)^4} + \frac{1}{(x-y)^2} {}^\circ_\circ \alpha(x)\alpha(y) {}^\circ_\circ + \frac{1}{4} {}^\circ_\circ \alpha(x)^2\alpha(y)^2 {}^\circ_\circ .$$

Notice that the normally ordered product is commutative for the Heisenberg algebra. Thus, the commutator reads

$$[L(x), L(y)] = \frac{1}{2}\partial_y^{(3)} x^{-1}\delta(y/x) + {}^\circ_\circ \alpha(x)\alpha(y) {}^\circ_\circ \partial_y x^{-1}\delta(y/x).$$

Here, we apply a version of Taylor's formula (Exercise 4.3.1) to the last term to get

$$\begin{aligned}[L(x), L(y)] = {} & \frac{1}{2}\partial_y^{(3)} x^{-1}\delta(y/x) + {}^\circ_\circ \alpha(y)^2 {}^\circ_\circ \partial_y x^{-1}\delta(y/x) \\ & + {}^\circ_\circ (\partial\alpha(y))\alpha(y) {}^\circ_\circ x^{-1}\delta(y/x).\end{aligned}$$

This is precisely the desired result.

(VOA2) By a similar computation, we can verify the commutation relations

$$[L_n, \alpha(y)] = ((n+1)y^n + y^{n+1}\partial_y)\alpha(y), \quad n \in \mathbb{Z}, \tag{6.1}$$

which, in particular, contain the case $n = 0$ implying

$$[L_0, \alpha_m] = -m\alpha_m, \quad m \in \mathbb{Z}.$$

From the PBW theorem, the Fock space $\mathcal{F}_0$ is decomposed into

$$\mathcal{F}_0 = \bigoplus_{n=0}^{\infty} (\mathcal{F}_0)_n,$$

where each

$$(\mathcal{F}_0)_n = \mathrm{Span}\left\{\alpha_{-n_1} \cdots \alpha_{-n_l}\mathbf{1} \middle| n_1 \geq \cdots \geq n_l > 0, \sum_{j=1}^{l} n_j = n\right\}$$

is the eigenspace of L_0 corresponding to the eigenvalue n. Each $(\mathcal{F}_0)_n$ is clearly finite-dimensional, and the eigenvalues of L_0 are all non-negative.

(VOA3) When we set $n = -1$ in (6.1), we obtain

$$[L_{-1}, \alpha(y)] = \frac{d}{dy}\alpha(y).$$

We have also seen that $L_{-1} |0\rangle = 0$. From the proof of the reconstruction theorem (Theorem 5.14) and the fact that the Heisenberg vertex algebra is generated by $\alpha_{-1} |0\rangle$ (Exercise 4.1.1), such an operator must coincide with T.

Therefore, the desired result follows. □

7 Interlude: Wightman axioms

We have defined vertex (operator) algebras and seen examples. In case they look just ad hoc and technical, let us see how they are motivated by conformal field theory.

This section is intended to present motivation, but should not be entirely precise. The text here mainly follows the heuristics in [Kac98, Chapter 1]. There is also an interesting math paper [RTT22].

7.1 Wightman axioms of quantum field theory

There are several mathematical formulations of quantum field theory. Among those, Wightman's set of axioms is the most classical one.

Wightman's axioms define a quantum field theory on a Minkowski space of any dimension. However, the case of two dimensions is relevant for us. The two dimensional Minkowski space $M = \mathbb{R}^{1,1}$ is $\mathbb{R}^2$ equipped with the Minkowski inner product:

$$\langle x, y \rangle = x_0 y_0 - x_1 y_1, \quad x = (x_0, x_1),\, y = (y_0, y_1) \in M.$$

The *zero-th* component x_0 is considered time, the other is the physical space. The Lorentz group $\mathcal{L}$ is the group of isometries of M. We, in particular, take the following connected component:

$$\begin{aligned} \mathcal{L}_+^\uparrow &= \left\{ \Lambda = \begin{pmatrix} \Lambda_{00} & \Lambda_{01} \\ \Lambda_{10} & \Lambda_{11} \end{pmatrix} \in \mathcal{L} \middle| \Lambda_{00} > 0,\, \det \Lambda = 1 \right\} \\ &= \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} \middle| a > 0,\, b \in \mathbb{R},\, a^2 - b^2 = 1 \right\} \end{aligned}$$

and call it the restricted Lorentz group. Taking the semi-direct product with the translation group, we get the restricted Poincaré group:

$$\mathcal{P}_+^\uparrow = M \rtimes \mathcal{L}_+^\uparrow$$

that acts on M by $(a, \Lambda)\colon x \mapsto \Lambda x + a$, $(a, \Lambda) \in \mathcal{P}_+^\uparrow$, $x \in M$.

Two subsets $A, B \subset M$ are said to be space-like separated if

$$\|x - y\|^2 < 0, \quad \text{for all } x \in A,\, y \in B.$$

The subset

$$V_+ = \{x = (x_0, x_1) | \, \|x\|^2 \geq 0,\, x_0 \geq 0\}$$

is called the forward cone.

Definition 7.1. A quantum field theory on M consists of

- a Hilbert space $\mathcal{H}$ of states,
- a distinguished vector $|0\rangle \in \mathcal{H}$ called the vacuum,
- a unitary representation $U\colon \mathcal{P}_+^\uparrow \to U(\mathcal{H})$, and
- a collection of operator-valued distributions $\{\Phi_a\}_{a \in I}$ with test functions in $\mathcal{S}(M)$

satisfying the following axioms:

(W1) Poincaré covariance: For any $f \in \mathcal{S}(M)$ and $g \in \mathcal{P}_+^\uparrow$,

$$U(g)\Phi_a(f)U(g)^{-1} = \Phi_a(g.f).$$

By Stone's theorem, the are commuting self-adjoint operators P_0, P_1 such that

$$U((a,I)) = \exp(\mathbb{i}(a_0P_0 + a_1P_1)), \quad a = (a_0, a_1) \in M.$$

(W2) Stable vacuum: The vacuum is invariant under $\mathcal{P}_+^\uparrow$: $U(g)\,|0\rangle = |0\rangle$, $g \in \mathcal{P}_+^\uparrow$. Furthermore, $\mathrm{Spec}(P_0, P_1) \subset V_+$.

(W3) Completeness:

$$\mathcal{H} = \overline{\mathrm{Span}\{\Phi_{a_1}(f_1)\cdots\Phi_{a_l}(f_l)\,|0\rangle\}}.$$

(W4) Locality: For $f, g \in \mathcal{S}(M)$, if $\mathrm{supp}(f)$ and $\mathrm{supp}(g)$ are space-like separated, we have

$$\Phi_a(f)\Phi_b(g) = \Phi_b(g)\Phi_a(f), \quad a, b \in I.$$

A moral lesson from Wightman's axioms is that a quantum field theory is a collection of fields that transform properly under a given symmetry group. Therefore, if one changes the symmetry group, one may get another class of field theory.

7.2 Chiral Wightman Möbius CFT

The Poincaré group is the group of transformations that preserve the Minkowski metric. As a larger group, we may take the group of conformal transformations, which only preserve the angle, but possibly do not the distance.

It is known that conformal transformations of the two dimensional Minkowski space are decoupled into two pieces in the so-called light-cone coordinates: when we introduce $t = x_0 - x_1$ and $\bar{t} = x_0 + x_1$, the group of conformal transformations of M is isomorphic to

$$\mathrm{Diff}(\mathbb{R}) \times \mathrm{Diff}(\mathbb{R}),$$

where the first and second components act on t and $\bar{t}$, respectively. For this fact, see e.g. [Sch08, Chapter 2].

We are going to do two nontrivial operations to get the object we will be interested in. First, we only take one light ray $\mathbb{R}$ that is parametrized by t. Then, the group of conformal transformations is simply $\mathrm{Diff}(\mathbb{R})$. It is sometimes convenient if the space is compact. So, second, we compactify $\mathbb{R}$ to S^1 with the group of conformal transformations $\mathrm{Diff}(S^1)$. It is probably natural to take $\mathrm{Diff}_+(S^1)$, the group of orientation preserving diffeomorphisms, i.e., the connected component that contains the identity transformation. Because we have taken one of two rays, the CFT we will consider is called **chiral**.

By definition of the light-cone coordinates, the generator of translation along t is

$$P = \frac{1}{2}(P_0 - P_1).$$

If the joint spectrum of (P_0, P_1) is contained in V_+ as required in **(W2)**, the spectrum of P lies in $\mathbb{R}_{\geq 0}$.

We first focus on the finite dimensional subgroup $\mathrm{PSL}(2,\mathbb{R})$ of $\mathrm{Diff}_+(S^1)$ that acts on S^1 by Möbius transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : t \mapsto \frac{at+b}{ct+d}.$$

Definition 7.2. A chiral Möbius CFT on S^1 consists of

- a Hilbert space $\mathcal{H}$ of states,
- a distinguished vector $|0\rangle \in \mathcal{H}$ called the vacuum,
- a unitary representation $U\colon \mathrm{PSL}(2,\mathbb{R}) \to U(\mathcal{H})$, and
- a collection of operator-valued distributions $\{\Phi_a\}_{a\in I}$ on S^1

satisfying the following axioms:

(W1') PSL(2, ℝ) covariance: For each $a \in I$, there is $\Delta_a \in \mathbb{R}$, such that

$$U(g)\Phi_a(t)U(g)^{-1} = \frac{1}{(ct+d)^{2\Delta_a}}\Phi_a(g.t), \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2,\mathbb{R}).$$

The generator P of translations is defined by

$$U\left(\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)\right) = \exp(\mathbb{i}bP), \quad b \in \mathbb{R}.$$

(W2') Stable vacuum: The vacuum is invariant under $\mathrm{PSL}(2,\mathbb{R})$: $U(g)\,|0\rangle = |0\rangle$, $g \in \mathrm{PSL}(2,\mathbb{R})$. Furthermore, $\mathrm{Spec}(P) \subset \mathbb{R}_{\geq 0}$.

