Conformal symmetry and the Virasoro algebra

THESIS

submitted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE
in
PHYSICS

Author :	B.W. van Kootwijk
Student ID :	s1852337
Supervisors :	Prof.dr. K.E. Schalm,
	Dr. R.I. van der Veen

Leiden, The Netherlands, July 11, 2020

Conformal symmetry and the Virasoro algebra

B.W. van Kootwijk

Huygens-Kamerlingh Onnes Laboratory, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

July 11, 2020

Abstract

In this thesis we study conformal invariance of two-dimensional physical systems, and its connections to the Virasoro algebra. We first discuss how symmetries relate to Lie algebras, and how the Virasoro algebra corresponds to conformal symmetries. We then study the theory of the free open bosonic string. Starting with the Nambu-Goto action, we first solve the equations of motion of the classical string using the light-cone gauge. From there, we construct the quantum theory and show it is indeed a representation of the Virasoro algebra. We then look more directly into representations of the Virasoro algebra. We define lowest weight representations and Verma modules, and discuss reducibility of the latter.

Contents

Version of July 11, 2020– Created July 11, 2020 - 11:00

Introduction

Many algebraic structures in theoretical physics arise through symmetries of physical systems. Symmetry is a fundamental concept in modern physics, and it is no wonder Nobel laureate P. W. Anderson once said: *"it is only slightly overstating the case to say that physics is the study of symmetry"* [And72]. The essential role of symmetry becomes clear via Noether's theorem, which states that families of spacetime symmetries correspond to conserved quantities. For example, invariance under translations induces conservation of momentum, and similarly rotational symmetry results in conserved angular momentum. In these two cases, the set of symmetries forms an algebraic group under composition, since both translations and rotations are invertible. However, one often encounters symmetries which do not follow this pattern: these symmetries may not be invertible or may only be defined locally. A set of such symmetries does not generally form a group, yet one can still study them by looking at their infinitesimal generators. These generators naturally form a *Lie algebra*, a vector space endowed with a certain antisymmetric bilinear operation.

Apart from translating or rotating the target space, another natural transformation is scaling. Scale invariance might at first glance seem an unlikely property of any physical system, as physical systems usually have some defined notions of size and mass. However, certain theories such as the Ising model do, under certain conditions, show scale invariance. A generalisation of proper scaling transformations are transformations which infinitesimally only scale, but globally may exhibit more involved behaviour. Such transformations are exactly those transformations that leave angles invariant, which are referred to as *conformal transformations*. Generally we may take the target space to be $\mathbb{R}^n$ for some $n \in \mathbb{Z}_{\geq 2}$. For $n > 2$, any invertible conformal transformation which is only locally de-

fined, i.e. it is defined on some open subset of $\mathbb{R}^n$, can be extended to a globally defined invertible conformal transformation [Sch08]. This results in a *finite* number infinitesimal generators. In the two-dimensional case, however, such extensions are not possible in general, and one obtains an *infinite* set of generators. Conformal invariance in two dimensions is therefore very restrictive compared to higher dimensional cases, which makes it easier to determine the dynamics of conformally invariant systems in two dimensions. In this thesis, we solely focus on two-dimensional conformal theories and the Lie algebra structures they contain, namely the Witt and Virasoro Lie algebras. In particular, we investigate the algebraic structure of the simplest of such theories which is given by *lowest weight representations* of the Virasoro algebra. First, however, we will discuss one particular physical example of a conformal theory, given by bosonic string theory.

String theory is often advocated as a *theory of everything*. In this theory, the fundamental building blocks of nature are thought to be strings instead of elementary particles. While the theory was originally meant to merely describe the strong interaction, it was later found to contain particles with the properties of the photon and the hypothetical graviton, the mediator of gravity [Ric16]. This led to string theory being considered in a much broader context, namely as a theory in which both quantum mechanics and Einsteins theory of relativity are united. Although a multitude of string theories has risen over the years, for our purposes it suffices to study the simpler bosonic string theory, and leave out the supersymmetrical counterparts involving fermions. The strings in string theory are micoroscopic one-dimensional objects which trace out a two-dimensional surface in spacetime as they propagate through time. This surface is called the *world sheet*. Conformal transformations of the world sheet leave the dynamics of string invariant, which indeed makes string theory a conformal field theory in two dimensions. In the quantum theory of the string one therefore obtains a representation of the Virasoro Lie algebra, as we will show for the open bosonic string.

In chapter 1 we cover the necessary groundwork on which the following chapters are build. This includes a definition of metrics, the framework of Lagrangian mechanics and a short introduction into quantum mechanics. In chapter 2 we develop the basics of the theory of Lie algebras and their representations, and investigate the Lie algebras corresponding to conformal transformations. In chapter 3 we discuss the open bosonic string and formulate a quantum theory. We finish in chapter 4 by studying elementary representations of the Virasoro algebra, namely lowest weight representations.

Chapter 1

Preliminaries

In this chapter we give a short review of general concepts that will be used and discussed within this thesis. These concepts include metrics on vector spaces, Lagrangian mechanics and quantum theory.

1.1 Metrics

Throughout this thesis we will often encounter notions of length, area and angles on different sets. Although the reader should be familiar with these concepts, they might be surprised as we use a slightly generalised notion of length and area, which might result in negative values. We will therefore introduce precise definitions for these concepts.

1.1.1 Distances

Metrics are defined through inner products, for which we use a more general definition than is conventional.