(W3') Completeness:

$$\mathcal{H} = \overline{\mathrm{Span}\{\Phi_{a_1}(f_1)\cdots\Phi_{a_l}(f_l)\,|0\rangle\}}.$$

(W4') Locality: If $\mathrm{supp}(f) \cap \mathrm{supp}(g) = \emptyset$, we have

$$\Phi_a(f)\Phi_b(g) = \Phi_b(g)\Phi_a(f), \quad a, b \in I.$$

Remark 7.3. The multiplicative factor $\frac{1}{(ct+d)^2}$ in **(W1')** is, of course, the Jacobian of the transformation

$$t \mapsto \frac{at+b}{ct+d}.$$

The group $\mathrm{PSL}(2, \mathbb{R})$ is generated by

- translations $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$, $b \in \mathbb{R}$, and
- special conformal transformations $\left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)$, $c \in \mathbb{R}$.

The generator of the translations has been already introduced and denoted by P. By **(W1')**, it governs the infinitesimal translation of fields as

$$\mathbb{i}[P, \Phi_a(t)] = \partial_t \Phi_a(t).$$

Similarly, let Q be the generator of the special conformal transformations defined by

$$U\left(\left(\begin{smallmatrix} 1 & 0 \\ c & 1 \end{smallmatrix}\right)\right) = \exp(-\mathbb{i}cQ), \quad c \in \mathbb{R}.$$

By **(W1')**, the operator Q yields the infinitesimal transformation

$$\mathbb{i}[Q, \Phi_a(t)] = (2\Delta_a t + t^2 \partial_t)\Phi_a(t).$$

The Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ of $\mathrm{PSL}(2, \mathbb{R})$ is represented on $\mathcal{H}$ and

$$dU(\mathfrak{sl}(2, \mathbb{R})) = \mathrm{Span}_{\mathbb{R}}\{P, Q, [P, Q]\}.$$

As a coordinate on S^1, we may take

$$z = \frac{1 + \mathbb{i}t}{1 - \mathbb{i}t}.$$

The assignment $t \mapsto z$ is the extension of the conformal map

$$\mathbb{H} = \{t \in \mathbb{C} | \mathrm{Im} t \geq 0\} \to \mathbb{D} = \{z \in \mathbb{C} | |z| < 1\}$$

defined by the same formula to the boundary.

Let us define

$$Y(a, z) := \frac{1}{(1+z)^{2\Delta_a}} \Phi_a(t)$$

and apply it to the vacuum vector $|0\rangle$:

$$Y(a, z) |0\rangle = \frac{1}{(1+z)^{2\Delta_a}} e^{\mathbb{i}tP} \Phi_a(0) |0\rangle.$$

Due to **(W2)**, the operator $e^{\mathbb{i}tP}$ is bounded if $\mathrm{Im} t > 0$, so it makes sense to specialize at $t = \mathbb{i}$ to get a well-defined vector

$$a := Y(a, z) |0\rangle |_{z=0} \in \mathcal{H}.$$

This indeed looks like one of the vacuum axiom **(VA2)** for a vertex algebra.

We can find operators T, H, T^* by linear combination of $P, Q, [P, Q]$ such that

$$\begin{aligned}[T, Y(a,z)] &= \partial_z Y(a,z), \\ [H, Y(a,z)] &= (z\partial_z + \Delta_a)Y(a,z), \\ [T^*, Y(a,z)] &= (z^2\partial_z + 2\Delta_a z)Y(a,z).\end{aligned}$$

The first one is the translation axiom **(VA3)** for a vertex algebra. The second and third equations are not required for a vertex algebra, but when we apply them to the vacuum and set $z = 0$, we get interesting results:

$$Ha = \Delta_a a, \quad T^* a = 0.$$

As a consequence of the locality **(W4')**, we have

$$[\Phi_a(t), \Phi_b(t')] = \sum_{j\geq 0} \partial^{(j)}\delta(t - t')\Psi^j(t')$$

If the right-hand side is a finite sum, we get

$$(z - w)^N [Y(a,z), Y(b,w)] = 0.$$

7.3 Virasoro algebra

Of course, the full conformal symmetry is $\mathrm{Diff}_+(S^1)$, so a genuine CFT must carry an action of this group as well. Note that in a quantum theory, a group action might be only projective.

Definition 7.4. A chiral Möbius CFT is a chiral CFT if the unitary representation of $\mathrm{PSL}(2,\mathbb{R})$ can be extended to a projective unitary representation of $\mathrm{Diff}_+(S^1)$, under which the fields transform covariantly.

As $\mathrm{Diff}_+(S^1)$ is infinite dimensional, it is easier to deal with its Lie algebra. The Lie algebra of $\mathrm{Diff}_+(S^1)$ is $\mathrm{Vect}(S^1)$, the Lie algebra of vector fields on S^1. After extending the coefficients to $\mathbb{C}$, we may find its subalgebra

$$\mathfrak{Witt} = \bigoplus_{n\in\mathbb{Z}} \mathbb{C}\ell_n, \quad \ell_n = -z^{n+1}\partial_z$$

called the Witt algebra. The Lie bracket of $\mathfrak{Witt}$ is given by

$$[\ell_m, \ell_n] = (m - n)\ell_{m+n}, \quad m, n \in \mathbb{Z}.$$

A projective representation is equivalent to a representation of a central extension. In the case of the Witt algebra, its (essentially only) nontrivial central extension is given by the Virasoro algebra $\mathfrak{vir}$:

$$0 \to \mathbb{C}C \to \mathfrak{vir} \xrightarrow{\pi} \mathfrak{Witt} \to 0 \quad \text{(exact)}.$$

Let L_n be such that $\pi(L_n) = \ell_n$, $n \in \mathbb{Z}$. The Lie bracket of $\mathfrak{vir}$ is

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{m^3 - m}{12}\delta_{m+n,0}C, \quad m, n \in \mathbb{Z}.$$

Therefore, by the very definition of a chiral CFT, it carries a representation of the Virasoro algebra. Furthermore, the Fourier series

$$L(z) = \sum_{n \in \mathbb{Z}} L_n z^{-n-2}$$

is considered to be the stress-energy tensor and reasonably one of our quantum fields. In particular,

$$\omega = L(z) |0\rangle |_{z=0} \in \mathcal{H}$$

exists.

8 Modules over vertex algebras

If we agree that a vertex algebra is an 'algebra', it would make sense to consider its modules. In this section, we discuu modules over vertex algebras, but postpone modules for vertex operator algebras in a later section.

8.1 Motivation

Recall that the Heisenberg algebra admits a one-parameter family of Fock representations $\mathcal{F}_\lambda$, $\lambda \in \mathbb{C}$. Among those Fock representations, only $\mathcal{F}_0$ can be equipped with the structure of a vertex algebra in a reasonable way. This is due to the fact that, if we apply the field

$$\alpha(x) = \sum_{n \in \mathbb{Z}} \alpha_n x^{-n-1}$$

to the generating vector $|\lambda\rangle$, we get

$$\alpha(x) |\lambda\rangle = \lambda |\lambda\rangle x^{-1} + \alpha_{-1} |\lambda\rangle + \cdots .$$

Thus, in order for the vacuum axiom to hold, we must have $\lambda = 0$.

On the other hand, apart from the vacuum axiom, the field $\alpha(x)$ still seems to be a meaningful object acting on a general $\mathcal{F}_\lambda$. As we will see below, $\mathcal{F}_\lambda$, $\lambda \in \mathbb{C}$ are modules for the vertex algebra $\mathcal{F}_0$.

8.2 Definition, characterization

Before defining a module over a vertex algebra, let us recall that a vertex algebra is a triple $(V, \mathbf{1}, Y)$ satisfying axioms **(VA1)**–**(VA4)**. A general philosophy of defining a module over an algebra is to transfer all relations satisfied by the algebra to another vector space. However, as the Fock representations of the Heisenberg algebra suggest, we cannot expect analogues of some of the axioms on a modules; a vacuum vector might not exist in a module. This means that one of **(VA2)** cannot be imposed on a module, and **(VA3)** does not make any sense as the definition of T uses the vacuum vector. On the other hand, just dropping **(VA3)** does not seem to give a reasonable definition of a module (the locality alone is too weak).

To get a correct definition of a module, we should notice that, under **(VA1)** and **(VA2)**, **(VA3)** and **(VA4)** are equivalent to the Jacobi identities. Thus, a module should be defined as follows.

Definition 8.1. Let $(V, Y, \mathbf{1})$ be a vertex algebra. A module over the vertex algebra, or simply a V-module, is a pair (W, Y_W) of

- a vector space W, and
- a linear map

$$Y_W(-,x)\colon V \to \mathrm{End}(W)[[x^{\pm 1}]]; \quad a \mapsto Y_W(a,x) = \sum_{n\in\mathbb{Z}} a^W_{(n)} x^{-n-1}$$

that satisfy the following axioms:

(VAmod1) Field condition: For any $a \in V$ and $w \in W$,

$$Y_W(a,x)w \in W((x)).$$

(VAmod2) Vacuum axiom: The vacuum vector 'acts as identity':

$$Y_W(\mathbf{1},x) = \mathrm{Id}_W \cdot x^0.$$

(VAmod3) Jacobi identity: For any $a, b \in V$, we have

$$\begin{aligned} &x_0^{-1}\delta\left(\frac{x_1-x_2}{x_0}\right) Y_W(a,x_1)Y_W(b,x_2) \\ - &x_0^{-1}\delta\left(\frac{x_2-x_1}{-x_0}\right) Y_W(b,x_2)Y_W(a,x_1) \\ = &x_2^{-1}\delta\left(\frac{x_1-x_0}{x_2}\right) Y_W\left(Y(a,x_0)b,x_2\right). \end{aligned}$$

A problem with this definition is that, in order to show that a pair (W, Y_W) is a V-module, we need to check the complicated Jacobi identities. Of course, we would like to reduce the number of properties we have to check.

From the Jacobi identity, it follows that

$$Y_W(a_{(n)}b, x) = Y_W(a, x)_{(n)}Y_W(b, x)$$

for all $a, b \in V$ and $n \in \mathbb{Z}$. It turns out that they imply the Jacobi identity under assuming **(VAmod1)** and **(VAmod2)**.

Proposition 8.2. *Let $(V, \mathbf{1}, Y)$ be a vertex algebra and (W, Y_W) be a pair of a vector space W and a linear map*

$$Y_W(-, x)\colon V \to \mathrm{End}(W)[[x^{\pm 1}]].$$

If they satisfy **(VAmod1)**, **(VAmod2)** *and*

$$Y_W(a_{(n)}b, x) = Y_W(a, x)_{(n)}Y_W(b, x) \tag{8.1}$$

for all $a, b \in V$ and $n \in \mathbb{Z}$, then (W, Y_W) is a V-module.

Proof. **Step 1: Translation property.**

Lemma 8.3. *For any $a \in V$, we have*

$$Y(Ta, x) = \frac{d}{dx}Y(a, x).$$

Proof. This is straightforward from the definition of T and the assumptions:

$$Y_W(Ta, x) = Y_W(a_{(-2)}\mathbf{1}, x) = Y_W(a, x)_{(-2)}\mathrm{Id}_W.$$

From the definition of the (-2)-nd product and integration by parts, we obtain the desired result. $\square$

Step 2: Weak associativity.

Lemma 8.4. *For all $a \in V$ and $w \in W$, there exists $k \gg 0$ such that*

$$(x_0 + x_2)^k Y_W(Y(a, x_0)b, x_2)w = (x_0 + x_2)^k Y_W(a, x_0 + x_2)Y_W(b, x_2)w$$

for all $b \in V$.

Proof. The assumption gives

$$\begin{aligned} Y_W(Y(a, x_0)b, x_2)w = [x_1^{-1}]\Bigg\{ & x_0^{-1}\delta\left(\frac{x_1 - x_2}{x_0}\right) Y_W(a, x_1)Y_W(b, x_2)w \\ & - x_0^{-1}\delta\left(\frac{x_2 - x_1}{-x_0}\right) Y_W(b, x_2)Y_W(a, x_1)w \Bigg\} \end{aligned}$$

Multiplying $(x_0+x_2)^k$ on both sides and using some properties of the delta distribution, we can observe

$$\begin{aligned}
&(x_0+x_2)^k Y_W(Y(a,x_0)b,x_2)w \\
&= [x_1^{-1}]\Bigg\{ x_1^{-1}\delta\left(\frac{x_0+x_2}{x_1}\right)(x_0+x_2)^k Y_W(a,x_0+x_2)Y_W(b,x_2)w \\
&\qquad - x_0^{-1}\delta\left(\frac{x_2-x_1}{-x_0}\right) Y_W(b,x_2)x_1^k Y_W(a,x_1)w \Bigg\}.
\end{aligned}$$

As we have assumed **(VAmod1)**, there is a large enough k (depending on a and w) such that $x_1^k Y_W(a,x_1)w \in W[[x_1]]$. For such a choice, the right-hand side becomes the desired one. □

Step 3: Matrix elements. Let $a,b\in V$, $w\in W$ and $\varphi\in W^*$ be arbitrarily fixed. Then, by Lemma 8.4, there exists a unique element

$$f(x_1,x_2)\in\mathbb{C}[[x_1,x_2]][x_1^{-1},x_2^{-1},(x_1-x_2)^{-1}]$$

such that

$$\begin{aligned}
\langle\varphi, Y_W(a,x_1)Y_W(b,x_2)w\rangle &= \iota_{12}f(x_1,x_2) \\
\langle\varphi, Y_W(Y(a,x_0)b,x_2)w\rangle &= \iota_{20}f(x_0+x_2,x_2).
\end{aligned}$$

If we change the roles of (a,x_1) and (b,x_2) and replace x_0 by $-x_0$, we can see that there exists a unique element

$$g(x_1,x_2)\in\mathbb{C}[[x_1,x_2]][x_1^{-1},x_2^{-1},(x_1-x_2)^{-1}]$$

such that

$$\begin{aligned}
\langle\varphi, Y_W(b,x_2)Y_W(a,x_1)w\rangle &= \iota_{21}g(x_1,g_2) \\
\langle\varphi, Y_W(Y(b,-x_0)a,x_1)w\rangle &= \iota_{10}g(x_1,x_1-x_0).
\end{aligned}$$

If we can show that $f=g$, the Jacobi identity follows as before. Now we use Lemma 8.3 and the already known skew-symmetry on V as follows:

$$\begin{aligned}
\iota_{20}f(x_0+x_2,x_2) &= \langle\varphi, Y_W(Y(a,x_0)b,x_2)w\rangle, \\
&= \langle\varphi, Y_W(e^{x_0T}Y(b,-x_0)a,x_2)w\rangle \\
&= \langle\varphi, Y_W(Y(b,-x_0)a,x_2+x_0)w\rangle \\
&= \iota_{10}g(x_1,x_1-x_0)|_{x_1=x_2+x_0} \\
&= \iota_{20}g(x_2+x_0,x_2).
\end{aligned}$$

Therefore, $f=g$. □

Remark 8.5. It might be convenient to compare the derivation of the Jacobi identity here and the previous one for a vertex algebra. Previously, the Jacobi identity was a consequence of weak commutativity and weak associativity, but here, only (8.1) implies the Jacobi identity. This might seem strange since (8.1) does not imply weak commutativity. The trick is that we already know that V is a vertex algebra and, in fact, we used the skew-symmetry of a vertex algebra.

8.3 Example: Heisenberg vertex algebra

As we have already anticipated, Fock representations $\mathcal{F}_\lambda$ of the Heisenberg algebra are given the structure of modules over the Heisenberg vertex algebra $(\mathcal{F}_0, \mathbf{1}, Y)$. The task is, of course, to define a linear map

$$Y_{\mathcal{F}_\lambda}(-,x)\colon \mathcal{F}_0 \to \mathrm{End}(\mathcal{F}_\lambda)[[x^{\pm 1}]]$$

and check the conditions in Proposition 8.2.

Let us start with getting the linear map. Recall that, for a vector space W, the space of fields on W is denoted by $\mathcal{E}(W)$. The following proposition says that the Hensenberg algebra $\widehat{\mathfrak{h}}$ acts, not only on each $\mathcal{F}_\lambda$, but also on the space of fields on it, $\mathcal{E}(\mathcal{F}_\lambda)$.

Proposition 8.6. *The linear map* $\widehat{\mathfrak{h}} \to \mathrm{End}(\mathcal{E}(\mathcal{F}_\lambda))$ *defined by*

$$\alpha_n \mapsto \alpha(x)_{(n)}, \quad n \in \mathbb{Z}, \quad K \mapsto \mathrm{Id}_{\mathcal{E}(\mathcal{F}_\lambda)}$$

is a Lie algebra homomorphism. Furthermore, on this representation of $\widehat{\mathfrak{h}}$*, we have*

$$\alpha(x)_{(n)}\mathrm{Id}_{\mathcal{F}_\lambda} = 0, \quad n \geq 0.$$

Idea of Proof. For an arbitrary field $C(x) \in \mathcal{E}(\mathcal{F}_\lambda)$ and $m, n \in \mathbb{Z}$, we have

$$\begin{aligned}
\alpha(x)_{(m)}\alpha(x)_{(n)}C(x) = {}& [x_1^{-1}]\Big\{(x_1 - x)^m \alpha(x_1)\alpha(x)_{(n)}C(x) \\
& \qquad - (-x + x_1)^m \alpha(x)_{(n)}C(x)\alpha(x_1)\Big\} \\
= {}& [x_1^{-1}x_2^{-1}]\Big\{(x_1 - x)^m (x_2 - x)^n \alpha(x_1)\alpha(x_2)C(x) \\
& \qquad - (x_1 - x)^m (-x + x_2)^n \alpha(x_1)C(x)\alpha(x_2) \\
& \qquad - (-x + x_1)^m (x_2 - x)^n \alpha(x_2)C(x)\alpha(x_1) \\
& \qquad + (-x + x_1)^m (-x + x_2)^n C(x)\alpha(x_2)\alpha(x_1)\Big\}.
\end{aligned}$$

When we exchange m and n, the second and third terms in the last sum are exchanged too, so they are cancelled in the commutator. Therefore, computing the commutator of $\alpha(x)_{(m)}$ and $\alpha(x)_{(n)}$ amounts to *integrating* that of $\alpha(x_1)$ and $\alpha(x_2)$. Now, it is not surprising that the action of $\alpha(x)_{(n)}$ knows the Heisenberg algebra. A complete proof is left for Exercise 5.2.1. □

Proposition 8.6 implies (see also Exercise 5.3.1) that there exists a unique homomorphism

$$\Phi_\lambda \colon \mathcal{F}_0 \to \mathcal{E}(\mathcal{F}_\lambda)$$

of representations of $\widehat{\mathfrak{h}}$ such that

$$\Phi_\lambda(\mathbf{1}) = \mathrm{Id}_{\mathcal{F}_\lambda}.$$

As the target space of Φ_λ is the space of fields on $\mathcal{F}_\lambda$, we can view Φ_λ as a linear map

$$Y_{\mathcal{F}_\lambda}(-,x)\colon \mathcal{F}_0 \to \mathrm{End}(\mathcal{F}_\lambda)[[x^{\pm 1}]]; \quad a \mapsto \Phi_\lambda(a).$$

Then, $Y_{\mathcal{F}_\lambda}(-,x)$ clearly satisfies **(VAmod1)** and **(VAmod2)**. The claim is that (8.1) holds for $Y_{\mathcal{F}_\lambda}(-,x)$, which can be shown by the following trick.