Definition 1.1. *Let V be a real vector space. An **inner product** on V is a map $V \times V \to \mathbb{R}$, $(v, w) \mapsto v \cdot w$ which is bilinear, symmetric, and admits a basis B of V such that for all $b, c \in B$ we have*

$$\langle b, c \rangle = \begin{cases} 0 & \text{if } b \neq c, \\ \pm 1 & \text{if } b = c. \end{cases} \tag{1.1}$$

For finite dimensional vector spaces this last property simple means the inner product is non-degenerate, since it directly implies no nonzero vec-

tor is orthogonal to every basis vector. One should note we explicitly do not require an inner product to be positive definite, which is usually included in the definition. As a consequence, the inner product of a vector with itself might be zero or negative. The advantage of this is that if we want to view our vector space as spacetime, we can distinguish between time and space directions. Any vector whose inner product with itself is negative can be thought of as pointing in a timelike direction. In line with this idea, a vector $v \in V$ with $v^2 > 0$, $v^2 = 0$, $v^2 < 0$ will be called *spacelike*, *lightlike* or *timelike* respectively. We can now use inner products to define metrics.

Definition 1.2. *Let $n \in \mathbb{Z}_{\geq 1}$, let M be an n-dimensional real smooth manifold, and for each $x \in M$ let T_xM be its tangent space. A **metric** on M is a map $g : M \to \mathrm{Bil}(T_xM \times T_xM, \mathbb{R})$ such that for all $x \in M$ the map $g(x)$ is a real inner product and such that for any smooth curve $\gamma : (-1, 1) \to M$ the map $t \mapsto g(\gamma(t))(\gamma'(t), \gamma'(t))$ is smooth.*

A manifold with such a metric is called a pseudo-Riemannian manifold. Given two vectors v, w in some tangent space T_xM we will often shorten notation and write $v \cdot w$ instead of $g(x)(v, w)$, if it is clear what is meant by this. Accordingly, v^2 means $g(x)(v, v)$. A simple but very important example of a pseudo-Riemannian manifold is Minkowski space.

Definition 1.3. *For $d \in \mathbb{N}$, the $\mathbf{d+1}$**-dimensional Minkowski space** is the vector space $\mathbb{R}^{(1,d)} := \mathbb{R}^1 \times \mathbb{R}^d$. We write its elements as $x = (x^0, \vec{x})$ for $x^0 \in \mathbb{R}^1$ and $\vec{x} = (x^1, \ldots, x^d) \in \mathbb{R}^d$. This vector space is endowed with the dot product $\mathbb{R}^{(1,d)} \times \mathbb{R}^{(1,d)} \to \mathbb{R}$ given by*

$$x \cdot y = (x^0, \vec{x}) \cdot (y^0, \vec{y}) := -x^0 y^0 + \sum_{i=1}^{d} x^i y^i. \tag{1.2}$$

Here we identified $\mathbb{R}^{(1,d)}$ with its own tangent space.

A common alternative for the definition of this inner product is via matrix multiplication. In order to do this, one defines the $(d+1) \times (d+1)$ matrix

$$\eta = \begin{pmatrix} -1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{1.3}$$

and

$$x \cdot y = y^\intercal \eta x, \tag{1.4}$$

where all multiplication on the right-hand side is ordinary matrix multiplication. Yet another way of writing such an inner product is via Einstein notation. Within this notation, summation symbols are suppressed and instead implied by repeated higher and lower indices. For example

$$\eta^{\mu\nu}x_\mu y_\nu := \sum_{\mu=0}^{d}\sum_{\nu=0}^{d}\eta^{\mu\nu}x_\mu y_\nu = y^\intercal \eta x = x \cdot y. \tag{1.5}$$

Any linear map $L : \mathbb{R}^{(1,d)} \to \mathbb{R}^{(1,d)}$ which preserves this inner product is called a *Lorentz transformation*. In particular such transformations preserve the "distance squared". Special relativity tells us that any physical system is invariant under a Lorentz transformation; the laws of physics do not change when a Lorentz transformation is applied.

A pseudo-Riemannian manifold has more structure than a general manifold, as it provides a notion of distance and angles on the manifold. One should here distinguish between curves that follow spacelike trajectories and ones that follow timelike trajectories. That is, given a curve $\gamma : [0,1] \to M$ such that γ' is always spacelike, the length of γ is defined as

$$\text{length}(\gamma) := \int_{x=0}^{1} dx\sqrt{\gamma'(x)^2}, \tag{1.6}$$

whereas for such γ where γ' is always timelike, the *proper time* elapsed along γ is

$$\text{proper time}(\gamma) := \int_{t=0}^{1} dt\sqrt{-\gamma'(t)^2}. \tag{1.7}$$

Proper time is an important concept, as it can be used to determine the behaviour of a single relativistic point particle. However, it will not be sufficient in string theory, for the simple reason a string traces out a two-dimensional surface, which has both space and time directions. In this specific case, we will be faced with a map of the type $\gamma : [0,1]^2 \to M$, $(\sigma,\tau) \mapsto \gamma(\sigma,\tau)$ such that $\frac{\partial\gamma}{\partial\sigma}$ is spacelike and $\frac{\partial\gamma}{\partial\tau}$ is lightlike, and we need to define its area. This definition would have to depend on $\frac{\partial\gamma}{\partial\sigma}$ and $\frac{\partial\gamma}{\partial\tau}$ and be invariant under reparameterisations. The correct definition in this case is

$$\text{area}(\gamma) = \int_{\tau=0}^{1} d\tau \int_{\sigma=0}^{1} d\sigma \sqrt{\left(\frac{\partial\gamma}{\partial\tau}\cdot\frac{\partial\gamma}{\partial\sigma}\right)^2 - \left(\frac{\partial\gamma}{\partial\tau}\right)^2\left(\frac{\partial\gamma}{\partial\sigma}\right)^2}. \tag{1.8}$$

Since all of $\left(\frac{\partial\gamma}{\partial\tau}\cdot\frac{\partial\gamma}{\partial\sigma}\right)^2$, $-\left(\frac{\partial\gamma}{\partial\tau}\right)^2$, and $\left(\frac{\partial\gamma}{\partial\sigma}\right)^2$ are positive, the quantity in the square root is positive, so this expression is well-defined. Moreover, it depends, up to a sign, linearly on both derivatives, which is an expected property of any area functional.

1.1.2 Angles and conformal maps

Returning to the setting introduced in the previous section, where we have some manifold M with a metric, the metric gives rise to another interesting notion, namely that of angles. However, we will only be using angles in the case of a positive definite metric, so this will be assumed in the remaining of this chapter. Angles are defined in the following manner.

Definition 1.4. *Given two curves $\gamma, \delta : (-1,1) \to M$ intersecting at some $x \in M$ at time $t = 0$ (by which we mean that $\gamma(0) = \delta(0) = x$), the angle α between γ and δ is defined by*

$$\cos(\alpha) := \frac{g(x)(\gamma'(0),\delta'(0))}{\sqrt{g(x)(\gamma'(0),\gamma'(0)\cdot g(x)(\delta'(0),\delta'(0)}} \tag{1.9}$$

$$= \frac{\gamma'(0)\cdot\delta'(0)}{\sqrt{\gamma'(0)^2\delta'(0)^2}}. \tag{1.10}$$

This expression should be familiar, as it is simply a generalisation of the formula $x \cdot y = \cos(\alpha)\,\|x\|\,\|y\|$ which holds in Euclidean space. We can now define so called *conformal maps*, which are maps between manifolds that preserve angles.

Definition 1.5. *Let $n \in \mathbb{Z}_{\geq 1}$, let M and M' be n-manifolds and let g and g' be metrics on M and M' respectively. A **conformal map** from M to M' is a differentiable map $f : M \to M'$ such that for all $x \in M$ and for all $v, w \in T_xM$ we have*

$$g'\big(f(x)\big)\big(f'(x)v, f'(x)w\big) = \lambda^2(x)\cdot g(x)(v,w). \tag{1.11}$$

for some smooth map $\lambda : M \to \mathbb{R}_{>0}$.

It should be clear that a conformal map preserves angles as defined in (1.9). As we are specifically interested in conformal maps and conformal invariance in two dimensions, we will restrict ourselves to the case $M = \mathbb{R}^2$. Here $\mathbb{R}^2$ is equipped with the standard metric $g(x) = \langle\cdot,\cdot\rangle$, which assigns to every point the standard (Euclidean) inner product. Moreover, in order to classify the conformal maps on $\mathbb{R}^2$, we use the isometry $\mathbb{C} \xrightarrow{\sim}$

$\mathbb{R}^2$ given by $a + bi \mapsto (a, b)$ to identify $\mathbb{R}^2$ with the complex plane. Since we are working with a constant metric, (1.11) simplifies considerably to

$$\langle f'(x)v, f'(x)w \rangle = \lambda^2(x) \cdot \langle v, w \rangle. \tag{1.12}$$

The set of all conformal maps is now easily determined; we show all conformal maps on the plane are either holomorphic or antiholomorphic.

Proposition 1.6. *Let $U \subseteq \mathbb{R}^2$ be an open non-empty subset and let $f : U \to \mathbb{R}^2$ given by $z \mapsto (f_1(z), f_2(z))$ be a conformal map (where both U and $\mathbb{R}^2$ are equipped with the standard metric mentioned above). Then, if f is viewed as a complex function, f is either holomorphic or antiholomorphic. In other words, either f or its complex conjugate is holomorphic.*

Proof. First note that for all $z \in U$ the derivative of f is just its Jacobian matrix

$$df(z) = \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix}, \tag{1.13}$$

where $\partial_x f_i$ and $\partial_y f_i$ are the partial derivatives of f_i to the first respectively the second coordinate. By substituting the (orthonormal) standard basis vectors in the place of v and w in (1.11), the right-hand simplifies to either zero or $\lambda^2(x)$. To be more concrete, let $i, j \in \{1, 2\}$, and let $v = e_i$ and $w = e_j$. Then by (1.11) we get

$$\begin{aligned} \lambda^2(x) \cdot \delta_{i,j} = \lambda^2(x) \cdot \langle v, w \rangle &= \left\langle \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix} e_i, \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix} e_j \right\rangle \\ &= \left\langle \begin{pmatrix} \partial_i f_1 \\ \partial_i f_2 \end{pmatrix}, \begin{pmatrix} \partial_j f_1 \\ \partial_j f_2 \end{pmatrix} \right\rangle = \partial_i f_1 \cdot \partial_j f_1 + \partial_i f_2 \cdot \partial_j f_2, \end{aligned} \tag{1.14}$$

and in particular

$$(\partial_x f_1)^2 + (\partial_x f_2)^2 = \lambda^2(x) = (\partial_y f_1)^2 + (\partial_y f_2)^2 \tag{1.15}$$

and

$$\partial_x f_1 \cdot \partial_y f_1 + \partial_y f_2 \cdot \partial_x f_2 = 0. \tag{1.16}$$

To prove f is either holomorphic or antiholomorphic, we now transition explicitly to the complex plane. We write $z = x + iy$ and $\bar{z} = x - iy$ (and equivalently $x = \frac{z+\bar{z}}{2}$ and $y = \frac{z-\bar{z}}{2i}$), and we view f as a complex valued function via $f = f_1 + if_2$. We will finish the proof by showing that either

$\partial_z f = 0$, or $\partial_{\bar{z}} f = 0$, since these conditions are equivalent to f being holomorphic in the first case, and antiholomorphic in the second case. To do this, we will first need explicit expressions for ∂_z and $\partial_{\bar{z}}$.