Let us write $\mathcal{E}_0(\mathcal{F}_\lambda)$ for the image of Φ_λ. By definition, we have a spanning set of $\mathcal{E}_0(\mathcal{F}_\lambda)$:

$$\mathcal{E}_0(\mathcal{F}_\lambda) = \mathrm{Span}\big\{\alpha(x)_{(-n_1)} \cdots \alpha(x)_{(-n_l)}\mathrm{Id}_{\mathcal{F}_\lambda} | n_1 \geq \cdots \geq n_l > 0,\, l \in \mathbb{N}\big\}.$$

We fix an element

$$A = \alpha(x)_{(-1)}\mathrm{Id}_{\mathcal{F}_\lambda} = \alpha(x) \in \mathcal{E}_0(\mathcal{F}_\lambda)$$

and a field

$$A(\zeta) = \sum_{n\in\mathbb{Z}} a(x)_{(n)}\zeta^{-n-1} \in \mathcal{E}(\mathcal{E}_0(\mathcal{F}_\lambda)).$$

on $\mathcal{E}_0(\mathcal{F}_\lambda)$. This is indeed a field because the action of $\alpha(x)_{(n)}$ on $\mathcal{E}_0(\mathcal{F}_\lambda)$ is intertwined by Φ_λ with that of α_n on $\mathcal{F}_0$. We also take an operator

$$T' = \frac{d}{dx} \in \mathrm{End}(\mathcal{E}_0(\mathcal{F}_\lambda)).$$

One can check that T' is an operator on $\mathcal{E}_0(\mathcal{F}_\lambda)$ by observing

$$\alpha'(x)_{(n)} = -n\alpha(x)_{(n-1)}, \quad n \in \mathbb{Z}.$$

Proposition 8.7. *There exists a unique vertex algebra* $(\mathcal{E}_0(\mathcal{F}_\lambda), \mathrm{Id}_{\mathcal{F}_\lambda}, Y_{\mathcal{E}})$ *such that*

$$Y_{\mathcal{E}}\left(\alpha(x)_{(-n_1)} \cdots \alpha(x)_{(-n_l)}\mathrm{Id}_{\mathcal{F}_\lambda}, \zeta\right) = {}^{\circ}_{\circ}\partial^{(n_1-1)}A(\zeta)\cdots\partial^{(n_l-1)}A(\zeta){}^{\circ}_{\circ}$$

for $n_1 \geq \cdots \geq n_l$, $l \in \mathbb{N}$.

Proof. Exercise 5.4.1 □

As the last preliminary, we introduce a filtration of $\mathcal{F}_0$. For each $p \in \mathbb{N}$, we set

$$F_p\mathcal{F}_0 = \mathrm{Span}\{\alpha_{-n_1}\cdots\alpha_{-n_l}\mathbf{1} | n_1 \geq \cdots \geq n_l,\, l \leq p\}.$$

Then, the family $\{F_p\mathcal{F}_0\}_{p\in\mathbb{N}}$ forms an increasing and exhausting filtration of $\mathcal{F}_0$ called the PBW filtration.

We are now ready to get modules over the Heisenberg vertex algebra.

Theorem 8.8. *For each λ, the pair $(\mathcal{F}_\lambda, Y_{\mathcal{F}_\lambda})$ is a module over the Heisenberg vertex algebra $(\mathcal{F}_0, \mathbf{1}, Y)$.*

Proof. As we have already noticed, $Y_{\mathcal{F}_\lambda}(-,x)$ satisfies **(VAmod1)** and **(VAmod2)**, it suffices to show (8.1). To ease the notation, we shall use Φ_λ instead of $Y_{\mathcal{F}_\lambda}(-,x)$. Then, the conditions to be checked are

$$\Phi_\lambda(a_{(n)}b) = \Phi_\lambda(a)_{(n)}\Phi_\lambda(b), \quad a, b \in \mathcal{F}_0,\, n \in \mathbb{Z},$$

but they are equivalent to

$$\Phi_\lambda(Y(a,\zeta)b) = Y_{\mathcal{E}}(\Phi_\lambda(a),\zeta)\Phi_\lambda(b), \quad a, b \in \mathcal{F}_0.$$

Let us introduce the following subspace of $\mathcal{F}_0$:

$$U = \{a \in \mathcal{F}_0 | \Phi_\lambda(Y(a,\zeta)b) = Y_{\mathcal{E}}(\Phi_\lambda(a),\zeta)\Phi_\lambda(b) \,\text{for all}\, b \in \mathcal{F}_0\}.$$

We show that $U = \mathcal{F}_0$ by induction along the PBW filtration. It is clear that $F_1\mathcal{F}_0 \subset U$. Let us assume that $F_p\mathcal{F}_0 \subset U$. Any element in $F_{p+1}\mathcal{F}_0$ has the form $\alpha_{-n}c$ with some $n > 0$ and $c \in F_p\mathcal{F}_0$. Then, we can see

$$\begin{aligned}\Phi_\lambda(Y(\alpha_{-n}c,\zeta)b) &= \Phi_\lambda(\alpha(\zeta)_{(-n)}Y(c,\zeta)b)\\ &= [\xi^{-1}]\Big\{(\xi-\zeta)^{-n}\Phi_\lambda(\alpha(\xi)Y(c,\zeta)b)\\ &\qquad - (-\zeta+\xi)^{-n}\Phi_\lambda(Y(c,\zeta)\alpha(\xi)b)\Big\}.\end{aligned}$$

From the induction hypothesis, the right-hand side becomes

$$\begin{aligned}&[\xi^{-1}]\Big\{(\xi-\zeta)^{-n}A(\xi)Y_{\mathcal{E}}(\Phi_\lambda(c),\zeta)\Phi_\lambda(b)\\ &\qquad - (-\zeta+\xi)^{-n}Y_{\mathcal{E}}(\Phi_\lambda(c),\zeta)A(\xi)\Phi_\lambda(b)\Big\}\\ &= Y_{\mathcal{E}}(\alpha(x),\zeta)_{(-n)}Y_{\mathcal{E}}(\Phi_\lambda(c),\zeta)\Phi_\lambda(b)\end{aligned}$$

but we know that $(\mathcal{E}_0(\mathcal{F}_\lambda), \mathrm{Id}_{\mathcal{F}_\lambda}, Y_{\mathcal{E}})$ is a vertex algebra, so we eventually get

$$\begin{aligned}\Phi_\lambda(Y(\alpha_{-n}c,\zeta)b) &= Y_{\mathcal{E}}(\alpha(x)_{(-n)}\Phi_\lambda(c),\zeta)\Phi_\lambda(b)\\ &= Y_{\mathcal{E}}(\Phi_\lambda(\alpha_{-n}c),\zeta)\Phi_\lambda(b).\end{aligned}$$

Therefore, $F_{p+1}\mathcal{F}_0 \subset U$. □

8.4 Example: Virasoro vertex algebra

The previous construction of modules over the Heisenberg vertex algebra did not rely on specific properties of the Heisenberg algebra too much. Thus, the same strategy allows us to get modules over other vertex algebras especially when they are *generated* by Lie algebras.

Now, let us consider the Virasoro vertex algebra $(V_c, \mathbf{1}, Y)$ with a fixed central charge $c \in \mathbb{C}$. The Verma module of central charge c and conformal weight $h \in \mathbb{C}$ is given by (see Exercises 1.3)

$$M(c,h) = \mathcal{U}(\mathfrak{vir})/\mathcal{I}_{c,h},$$

where

$$\mathcal{I}_{c,h} = \sum_{n=1}^{\infty} \mathcal{U}(\mathfrak{vir})L_n + \mathcal{U}(\mathfrak{vir})(L_0 - h) + \mathcal{U}(\mathfrak{vir})(C - c).$$

The Verma module has the following PBW basis (Exercise 1.3.2):

$$L_{-n_1} \cdots L_{-n_l} \left|c,h\right\rangle, \quad n_1 \geq \cdots \geq n_l > 0,\, l \in \mathbb{N},$$

where $|c,h\rangle = 1 + \mathcal{I}_{c,h}$.

With a slight abuse of notation, we introduce the following generating series of the action of $\mathfrak{vir}$ on $M(c,h)$:

$$L(x) = \sum_{n\in\mathbb{Z}} L_n x^{-n-2} \in \mathcal{E}(M(c,h)).$$

This series is in fact a field on $M(c,h)$ (Exercise 1.3.3).

Theorem 8.9. *Define a linear map*

$$Y_{M(c,h)}(-,x) \colon V_c \to \operatorname{End}(M(c,h))[[x^{\pm 1}]]$$

by

$$Y_{M(c,h)}(L_{-n_1} \cdots L_{-n_l}\mathbf{1}_c, x) = {}^{\circ}_{\circ}\, \partial^{(n_1-2)}L(x) \cdots \partial^{(n_l-2)}L(x)\, {}^{\circ}_{\circ}$$

for $n_1 \geq \cdots \geq n_l > 0$, $l \in \mathbb{N}$. Then, the pair $(M(c,h), Y_{M(c,h)})$ is a V_c-module.