For a general function $g(x, y)$ we may apply the chain rule to find explicit expressions for $\partial_z = \frac{\partial}{\partial z}$ and $\partial_{\bar{z}} = \frac{\partial}{\partial \bar{z}}$ in terms of ∂_x and ∂_y, namely as

$$\partial_z g(x,y) = \frac{\partial g}{\partial x}\frac{\partial x}{\partial z} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial z} = \frac{\partial g}{\partial x}\frac{1}{2} + \frac{\partial g}{\partial y}\frac{1}{2i} = \frac{1}{2}(\partial_x - i\partial_y)g \tag{1.17}$$

and

$$\partial_{\bar{z}} g(x,y) = \frac{\partial g}{\partial x}\frac{\partial x}{\partial \bar{z}} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial \bar{z}} = \frac{\partial g}{\partial x}\frac{1}{2} - \frac{\partial g}{\partial y}\frac{1}{2i} = \frac{1}{2}(\partial_x + i\partial_y)g. \tag{1.18}$$

Hence we conclude $\partial_z = \frac{\partial_x - i\partial_y}{2}$ and $\partial_{\bar{z}} = \frac{\partial_x + i\partial_y}{2}$. We can now directly calculate the product $(\partial_z f)\overline{(\partial_{\bar{z}} f)}$ by splitting the real and imaginary part. We have

$$4(\partial_z f)\overline{(\partial_{\bar{z}} f)} = \big((\partial_x - i\partial_y)(f_1 + if_2)\big) \cdot \overline{\big((\partial_x + i\partial_y)(f_1 + if_2)\big)},$$

so*

$$\begin{aligned} 4\,\mathrm{Re}[(\partial_z f)\overline{(\partial_{\bar{z}} f)}] &= (\partial_x f_1 + \partial_y f_2)(\partial_x f_1 - \partial_y f_2) + (\partial_x f_2 - \partial_y f_1)(\partial_x f_2 + \partial_y f_1) \\ &= (\partial_x f_1)^2 - (\partial_y f_2)^2 + (\partial_x f_2)^2 - (\partial_y f_1)^2 \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} 4\,\mathrm{Im}[(\partial_z f)\overline{(\partial_{\bar{z}} f)}] &= (\partial_x f_2 - \partial_y f_1)(\partial_x f_1 - \partial_y f_2) - (\partial_x f_1 + \partial_y f_2)(\partial_x f_2 + \partial_y f_1) \\ &= -2\partial_x f_1 \partial_y f_1 - 2\partial_x f_2 \partial_y f_2 \\ &= 0. \end{aligned}$$

Both the "$= 0$" follow directly from (1.15). Hence either $\partial_z f = 0$ or $\partial_{\bar{z}} f = 0$, which completes the proof. Note that in the case where in some point both $\partial_z f$ and $\partial_{\bar{z}}$ are zero, we simply have $f' = 0$ which implies $\lambda = 0$, which is excluded by definition of conformal maps. □

Remarkably, the situation in two dimensions is unique, in the sense that the set of conformal transformations is "very large". It can be shown that

*Here we use the identities $\mathrm{Re}(z\overline{w}) = \mathrm{Re}(z)\,\mathrm{Re}(w) + \mathrm{Im}(z)\,\mathrm{Im}(w)$ and $\mathrm{Im}(z\overline{w}) = \mathrm{Im}(z)\,\mathrm{Re}(w) - \mathrm{Re}(z)\,\mathrm{Im}(w)$ for complex numbers $z, w \in \mathbb{C}$.

Version of July 11, 2020– Created July 11, 2020 - 11:00

in higher dimensions, all conformal transformations can be obtained by taking combinations of translations, rotations, dilatation, reflections, and so called special conformal transformations [Sch08]. The latter are given by

$$x \mapsto \frac{x - \langle x, x\rangle b}{1 - 2\langle x, b\rangle + \langle x, x\rangle\langle b, b\rangle}$$

for some $b \in \mathbb{R}^n$, and are obtained by doing an inversion ($x \mapsto \frac{x}{\langle x,x\rangle}$), translating by b and doing another inversion. The set of conformal transformations (in more than two dimensions) is therefore generated by only a finite set of infinitesimal transformations [Sch08]. More formally, the Lie algebra corresponding to these transformations is finite dimensional. The contrary is true in two dimensions, as the set of holomorphic transformations requires an *infinitely* set of generators. This implies conformal invariance in two dimensions is much more restrictive than it would be in higher dimensions, which motivates the study of this specific case.

1.2 Lagrangian mechanics

As we will see, string theory provides an example of conformal invariance. The theory of the open bosonic string in chapter 3 will be formulated via a Lagrangian. We will give a brief introduction to Lagrangian mechanics and treat the example of the relativistic point particle.

1.2.1 The Lagrangian

In order to formulate Lagrangian mechanics we need a way to describe physical objects and their motions. We do this by giving a parameterisation, which requires two main ingredients: a parameter space and a target space. The parameter space tells us what the object looks like and the time period we are interested in. For simplicity we will always assume our object is a smooth connected manifold M, which justifies us to take the parameter space to be $M \times \mathbb{R}$. Here $\mathbb{R}$ represents the time domain. The target space is the space in which the object moves. In the simplest case this is flat spacetime, where the metric is constant, which is exactly Minkowski space. These observations justify the following definition.

Definition 1.7. *A **parameterisation** is a differentiable map $X : \mathbb{R} \times M \to \mathbb{R}^{(1,d)}$ written as $X = (X^0, X^1, \ldots, X^d) = (X^0, \vec{X})$ such that the following hold:*

- *For each $t \in \mathbb{R}$ the map $M \to \mathbb{R}^{(1,d)}$ given by $m \mapsto X(t, m)$ is injective.*

Version of July 11, 2020– Created July 11, 2020 - 11:00

- *For each* $(t, m) \in \mathbb{R} \times M$ *the vector* $\frac{\partial}{\partial t} X(t, m)$ *is either timelike of lightlike.*

The above requirements are clearly necessary for our definition to represent any physical object, but they will not be sufficient for specific purposes. For instance, no physical laws are incorporated within this definition, hence additional requirements will be made dependent on the specific context. The way this will be done is through a Lagrangian.

We will formulate Lagrangian mechanics in natural units. This means we set

$$c = \hbar = 1, \tag{1.19}$$

where c is the speed of light, and $\hbar$ the reduced Planck constant. This means we measure time, distance and inverse energy in the same units. Especially the fact that we do not distinguish between distance and time units seems reasonable, since relativity tells us space and time are not fully distinct but can be mixed instead. The biggest advantage of natural units is that it cleans up formulas without loss of information, as the correct quantities of c and $\hbar$ can always be recovered by comparing units.

Definition 1.8. *A* ***Lagrangian density*** *is a map which assigns to each parmaterisation* X *a function* $\mathcal{L}(X) : \mathbb{R} \times M \to \mathbb{R}$*, which depends differentiably on* X *and the derivatives of* X*. The associated* ***Lagrangian*** *is the function* $L(X)$ *given by*

$$L(X) := \int_M \mathcal{L}(X)\, dV.$$

The Lagrangian has units of energy.

Definition 1.9. *Given a parameterisation* X *and a Lagrangian* L *as in definition 1.8 and a closed interval* $[t_i, t_f]$ *with more than one point, the* ***action*** *is given by*

$$S_{[t_i,t_f]}(X) := \int_{t_i}^{t_f} \int_M \mathcal{L}(X)\, dV\, dt = \int_{t_i}^{t_f} L(X)\, dt. \tag{1.20}$$

Often the subscript will be omitted if the interval is clear.

It follows that the action is dimensionless in natural units. We are now able to state the primary axiom of Lagrangian mechanics.

Definition 1.10. *A parameterisation* X *is said to describe a (classical) physical object if for every closed interval* $[t_i, t_f]$ *and every differentiable map* $Y : \mathbb{R} \times M \to \mathbb{R}^{(1,d)}$ *such that for all* $m \in M$*,* $Y(m, t_i) = Y(m, t_f) = 0$*, we have*

$$\frac{\partial}{\partial \epsilon} S(X + \epsilon Y) = 0, \quad \text{with } \epsilon \in \mathbb{R}. \tag{1.21}$$

Solving this equation generally fixes a set of equations of motion, which may need to be supplemented with separate boundary conditions. It is important to note that although an action can be used to fully fix the equations of motions, the action that does this is not necessarily unique. For example, scalar multiples of an action give the same equations of motion, but actions which look vastly different may also have the same solutions. This can be exploited when studying a system, as different actions can expose different properties, such as symmetries, of a system. This is something we will use in the study of the conformal properties of the free string, in chapter 3.

Another useful result from Lagrangian mechanics states that if $L(X)$ only depends on X and the first order derivatives of X, and if X explicitly depends on a set of parameters $\zeta_1, ..., \zeta_n$, then the requirement of the action being stationary in X is equivalent to the *Euler-Lagrange equation*, which is given by

$$\frac{\partial L}{\partial X} - \sum_{i=1}^{n} \frac{\partial}{\partial \zeta_i} \mathcal{P}^{\zeta_i} = 0 \tag{1.22}$$

where $\mathcal{P}^{\zeta_i}$ are the conjugate momenta defined by

$$\mathcal{P}^{\zeta_i} = \frac{\partial \mathcal{L}}{\partial(\frac{\partial X}{\partial \zeta_i})}. \tag{1.23}$$

1.2.2 The free relativistic point particle

We will discuss a short example of Lagrangian mechanics, namely the free relativistic point particle. By definition 1.7 a parameterisation of a point particle is a map $X : \mathbb{R} \times \{point\} \to \mathbb{R}^{(1,d)}$, but to simplify notation we will view this as a map $X : \mathbb{R} \to \mathbb{R}^{(1,d)}$. Given a time interval $[t_i, t_f] \subseteq \mathbb{R}$, the action is then given by the *proper time* that has elapsed on the path that our particle traces out in the target space. This means we set

$$L = m\sqrt{-\left(\frac{\partial X}{\partial t}\right)^2} \tag{1.24}$$

and consequently

$$S(X) = m\int_{t_i}^{t_f} \sqrt{-\left(\frac{\partial X}{\partial t}\right)^2}\, dt. \tag{1.25}$$

Version of July 11, 2020– Created July 11, 2020 - 11:00

The mass is included to make sure the dimensions are correct. For simple parameterisations which are taken in the *static guage*, which means $X(t) = (t, \vec{X}(t))$ for all t, this means we take

$$S(X) = m \int_{t_i}^{t_f} \sqrt{1 - \left(\frac{\partial \vec{X}}{\partial t}\right)^2}\, dt. \tag{1.26}$$