Proof. Exercises 5.5. □

9 Zhu algebra

In this section, let us make the first step towards studying the module category of a vertex operator algebra. For that, we actually need to narrow down our attention. Recall that a vertex operator algebra V is $\mathbb{Z}$-graded, but we assume that it only has non-negative degrees:

$$V = \bigoplus_{n=0}^{\infty} V_n.$$

We also focus on modules that have a non-negative grading.

In such a setting, we can define an associative algebra $A(V)$ associated to V, called the Zhu algebra of V and see that finite-dimensional simple $A(V)$-modules and simple V-modules are in one-to-one correspondence up to isomorphism. The reference for this section is [Zhu96].

9.1 Modules over VOAs

Let $(V, \mathbf{1}, Y, \omega)$ be a vertex operator algebra. As we have already declared, V has a non-negative grading. Recall that we write $\deg a = n$ if $a \in V_n$. Also, for such a homogeneous element a, the corresponding operators have clear degrees

$$\deg a_{(n)} = \deg a - n - 1, \quad n \in \mathbb{Z}.$$

To define a module over such a VOA, we require that these degrees of operators are transferred to a representation space.

Definition 9.1. Let $(V, \mathbf{1}, Y, \omega)$ be a vertex operator algebra as above. An $\mathbb{N}$-graded module over V is a module (W, Y_W) over V as a vertex algebra such that the following conditions are satisfied:

- the vector space W is $\mathbb{N}$-graded:

$$W = \bigoplus_{n=0}^{\infty} W_n, \quad \dim W_n < \infty, n \in \mathbb{N}, \quad W_0 \neq 0,$$

- with respect to this grading, each operator $a^W_{(n)}$ has the degree

$$\deg a^W_{(n)} = \deg a - n - 1,$$

  for a homogeneous $a \in V$.

If $M = \bigoplus_{n=0}^{\infty} M_n$ and $N = \bigoplus_{n=0}^{\infty} N_n$ are $\mathbb{N}$-graded V-modules, a morphism $f\colon M \to N$ is a linear map with the following properties:

- for all $a \in V$, $w \in M$, $n \in \mathbb{Z}$,

$$f(a^M_{(n)} w) = a^N_{(n)} f(w).$$

- there is $k \in \mathbb{Z}$ such that $f(M_n) \subset N_{n+k}$ for all $n \in \mathbb{Z}$.

The image of a morphism $f\colon M \to N$ is called a **submodule** of N (so the category of $\mathbb{N}$-graded modules is closed under quotient). We say that an $\mathbb{N}$-graded V-module $M (\neq 0)$ is **simple** if its submodule is either 0 or M.

By the Jacobi identity, when we set $L^M_n = \omega^M_{(n+1)}$, $n \in \mathbb{Z}$, they form a representation of the Virasoro algebra on M, but the grading parameter of M has a priori nothing to do with the eigenvalues of L^M_0, and we do not even require that L^M_0 is diagonalizable in contrast to the VOA itself. However, for a simple module, the grading has something to do with the eigenvalues of L^M_0.

Theorem 9.2. *Let $M = \bigoplus_{n=0}^{\infty} M_n$ be a simple $\mathbb{N}$-gradable V-module. Then, there exists $h \in \mathbb{C}$ such that*

$$L_0^M|_{M_n} = (h+n)\mathrm{Id}_{M_n}, \quad n \in \mathbb{Z}.$$

Proof. Exercises 6.2. $\square$

Example 9.3. Let $(\mathcal{F}_0, \mathbf{1}, Y, \omega)$ be the Heisenberg VOA. Recall that the pair $(\mathcal{F}_\lambda, Y_{\mathcal{F}_\lambda})$ for each $\lambda \in \mathbb{C}$ is a module over $\mathcal{F}_0$ as a vertex algebra. For each $n \in \mathbb{N}$, let us set

$$(\mathcal{F}_\lambda)_n = \mathrm{Span}\Big\{\alpha_{-n_1} \cdots \alpha_{-n_l} |\lambda\rangle \,\Big|\, \sum_{i=1}^{l} n_i = n\Big\}.$$

Then, the pair $(\mathcal{F}_\lambda = \bigoplus_{n=0}^{\infty} (\mathcal{F}_\lambda)_n, Y_{\mathcal{F}_\lambda})$ is a simple $\mathbb{N}$-gradable $\mathcal{F}_0$-module.

In fact, from the Jacobi identity, we get

$$[L_0^{\mathcal{F}_\lambda}, a_{(n)}^{\mathcal{F}_\lambda}] = (\deg a - n - 1) a_{(n)}^{\mathcal{F}_\lambda}$$

for all homogeneous $a \in V$ and $n \in \mathbb{Z}$. This proves

$$L_0^{\mathcal{F}_\lambda}|_{(\mathcal{F}_\lambda)_n} = \left(\frac{\lambda^2}{2} + n\right) \mathrm{Id}_{\mathcal{F}_\lambda}, \quad n \in \mathbb{N}$$

and

$$\deg a_{(n)}^{\mathcal{F}_\lambda} = \deg a - n - 1$$

for a homogeneous $a \in V$ and $n \in \mathbb{Z}$.

A submodule $N \subset \mathcal{F}_\lambda$ is preserved by all operators $a_{(n)}^{\mathcal{F}_\lambda}$, $a \in V$, $n \in \mathbb{Z}$, in particular by $\widehat{\mathfrak{h}}$. Since $\mathcal{F}_\lambda$ is an irreducible representation of $\widehat{\mathfrak{h}}$ (Exercise 1.2.3), it must be a simple $\mathcal{F}_0$-module as well.

9.2 Zhu algebra and modules

Now, we are to introduce the Zhu algebra of a VOA. We start with defining the following bilinear maps on the VOA.

Definition 9.4. Define bilinear maps

$$*\colon V \times V \to V, \quad \circ\colon V \times V \to V$$

by

$$a * b = [x^{-1}]\left(Y(a,x)\frac{(x+1)^{\deg a}}{x} b\right),$$
$$a \circ b = [x^{-1}]\left(Y(a,x)\frac{(x+1)^{\deg a}}{x^2} b\right)$$

for homogeneous $a \in V$ and $b \in V$.

Note that $(x+1)^{\deg a}$ is a polynomial in x under our assumption that V only has non-negative degrees. The following theorem is the key for the Zhu algebra.

Theorem 9.5. *Let us set $O(V) = \{a \circ b | a, b \in V\}$ and $A(V) = V/O(V)$. Then, we have the following.*

(1) $(A(V), *)$ is an associative algebra, i.e.,

$$O(V) * V \subset O(V), \quad V * O(V) \subset O(V),$$
$$a * (b * c) - (a * b) * c \in O(V), \quad a, b, c \in V.$$

(2) $[\mathbf{1}] = \mathbf{1} + O(V)$ is the unit of $(A(V), *)$.

(3) $[\omega] = \omega + O(V)$ is a central element of $(A(V), *)$.

We do not prove this theorem here. A proof can be found in [Zhu96]. Of course, the above associative algebra is the Zhu algebra:

Definition 9.6. The associative algebra $(A(V), *)$ is called the Zhu algebra of V.

We will see that there is a correspondence between representations of $A(V)$ and modules over V. One direction is rather straightforward; we can extract a representation of $A(V)$ out of an $\mathbb{N}$-graded V-module.

Theorem 9.7. *Let $M = \bigoplus_{n=0}^{\infty} M_n$ be an $\mathbb{N}$-gradable V-module. For a homogeneous $a \in V$, we write $o(a) = a^M_{(\deg a - 1)}$ and extend the symbol linearly. Then,*

$$V \to \mathrm{End}(M_0); \quad a \mapsto o(a)$$

induces an action of $A(V)$ on M_0, i.e., $o(a)|_{M_0} = 0$ for all $a \in O(V)$ and $o(a * b)|_{M_0} = o(a)o(b)|_{M_0}$ for $a, b \in V$.

The other direction is more complicated, but still we have the following.

Theorem 9.8. *Let W be a finite dimensional $A(V)$-module. Then, there exists an $\mathbb{N}$-gradable V-module $M = \bigoplus_{n=0}^{\infty} M_n$ with the following properties.*

(1) $M_0 \simeq W$ as $A(V)$-modules.

(2) M is generated by the top space M_0.

(3) If $N \subset M$ is a submodule such that $N \cap M_0 = 0$, then $N = 0$.

We will sketch the proof of this theorem in Section 9.5. Now, combining Theorems 9.7 and 9.8, we get the following.

Theorem 9.9. *Theorems 9.7 and 9.8 induce a one-to-one correspondence between finite-dimensional simple $A(V)$-modules and simple $\mathbb{N}$-gradable V-modules.*

Sketch of Proof. Let W be a finite-dimensional $A(V)$-module. If we apply Theorem 9.8, we get an $\mathbb{N}$-gradable V-module $M = \bigoplus_{n=0}^{\infty} M_n$ with $M_0 \simeq W$ as $A(V)$-modules. Furthermore, if W is simple, then so is M.