Note that since the particle is assumed to travel slower than the speed of light, $\frac{\partial X}{\partial t}$ is assumed to be timelike. Therefore, the term within the square root is indeed positive, and this expression is well-defined. If we were now given a path Y as in definition 1.10, we have

$$\begin{aligned} S(X + \epsilon Y) &= m \int_{t_i}^{t_f} \sqrt{-\left(\frac{\partial (X + \epsilon Y)}{\partial t}\right)^2}\, dt. \\ &= m \int_{t_i}^{t_f} \sqrt{-\left(\frac{\partial X}{\partial t}\right)^2 - 2\epsilon \left(\frac{\partial X}{\partial t}\right)\left(\frac{\partial Y}{\partial t}\right) - \epsilon^2 \left(\frac{\partial Y}{\partial t}\right)^2}\, dt. \end{aligned}$$

Here we can do a Taylor expansion with respect to ϵ, and discard any terms of order 2 or higher. We will also introduce the notation $\dot{X} = \frac{\partial X}{\partial t}$. This gives

$$S(X + \epsilon Y) \approx m \int_{t_i}^{t_f} \left(\sqrt{-(\dot{X})^2} - \epsilon \frac{\dot{X}}{\sqrt{-(\dot{X})^2}} \dot{Y} \right) dt.$$

We can now calculate $\frac{\partial}{\partial \epsilon} S(X + \epsilon Y)$ directly, were we use integration by parts in the third equality:

$$\begin{aligned} \frac{\partial}{\partial \epsilon} S(X + \epsilon Y) &= \lim_{\epsilon \to 0} \frac{S(X + \epsilon Y) - S(X)}{\epsilon} \\ &= m \int_{t_i}^{t_f} \frac{\dot{X}}{\sqrt{-(\dot{X})^2}} \dot{Y}\, dt \\ &= m \left[\frac{\dot{X}}{\sqrt{-(\dot{X})^2}} Y \right]_{t_i}^{t_f} - m \int_{t_i}^{t_f} \frac{d}{dt} \left(\frac{\dot{X}}{\sqrt{-(\dot{X})^2}} \right) Y\, dt. \end{aligned}$$

But by the assumptions made on Y in definition 1.10 we have $Y(t_i) = Y(t_f) = 0$, so the first term vanishes. This implies the second term also

has to vanish for every path Y, and this gives us

$$m\frac{d}{dt}\left(\frac{\dot{X}}{\sqrt{-(\dot{X})^2}}\right) = 0. \tag{1.27}$$

To interpret this result, we return to the static gauge, where $X(t) = (t, \vec{X}(t))$. Here we can define $\vec{V}(t) = \frac{d}{dt}\vec{X}$, which gives us $\dot{X} = (1, \vec{V}(t))$. Additionally $\frac{1}{\sqrt{-(\dot{X})^2}}$ becomes $\frac{1}{\sqrt{1-\vec{V}^2}}$, which can be recognised as the relativistic Lorentz factor γ. If we now look at the spacial part of (1.27), we find $\frac{d}{dt}(\gamma\vec{V}) = 0$, which tells us exactly that the relativistic momentum

$$\vec{p} = \gamma m\vec{V} \tag{1.28}$$

is conserved. This is the known equation of motion for a single particle in the absence of forces, which establishes the idea the Lagrangian was chosen properly. Moreover, this implies the direction of $\vec{V}$ is constant. Additionally, since we have

$$\frac{1}{1+(\gamma\vec{V})^2} = \frac{1}{1+\frac{\vec{V}^2}{1-\vec{V}^2}} = \frac{1-\vec{V}^2}{1-\vec{V}^2+\vec{V}^2} = 1-\vec{V}^2, \tag{1.29}$$

the conservation of $\gamma\vec{V}$ implies the conservation of $\vec{V}^2$, and so $\vec{V}$ is itself constant. Therefore, a full description of the trajectory of the point particle can now be described by fixing a value for $\vec{V}(0) = \vec{V}_0$, from which it follows that

$$\vec{X}(t) = \vec{X}(0) + \int_0^t \vec{V}(t')dt' = \vec{X}(0) + \int_0^t \vec{V}_0 dt' = \vec{X}(0) + t\vec{V}_0. \tag{1.30}$$

Hence the full trajectory is determined by the initial position and the initial velocity, or equivalently by the initial position and the initial momentum.

1.3 The quantum formalism

The Lagrangian formalism described in the previous section can be used to describe a classical system. Given a Lagrangian, equations of motion can be derived, which tell us exactly which parameterisations are allowed. These equations of motion usually take the shape of a set of differential equations, which in general do not have a single unique solution, but more

so a family of solutions, characterised by some free variables. For example, the relativistic point particle can be completely described by fixing an initial position $\vec{X}_0$ and an initial momentum $\vec{p}_0$. In a quantum theory of this classical system these free variables are now turned into Hermitian operators on some state space, with appropriate commutation relations. Here, one can opt for either Schrödinger or Heisenberg operators. Schrödinger operators remain constant in time while their states change, while Heisenberg operators change over time while their states remain the same. Since in chapter 3 we will work in the Heisenberg picture, we will now illustrate how one can transition between the two pictures.

In the case where the Schrödinger Hamiltonian is time-independent, any state ψ evolving according to the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(t) = H\psi(t) \tag{1.31}$$

satisfies $\psi(t) = e^{-iHt}\psi$. Therefore, the operator e^{iHt} sends any time dependent state to a time independent one. Hence, given any Schrödinger operator $\mathcal{O}$, the corresponding Heisenberg operator is $\mathcal{O}(t) = e^{iHt}\mathcal{O}e^{-iHt}$, since for any state ψ

$$\mathcal{O}(t)\psi = e^{iHt}\mathcal{O}e^{-iHt}\psi = e^{iHt}(\mathcal{O}\psi(t)), \tag{1.32}$$

which is indeed $\mathcal{O}\psi(t)$ brought to rest. Consequently, the time evolution of a Heisenberg operator can neatly be calculated from its commutator with the Hamiltonian, via

$$\begin{aligned}\frac{\partial}{\partial t}\mathcal{O}(t) &= \frac{\partial}{\partial t}(e^{iHt}\mathcal{O}e^{-iHt}) \\ &= (iHe^{iHt})\mathcal{O}e^{-iHt} + e^{iHt}\mathcal{O}(-iHe^{-iHt}) \\ &= iH\mathcal{O}(t) - i\mathcal{O}(t)H \\ &= i[H,\mathcal{O}(t)].\end{aligned} \tag{1.33}$$

We will now derive one final result, which states that for any two Schrödinger operators $\mathcal{O}$ and $\mathcal{U}$ the equality $[\mathcal{O}(t),\mathcal{U}(t)] = [\mathcal{O},\mathcal{U}](t)$ holds. This simply follows from the following calculation:

$$\begin{aligned}[\mathcal{O}(t),\mathcal{U}(t)] &= \mathcal{O}(t)\mathcal{U}(t) - \mathcal{U}(t)\mathcal{O}(t) \\ &= e^{iHt}\mathcal{U}e^{-iHt}e^{iHt}\mathcal{O}e^{-iHt} - e^{iHt}\mathcal{U}e^{-iHt}e^{iHt}\mathcal{O}e^{-iHt} \\ &= e^{iHt}\mathcal{U}\mathcal{O}e^{-iHt} - e^{iHt}\mathcal{U}\mathcal{O}e^{-iHt} \\ &= e^{iHt}[\mathcal{O},\mathcal{U}]e^{-iHt} \\ &= [\mathcal{O},\mathcal{U}](t).\end{aligned} \tag{1.34}$$

This result is for example useful in the common case when one has two Schrödinger operators q and p, which represent some position and the corresponding conjugate momentum. These operators should have the canonical commutation relation $[q, p] = i$, hence it follows their Heisenberg counterparts satisfy $[q(t), p(t)] = i$. More generally, (1.34) allows us to (often implicitly) switch between the Schrödinger and Heisenberg pictures more easily. This will be an advantage in our discussion of the quantum theory of the bosonic string in chapter 3. First however, we treat the shorter example of the quantum harmonic oscillator.

1.3.1 The quantum harmonic oscillator

In order to establish a quantum theory of the harmonic oscillator, we first describe the classical theory. A classical harmonic oscillator is a physical system whose kinetic energy is proportional to its velocity squared, and its potential energy is proportional to its position squared. Hence, if the position of the particle is given by $q(t)$, the Lagrangian can be written as $L(t) = \frac{1}{2}\zeta\dot{q}(t)^2 - \frac{1}{2}\eta q(t)^2$, for suitable constants ζ and η. Since scaling of L does not change the behaviour of the system, we may assume it is of the form

$$L = \frac{1}{2\omega}\dot{q}^2 - \frac{\omega}{2}q^2. \tag{1.35}$$

Then the conjugate momentum to q is

$$p = \frac{\partial L}{\partial \dot{q}} = \frac{1}{\omega}\dot{q}. \tag{1.36}$$

and the Hamiltonian is given by

$$H = p\dot{q} - L = \frac{1}{2\omega}\dot{q}^2 + \frac{\omega}{2}q^2 = \frac{\omega}{2}(p^2 + q^2). \tag{1.37}$$

In order to quantise this oscillator, we introduce Hermitian Schrödinger operators p and q, with canonical commutation relation $[q, p] = i$. We then define the operators $a = \frac{1}{\sqrt{2}}(p - iq)$ and its adjoint $a^\dagger = \frac{1}{\sqrt{2}}(p + iq)$. A direct calculation then tells us that $[a, a^\dagger] = 1$, and we can retrieve p and q as $p = \frac{1}{\sqrt{2}}(a + a^\dagger)$ and $q = \frac{i}{\sqrt{2}}(a - a^\dagger)$. Thus we can rewrite the

Hamiltonian as

$$\begin{aligned} H &= \frac{\omega}{2}(p^2+q^2) = \frac{\omega}{4}\left((a+a^\dagger)^2-(a-a^\dagger)^2\right) \\ &= \frac{\omega}{4}\left((a^2+a^{\dagger 2}+aa^\dagger+a^\dagger a)-(a^2+a^{\dagger 2}-aa^\dagger-a^\dagger a)\right) \\ &= \frac{\omega}{2}\left(aa^\dagger+a^\dagger a\right) = \frac{\omega}{2}\left(aa^\dagger+aa^\dagger-[a,a^\dagger]\right) = \omega(aa^\dagger-\frac{1}{2}). \end{aligned} \tag{1.38}$$

Here we recognise $a^\dagger$ and a as *raising and lowering operators* respectively, since for any eigenstate $|h\rangle$ of H with eigenvalue h, we have

$$\begin{aligned} H(a|h\rangle) &= \omega(aa^\dagger-\frac{1}{2})a|h\rangle = a\left(\omega(a^\dagger a-\frac{1}{2})|h\rangle\right) \\ &= a(H-\omega)|h\rangle) = (h-\omega)(a|h\rangle) \end{aligned} \tag{1.39}$$

and

$$\begin{aligned} H(a^\dagger|h\rangle) &= \omega(aa^\dagger-\frac{1}{2})a^\dagger|h\rangle = \omega(a^\dagger a+\frac{1}{2})a^\dagger|h\rangle \\ &= a^\dagger(H+\omega)|h\rangle) = (h+\omega)(a^\dagger|h\rangle). \end{aligned} \tag{1.40}$$