Conversely, let us start with an $\mathbb{N}$-gradable V-module $M = \bigoplus_{n=0}^{\infty} M_n$. If we apply Theorem 9.8 to M_0, we get an $\mathbb{N}$-gradable V-module $\widetilde{M} = \bigoplus_{n=0}^{\infty} \widetilde{M}_n$ with $\widetilde{M}_0 \simeq M_0$ as $A(V)$-modules, but we cannot immediately say that $\widetilde{M} \simeq M$. However, if M is simple, then M_0 is simple as a $A(V)$-module and $\widetilde{M}$ is also a simple module. Furthermore, we can construct a non-zero morphism $\widetilde{M} \to M$, but since they are irreducible, we must have $\widetilde{M} \simeq M$. □

9.3 Example: Heisenberg VOA

Theorem 9.10. *We have an isomorphism of associative algebras*

$$\mathbb{C}[\alpha] \xrightarrow{\sim} A(\mathcal{F}_0); \quad \alpha \mapsto [\alpha_{-1}\mathbf{1}].$$

Proof. See [FZ92]. (The Heisenberg algebra is an affine Lie algebra.) □

Since $\mathbb{C}[\alpha]$ is commutative, finite-dimentional simple $\mathbb{C}[\alpha]$-modules are all one-dimensional. They are labelled by $\lambda \in \mathbb{C}$: for each $\lambda \in \mathbb{C}$, the one-dimensional space $\mathbb{C}_\lambda = \mathbb{C}$ is equipped with the action of $\mathbb{C}[\alpha]$,

$$\alpha \mapsto \lambda \cdot \mathrm{Id}.$$

On the other hand, $[\alpha_{-1}\mathbf{1}] \in A(\mathcal{F}_0)$ acts on $(\mathcal{F}_\lambda)_0 = \mathbb{C}\,|\lambda\rangle$ by $\lambda{\cdot}\mathrm{Id}$. Therefore, $\mathbb{C}_\lambda$ and $\mathcal{F}_\lambda$ correspond to each other under Theorem 9.9.

9.4 Example: Virasoro VOA

9.4.1 Universal Virasoro VOA

Recall that the Virasoro VOA of central charge c is built on

$$V_c = \mathcal{U}(\mathfrak{vir})/\Big(\sum_{n\geq -1} \mathcal{U}(\mathfrak{vir})L_n + \mathcal{U}(\mathfrak{vir})(C-c)\Big)$$

and the Verma modules

$$M(c,h) = \mathcal{U}(\mathfrak{vir})/\Big(\sum_{n>0} \mathcal{U}(\mathfrak{vir})L_n + \mathcal{U}(\mathfrak{vir})(L_0-h) + \mathcal{U}(\mathfrak{vir})(C-c)\Big)$$

are modules over V_c as a vertex algebra. It is not difficult to see that they are $\mathbb{N}$-gradable V_c-modules, but are not necessarily simple (in contrast to the Heisenberg case). In fact, for a certain choice of (c,h), the Verma module $M(c,h)$ is reducible as a representation of the Virasoro algebra, hence cannot be simple as V_c-module. (The V_c action goes through the action of the Virasoro algebra.) Nevertheless, we can always take the simple quotient of $M(c,h)$ and write it as $L(c,h)$.

Theorem 9.11. *There is an isomorphism of associative algebras*

$$\mathbb{C}[\mathsf{h}] \xrightarrow{\sim} A(V_c); \quad \mathsf{h} \mapsto [\omega].$$

Proof. See [FZ92] or [Wan93]. □

Exactly by the same reasoning as the Heisenberg case, we can conclude that $L(c, h)$, $h \in \mathbb{C}$ are all of the simple $\mathbb{N}$-gradable V_c-modules.

9.4.2 Minimal Virasoro VOA

The representation V_c of the Virasoro algebra is not always irreducible, or in other words, it is not a simple VOA. It is known from the representation theory of the Virasoro algebra, when we set

$$c = c_{p,q} = 1 - 6\frac{(p-q)^2}{pq}, \quad p, q \in \{2, 3, \dots\} : \text{coprime},$$

V_c is reducible, and its maximal proper submodule is generated by a singular vector $v_{p,q} \in (V_c)_{(p-1)(q-1)}$. For this fact, see [IK10]. Let us write L_c for the simple quotient. As the formula of $c_{p,q}$ is symmetric under exchanging p and q, we may assume $p < q$. If $(p, q) = (2, 3)$, $v_{p,q}$ lives in $(V_c)_2$, which is spanned by $L_{-2}\mathbf{1}_c$ and L_c is the trivial representation. Usually, we exclude this case. The submodule generated by $v_{p,q}$ is also an ideal of the VOA V_c, hence L_c is still a VOA. The simple VOA L_c with some $c = c_{p,q}$ is called a minimal Virasoro VOA.

From the general theory of the Zhu algebra, the Zhu algebra $A(L_c)$ is the following quotient of $A(V_c)$:

$$A(V_c)/([v_{p,q}]).$$

It can be shown [Wan93] that, under the identification $\mathbb{C}[\mathsf{h}] \simeq A(V_c)$, we can take a polynomial $G_{p,q}(\mathsf{h})$ of degree $\frac{1}{2}(p-1)(q-1)$ as a generator of the ideal.

Theorem 9.12 ([Wan93])**.** *The following polynomial works as a generator of the ideal:*

$$G_{p,q}(\mathsf{h}) = \Big(\prod_{r=1}^{p-1} \prod_{s=1}^{q-1} (\mathsf{h} - h_{r,s}) \Big)^{1/2},$$

where

$$h_{r,s} = \frac{(sp - rq)^2 - (p-q)^2}{4pq}, \quad r, s \in \mathbb{Z}.$$

Although the formula for $G_{p,q}(\mathsf{h})$ involves a square root, we can check that it is a polynomial of h (Exercise 6.4.1).

Corollary 9.13. *The minimal Virasoro VOA $L_{c_{p,q}}$ has $\frac{1}{2}(p-1)(q-1)$ number of simple $\mathbb{N}$-gradable modules, and they are covered by*

$$L(c_{p,q}, h_{r,s}), \quad r = 1, \dots, p-1,\ 1 \leq s < \frac{q}{p}.$$

Example 9.14 (Ising model: $(p,q) = (3,4)$). The central charge is $c = \frac{1}{2}$. The VOA L_c itself is a simple $\mathbb{N}$-gradable module, so we of course have

$$h_{1,1} = h_{2,3} = 0.$$

The other non-trivial modules have the conformal weights

$$h_{2,1} = h_{1,3} = \frac{1}{2} \quad \text{and} \quad h_{2,2} = h_{1,2} = \frac{1}{16}.$$

Consequently, the polynomial $G_{3,4}(\mathsf{h})$ is

$$G_{3,4}(\mathsf{h}) = \mathsf{h}\Big(\mathsf{h} - \frac{1}{2}\Big)\Big(\mathsf{h} - \frac{1}{16}\Big).$$

9.5 Proof of Theorem 9.8 (Sketch)

It is instructive to see the following for an $\mathbb{N}$-gradable V-module. Let (M, Y_M) be an $\mathbb{N}$-gradable V-module. We suppress M from the notation if there is no confusion. According to the direct sum $M = \bigoplus_{n=0}^{\infty} M_n$, a functional $\varphi \in M_0^*$ is naturally extended to the whole M.

Lemma 9.15. *For any $a^1, \dots, a^n \in V$, $w \in M_0$, and $\varphi \in M_0^*$, we have*

$$\begin{aligned} &\big\langle \varphi, Y(a^1, x_1) Y(a^2, x_2) \dots Y(a^n, x_n) w \big\rangle \\ =& \big\langle o(a^1)^* \varphi, Y(a^2, x_2) \dots Y(a^n, x_n) w \big\rangle \\ &+ \sum_{k=2}^{n} \sum_{i=0}^{\infty} \iota_{1k} F_{\deg a^1, i}(x_1, x_k) \cdot \Big\langle \varphi, Y(a^2, x_2) \cdots Y(a^1_{(i)} a^k, x_k) \cdots Y(a^n, x_n) w \Big\rangle, \end{aligned}$$

where

$$F_{n,i}(x,y) = x^{-n} \partial_y^{(i)} \frac{y^n}{x-y} \in \mathbb{C}[x,y][x^{-1}, y^{-1}, (x-y)^{-1}]$$

for $n, i \in \mathbb{N}$.

Proof. Exercises 6.5. □

Notice that, in this formula, the right-most M_0 and left-most M_0^* are preserved, but the number of Y inserted in the matrix element is reduced by one. Therefore, we can recover the matrix element

$$\big\langle \varphi, Y(a^1, x_1) Y(a^2, x_2) \dots Y(a^n, x_n) w \big\rangle$$

only from the knowledge of the top space M_0.

The proof of Theorem 9.8 goes as follows.

Step 1: We first construct functionals

$$S\colon W^* \otimes V^{\otimes n} \otimes W \to \mathbb{C}[x_1, \ldots, x_n][x_i^{-1}, (x_i - x_j)^{-1}]$$

that pretend matrix elements by recursion in n.

For $n = 0$, there is a natural pairing

$$S = \langle -, - \rangle : W^* \otimes W \to \mathbb{C}.$$

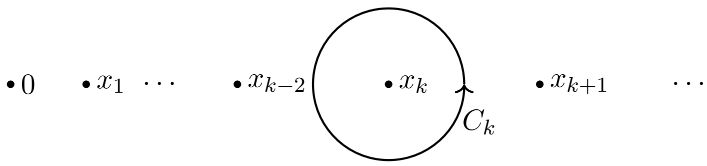
Assuming that S is defined up to $n-1$, we define, for $a^1, \ldots, a^n \in V$, $w \in W$ and $\varphi \in W^*$,

$$\begin{aligned}
&S(\varphi, (a^1, x_1)(a^2, x_2) \cdots (a^n, x_n)w) \\
=\,& S(o(a^1)^* \varphi, (a^2, x_2) \ldots (a^n, x_n)w) \\
+\,& \sum_{k=2}^{n} \sum_{i=0}^{\infty} F_{\deg a^1, i}(x_1, x_k) \cdot S(\varphi, (a^2, x_2) \cdots (a^1_{(i)} a^k, x_k) \cdots (a^n, x_n)w).
\end{aligned}$$

Step 2: It can be shown that the functions S satisfy the following property:

$$\begin{aligned}
&\int_{C_{k+1}} S(\varphi, \cdots (a^{k-1}, x_{k-1})(a^k, x_k) \cdots w)(x_{k-1} - x_k)^n dx_{k-1} \\
=\,& S(\varphi, \cdots (a^{k-1}_{(n)} a^k, x_k) \cdots w) \qquad (9.1)
\end{aligned}$$

for any k. Here C_k is an integral contour that encloses only x_k, but not others nor 0.



We omitted the numerical factor $\frac{1}{2\pi \mathbb{i}}$ from the integral.

Step 3: We define $\overline{M}$ by the formal span

$$\overline{M} = \mathrm{Span}\left\{ b^1_{(i_1)} \cdots b^l_{(i_l)} w \,\middle|\, b^1, \ldots, b^l \in V,\ i_1, \ldots, i_l \in \mathbb{Z},\ w \in W \right\}$$

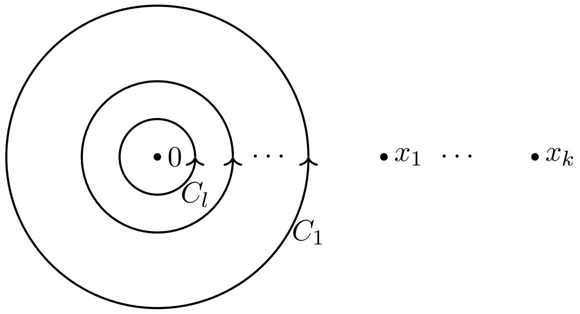
and extend S to $W^* \otimes V^{\otimes n} \otimes \overline{M}$ as follows: for

$$m = b^1_{(i_1)} \cdots b^l_{(i_l)} w,$$

we set

$$\begin{aligned}
&S(\varphi, (a^1, x_1) \cdots (a^k, x_k)m) \\
=& \int_{C_1} \cdots \int_{C_l} S(\varphi, (a^1, x_1) \cdots (a^k, x_k)(b^1, y_1) \cdots (b^l, y_l)w) y_1^{i_i} \cdots y_l^{i_l} dy_1 \cdots dy_l.
\end{aligned}$$

Here, each y_i is integrated along C_i that encloses 0 and all C_j of $j > i$. The points $x_1, \dots, x_k$ are outside these contours.



It is clear that (9.1) still holds even if we replace $w \in W$ with $m \in \overline{M}$:

$$\begin{aligned}&\int_{C_{k+1}} S(\varphi, \cdots (a^{k-1}, x_{k-1})(a^k, x_k) \cdots m)(x_{k-1} - x_k)^n dx_{k-1} \\ &= S(\varphi, \cdots (a^{k-1}_{(n)} a^k, x_k) \cdots m).\end{aligned} \tag{9.2}$$

Step 4: We define

$$\mathrm{Rad}(\overline{M}) = \left\{ m \in \overline{M} \,\middle|\, S(\varphi, (a^1, x_1) \cdots (a^n, x_n) m) = 0, \ {\scriptstyle \text{for all } \varphi \in W^*, \atop a^1, \dots, a^n \in V,\, n \in \mathbb{N}} \right\}$$

and

$$M = \overline{M} / \mathrm{Rad}(\overline{M}).$$

It is clear that $\mathrm{Rad}(\overline{M})$ is stable under adding a symbol $a_{(n)}$, $a \in V$, $n \in \mathbb{Z}$ on the left. Therefore, if we define

$$Y_M(a, x) = \sum_{n \in \mathbb{Z}} a_{(n)} x^{-n-1}$$

for $a \in V$, it is considered in $\mathrm{End}(M)[[x^{\pm 1}]]$.

The claim is, of course, that the pair (M, Y_M) is the desired $\mathbb{N}$-gradable V-module. Proving that requires few extra works, but most importantly, (9.2) implements the property

$$Y_M(a_{(n)} b, x) = Y_M(a, x)_{(n)} Y_M(b, x), \quad a, b \in V,\, n \in \mathbb{Z}$$

on M. (Recall Proposition 8.2.)

10 Monster moonshine

In this section, we overview the application of VOA to the monster moonshine. References are [Mas14, Gan07].

10.1 Monster

A group $G(\neq \{e\})$ is said to be simple, if its normal subgroup is either $\{e\}$ or G. Finite simple groups are completely classified and they are covered by

- cyclic groups $\mathbb{Z}_p$ with prime p,
- alternating groups $\mathcal{A}_n$ with $n \geq 5$,
- Lie type ($\mathrm{PSL}_n(\mathbb{F}_q)$, ...),
- 26 sporadic groups.

Among the 26 sporadic groups, the biggest one $\mathbb{M}$ is called the monster. Its degree is $|\mathbb{M}| \simeq 8 \times 10^{53}$ and it has 194 conjugacy classes. Therefore, it has 194 number of irreducible representations $\{\rho_i\}_{i=0}^{193}$, which we number in the increasing order in dimensions. The list, of course, starts from the trivial representation, but the next one already has a very big dimension:

$$\begin{aligned}
\dim \rho_0 &= 1, \\
\dim \rho_1 &= 196883, \\
\dim \rho_2 &= 21296876, \\
&\vdots
\end{aligned}$$

10.2 j-function

The complex upper-half plane is

$$\mathbb{H} = \{\tau \in \mathbb{C} | \mathrm{Im}\tau > 0\},$$

on which $\mathrm{SL}_2(\mathbb{R})$ acts by the Möbius transformation:

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

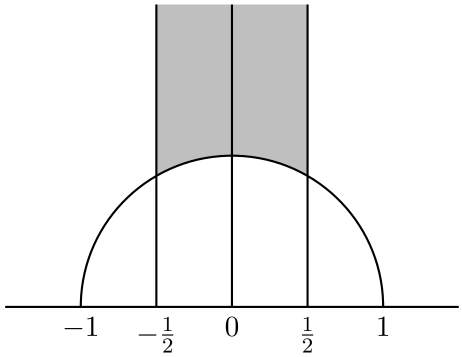
As a subgroup, we have $\mathrm{SL}_2(\mathbb{Z})$ and set

$$\Sigma = \mathrm{SL}_2(\mathbb{Z}) \backslash \overline{\mathbb{H}}$$

with

$$\overline{\mathbb{H}} = \mathbb{H} \cup \mathbb{Q} \cup \mathbb{i}\{\infty\}.$$

The most prominent choice of a fundamental domain of this action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$ is the following one:



A modular function for $\mathrm{SL}_2(\mathbb{Z})$ is a meromorphic function on Σ, in other words, it is a meromorphic function

$$f\colon \overline{\mathbb{H}} \to \mathbb{C}$$

such that

$$f(A\tau) = f(\tau), \quad A \in \mathrm{SL}_2(\mathbb{Z}),\ \tau \in \overline{\mathbb{H}}.$$

We can actually show that $\Sigma \simeq \mathbb{P}^1$. Therefore, once we find a function $J\colon \Sigma \xrightarrow{\sim} \mathbb{P}^1$, any meromorphic function on Σ has the form

$$\frac{P(J(\tau))}{Q(J(\tau))}, \quad P, Q \in \mathbb{C}[w].$$

There is a historical choice of such J called the j-function:

$$j(\tau) = 1728 \frac{20G_4(\tau)^3}{20G_4(\tau)^3 - 49G_6(\tau)^2}, \tag{10.1}$$

where $G_k(\tau)$ is the k-th Eisenstein series

$$G_k(\tau) = \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} (m\tau + n)^{-k}.$$

Eta and theta

Dedekind's eta function is given by

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n), \quad q = e^{2\pi i\tau}.$$

It is common to write the 24th power of η as

$$\Delta(\tau) = \eta(\tau)^{24}.$$

Let L be an even lattice, i.e., it is a finite-rank free $\mathbb{Z}$-module:

$$L = \sum_{i=1}^{l} \mathbb{Z}\alpha_i$$

along with a symmetric bilinear form

$$(-,-)\colon L \times L \to \mathbb{Z}, \quad (\alpha, \alpha) \in 2\mathbb{Z}_{>0},\ \alpha \neq 0.$$

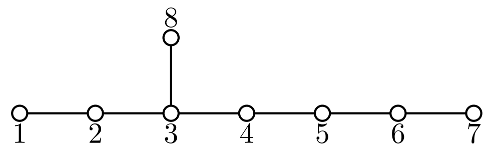
The theta function of the lattice L is defined by

$$\Theta_L(\tau) = \sum_{\alpha \in L} q^{(\alpha,\alpha)/2}, \quad q = e^{2\pi \mathrm{i} \tau}.$$

Let us take the E_8-lattice. As a $\mathbb{Z}$-module, it is

$$L_{E_8} = \sum_{i=1}^{8} \mathbb{Z}\alpha_i.$$

The bilinear form $(-,-)$ on L_{E_8} can be recovered from the E_8-Dynkin diagram:



as

$$(\alpha_i, \alpha_j) = \begin{cases} 2, & i = j, \\ -1, & (i,j) \text{ are connected}, \\ 0, & \text{otherwise} \end{cases}$$

The E_8-lattice is, of course, the root lattice of the simple Lie algebra E_8, but also gives an optimal S^7-packing in $\mathbb{R}^8$.

It is kwnon that, using these eta and theta functions, the j-function is written as

$$j(\tau) = \frac{\Theta_{L_{E_8}}(\tau)^3}{\Delta(\tau)}. \tag{10.2}$$

10.3 Monster moonshine

The expression (10.2) allows one to expand the j-function in $q = e^{2\pi \mathrm{i} \tau}$ (while the modular invariance is clear in (10.1)). As a result, we get

$$j(\tau) = q^{-1} + 744 + 196844q + 21493760q^2 + \cdots.$$

We can observe that

$$
\begin{aligned}
[q^{-1}]j(\tau) &= 1 = \dim \rho_0, \\
[q^{1}]j(\tau) &= 194844 = \dim \rho_0 + \dim \rho_1, \\
[q^{2}]j(\tau) &= 21493760 = \dim \rho_0 + \dim \rho_1 + \dim \rho_2.
\end{aligned}
$$

The coefficients of higher order terms are also written by combining *small* number of $\dim \rho_i$, $i = 0, 1, \ldots, 193$. Of course, we can express any natural number n as $n = n \cdot \dim \rho_0$, but the number of $\dim \rho_i$'s we need to use is surprisingly small.

The conjecture by Conway–Norton proved by Borcherds goes as follows:

Conjecture/Theorem 10.1 (Conway–Norton, Borcherds)**.** *(1) There is a graded vector space*

$$
V = \bigoplus_{n=0}^{\infty} V_n, \quad (\dim V_n < \infty)
$$

with a "natural algebraic structure" and $\mathrm{Aut}(V) = \mathbb{M}$*.*
(2) The moster $\mathbb{M}$ *acts on each* V_n*. For each* $g \in \mathbb{M}$*,*

$$
J_g(\tau) = \sum_{n=0}^{\infty} \mathrm{Tr}_{V_n}(g) q^{n-1}
$$

is a Hauptmodul of genus 0. Furthermore,

$$
J_e(\tau) = J(\tau) := j(\tau) - 744.
$$

10.4 Character of a VOA

Let $V = \bigoplus_{n=0}^{\infty} V_n$ be a VOA of central charge c. The character of V is defined by

$$
\mathrm{ch}[V](\tau) = q^{-c/24} \sum_{n=0}^{\infty} \dim(V_n) q^n, \quad q = e^{2\pi \mathbb{i} \tau}.
$$

10.5 Lattice VOA

We can construct a VOA out of a lattice. Let L be a lattice of rank l with a symmetric bilinear form $(-,-)\colon L \times L \to \mathbb{Z}$. We define the associated Heisenberg algebra by

$$
\widehat{\mathfrak{h}}_L = L \otimes_{\mathbb{Z}} \mathbb{C}[t^{\pm 1}] \oplus \mathbb{C}K
$$

along with the Lie bracket

$$
\begin{aligned}
[\alpha_m, \beta_n] &= m(\alpha, \beta) \delta_{m+n,0} K, \quad m, n \in \mathbb{Z}, \\
[K, \widehat{\mathfrak{h}}_L] &= 0,
\end{aligned}
$$

where we wrote $\alpha_n = \alpha \otimes t^n$, $\alpha \in L$, $n \in \mathbb{Z}$. Then the vacuum Fock representation of $\widehat{\mathfrak{h}}_L$ is

$$\mathcal{F}_0^L = \mathcal{U}(\widehat{\mathfrak{h}}_L)/\big(\mathcal{U}(\widehat{\mathfrak{h}}_L) \cdot (L \otimes \mathbb{C}[t]) + \mathcal{U}(\widehat{\mathfrak{h}}) \cdot (K-1)\big).$$

We set

$$V_L = \mathcal{F}_0^L \otimes \mathbb{C}_\epsilon[L],$$

where $\mathbb{C}_\epsilon[L]$ is the group algebra of L twisted by a $\mathbb{C}^\times$-valued 2-cocycle ϵ. It is convenient (and common) to realize the group algebra as

$$\mathbb{C}_\epsilon[L] = \bigoplus_{\alpha \in L} \mathbb{C} e^\alpha$$

introducing the symbols e^α, $\alpha \in L$. Then, the multiplication rule reads

$$e^\alpha e^\beta = \epsilon(\alpha, \beta) e^{\alpha+\beta}.$$

The Heisenberg algebra $\widehat{\mathfrak{h}}_L$ acts on V_L mostly naturally except α_0:

$$\begin{aligned} \alpha_n \cdot (v \otimes e^\beta) &= (\alpha_n v) \otimes e^\beta, \quad n \neq 0, \\ \alpha_0 \cdot (v \otimes e^\beta) &= (\alpha, \beta) v \otimes e^\beta. \end{aligned}$$

There is a unique cohomology class $[\epsilon] \in H^2(L, \mathbb{C}^\times)$ such that V_L is equipped with the structure of a VOA in the following way. It is generated by

$$\begin{aligned} Y(\alpha_{-1}\mathbf{1} \otimes e^0, x) &= \sum_{n \in \mathbb{Z}} \alpha_n x^{-n-1}, \\ Y(\mathbf{1} \otimes e^\alpha, x) &= e^\alpha x^{\alpha_0} \exp\Big(-\sum_{j<0} \frac{\alpha_j}{j} x^{-j}\Big) \exp\Big(-\sum_{j>0} \frac{\alpha_j}{j} x^{-j}\Big) \end{aligned}$$

with $\alpha \in L$. The conformal vector is given by

$$\omega = \frac{1}{2} \sum_{i=1}^{l} (\alpha_i)_{-1} (\alpha_i^\vee)_{-1} \mathbf{1} \otimes e^0,$$

where $\{\alpha_i\}$ and $\{\alpha_i^\vee\}$ are bases of $L \otimes_{\mathbb{Z}} \mathbb{C}$ that are dual to each other with respect to $(-,-)$. Then the central charge is $c = l$.

It is clear from the construction that the character of V_L becomes

$$\mathrm{ch}[V_L](\tau) = \frac{\Theta_L(\tau)}{\eta(\tau)^l}.$$

10.6 Automorphism

Let $(V, \mathbf{1}, Y, \omega)$ be a vertex operator algebra. We say that $g \in \mathrm{GL}(V)$ is an automorphism of the VOA if

$$gY(a,x)g^{-1} = Y(ga,x), \quad a \in V,$$
$$g\omega = \omega.$$

Given an automorphism $g \in \mathrm{Aut}(V)$, we can consider two things:

- The fixed point set V^g is again a VOA, called an orbifold VOA,
- There is a notion of a g-twisted V-module. While a g-twisted V-module is not a genuine V-module, it is a honest module over the orbifold VOA.

In particular, an automorphism of a lattice L induces a VOA automorphism of the lattice VOA. For example, we always have the automorphism $\theta \in \mathrm{Aut}(V_L)$ that comes from

$$L \to L; \quad \alpha \mapsto -\alpha.$$

10.7 Moonshine module $V^\natural$

We already know that the j-function has the expression (10.2), where the E_8-lattice is manifest, but we would rather take another lattice, namely, the Leech lattice Λ instead. We omit a precise definition of the Leech lattice, but it is a lattice of rank 24 without a lattice point α such that $(\alpha, \alpha) = 2$, and gives an optimal S^{23}-packing in $\mathbb{R}^{24}$.

The character of V_Λ is immediately

$$\mathrm{ch}[V_\Lambda](\tau) = \frac{\Theta_\Lambda(\tau)}{\Delta(\tau)} = J(\tau) + 24.$$

It is known that there exists a unique θ-twisted V_Λ-module $V_\Lambda(\theta)$. As a vector space, is looks like

$$V_\Lambda(\theta) = S(L \otimes_{\mathbb{Z}} \mathbb{C}[t^{-1/2}]) \otimes U,$$

where the first tensor component is the symmetric algebra over $L \otimes_{\mathbb{Z}} \mathbb{C}[t^{-1/2}]$, and the second one a is 2^{12}-dimensional vector space. On the twisted module $V_\Lambda(\theta)$, θ is acting naturally on the first component and by (-1) to the second component.

The so-called moonshine module $V^\natural$ is given by

$$V^\natural = V_\Lambda^\theta \oplus V_\Lambda(\theta)^\theta.$$

The remaining tasks are

(1) $V^\natural$ is a VOA.

(2) $V_1^\natural = 0$.

(3) $\mathrm{ch}[V^\natural](\tau) = q^{-1} + \cdots$ is modular. (Either directly, or due to Zhu's modularity theorem [Zhu96].) Together with (2), we get

$$\mathrm{ch}[V^\natural](\tau) = J(\tau).$$

(4) (Hard) $\mathrm{Aut}(V^\natural) = \mathbb{M}$.

References

[FBZ04] Edward Frenkel and David Ben-Zvi. *Vertex algebras and algebraic curves*. American Mathematical Soc., 2004.

[FZ92] Igor B Frenkel and Yongchang Zhu. Vertex operator algebras associated to representations of affine and virasoro algebras. *Duke Math. J*, 66:123–168, 1992.

[Gan07] Terry Gannon. *Moonshine beyond the Monster: The bridge connecting algebra, modular forms and physics*. Cambridge University Press, 2007.

[Hum12] James E Humphreys. *Introduction to Lie algebras and representation theory*. Springer Science & Business Media, 2012.

[IK10] Kenji Iohara and Yoshiyuki Koga. *Representation theory of the Virasoro algebra*. Springer Science & Business Media, 2010.

[Kac98] Victor G Kac. *Vertex algebras for beginners*. American Mathematical Soc., 1998.

[LL04] James Lepowsky and Haisheng Li. *Introduction to vertex operator algebras and their representations*. Springer Science & Business Media, 2004.

[Mas14] Geoffrey Mason. Vertex operator algebras, modular forms and moonshine. In *Conformal Field Theory, Automorphic Forms and Related Topics: CFT, Heidelberg, September 19-23, 2011*, pages 335–365. Springer, 2014.

[RTT22] Christopher Raymond, Yoh Tanimoto, and James E Tener. Unitary vertex algebras and wightman conformal field theories. *Communications in Mathematical Physics*, 395(1):299–330, 2022.

[Sch08] Martin Schottenloher. *A mathematical introduction to conformal field theory*, volume 66. Springer, 2008.

[Wan93] Weiqiang Wang. Rationality of virasoro vertex operator algebras. *International Mathematics Research Notices*, 1993(7):197–211, 1993.

[Zhu96] Yongchang Zhu. Modular invariance of characters of vertex operator algebras. *Journal of the American Mathematical Society*, 9(1):237–302, 1996.