Hence $a^\dagger$ and a raise and lower the H eigenvalue (or energy) of states, which justifies the name. They are also often called *creation and annihilation operators*, as they take this role in a more general quantum field theory. Additionally, (1.38) gives us

$$\omega a(t)a^\dagger(t) = e^{iHt}aa^\dagger e^{-iHt} = e^{iHt}(H+\frac{\omega}{2})e^{-iHt} = H+\frac{\omega}{2} = \omega aa^\dagger. \tag{1.41}$$

We can now compute $\dot{a}(t)$ using all previous results:

$$\begin{aligned} \frac{\partial}{\partial t}a(t) &\overset{(1.33)}{=} i[H,a(t)] \overset{(1.38)}{=} i\omega[aa^\dagger-1/2,a(t)] \\ &\overset{(1.41)}{=} i\omega[a(t)a^\dagger(t),a(t)] = i\omega a(t)[a^\dagger(t),a(t)] \\ &\overset{(1.34)}{=} i\omega a(t)[a^\dagger,a](t) = -i\omega a(t). \end{aligned} \tag{1.42}$$

The differential equation is solved by

$$a(t) = e^{-i\omega t}a, \tag{1.43}$$

and consequently

$$a^\dagger(t) = e^{i\omega t}a^\dagger. \tag{1.44}$$

We then obtain

$$q(t) = \frac{i}{\sqrt{2}}(ae^{-i\omega t}-a^\dagger e^{i\omega t}). \tag{1.45}$$

1.3.2 Symmetries and Noether's theorem

So far we have formally defined what we view as a classical or quantum system. In this thesis we are primarily interested in systems that have conformal symmetry. The precise definition of a symmetry depends on the context. In our framework, where we have some Lagrangian L, a symmetry is a smooth map $h : \mathbb{R}^{(1,d)} \to \mathbb{R}^{(1,d)}$ such that $L(h \circ X) = L(X)$, and h is everywhere locally bijective, which means its derivative is always invertible. Symmetries can often be described by *one-parameter families*, which are sets $\{h_s\}_{s\in\mathbb{R}}$ of symmetries such that $h_{s+t} = h_s \circ h_t$, and in particular $h_0 = id$. It is important to note such families are fully determined once they are known in *any* neighbourhood of $s = 0$. Hence, a one-parameter family of symmetries is already fully fixed once one knows its behaviour around $s = 0$, i.e. once one knows

$$\delta h := \left.\frac{\partial h_s}{\partial s}\right|_{s=0}.$$

The map δh is called the infinitesimal transformation corresponding to the family $\{h_s\}_{s\in\mathbb{R}}$. The *generator* of the transformations is the differential operator $\delta h\nabla$, where ∇ is some formal differential operator acting on the parameter space, whose precise definition depends on the specific context. It can be shown that the set of the differential operators obtained this way is closed under the commutator, and hence this set forms an algebraic structure which is known as a Lie algebra [BK13].

On the other hand, *Noether's theorem* tells us that for such a family $\{h_s\}_{s\in\mathbb{R}}$ we get a conserved charge Q, which in the quantum theory becomes a Hermitian operator. This operator then generates the symmetry transformation the classical charge originated from, and the set of these operators is also closed under the commutator, and we again obtain a Lie algebra. The two Lie algebras we have obtained are connected: the algebra of the quantum operators is isomorphic to (a central extension of) the complexification of the algebra of classical generators. For example, in the next chapter we will show that the classical generators of conformal transformations form the Witt algebra, which implies that the corresponding quantum theory should be a representation of the central extension of the Witt algebra, which is the Virasoro algebra.

A similar argument can be made for Lorentz invariance. The Lagrangian of the harmonic oscillator which we previously analysed is invariant under Lorentz transformations, and Noether's theorem hence implies there

are conserved charges. Explicitly, these charges are given by

$$M^{\mu\nu} = x^\mu p^\nu - p^\nu x^\mu, \tag{1.46}$$

which in the quantum theory turn into operators with commutation relations

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\eta^{\mu\rho} M^{\nu\sigma} - i\eta^{\nu\rho} M^{\mu\sigma} + i\eta^{\mu\sigma} M^{\rho\nu} - i\eta^{\nu\sigma} M^{\rho\mu}. \tag{1.47}$$

These commutation relations defines the Lorentz lie algebra. Any quantum theory which is compatible with special relativity has to be Lorentz invariant, and therefore it has to contain a set of operators satisfying these relations. In chapter 3, this fact will be used to conclude bosonic string theory can only exist in 26 spacetime dimensions.

In conclusion, we can describe symmetries of both classical and quantum systems through Lie algebras. This inspires us to study Lie algebras in more detail. In particular, we will in detail study the Lie algebra associated to conformal symmetries, namely the Virasoro algebra.

Chapter 2

The Virasoro algebra

In this chapter we study Lie algebras, the natural algebraic structure associated to symmetries. In particular, we will study the Virasoro algebra, which is the Lie algebra associated to conformal transformations. We start with the general theory of algebras and Lie algebras, including the necessary concepts of homomorphisms and ideals. We then define representations of Lie algebras and the universal enveloping algebra, and we state the Poincaré-Birkhoff-Witt theorem. We conclude by deriving the Witt and Virasoro algebras as the Lie algebras corresponding to conformal transformations.

2.1 Lie algebras

In this section, $\mathbb{F}$ is a field, and all vector spaces are taken over $\mathbb{F}$. However, the less mathematically oriented reader can always think of $\mathbb{F}$ as either $\mathbb{R}$ or $\mathbb{C}$, as in physics we only encounter Lie algebras over these two fields. Before we define Lie algebras, we first define the more general concept of *algebras*.

Definition 2.1. *An **algebra** is a vector space $\mathfrak{g}$ equipped with a bilinear multiplication $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which is called the **bracket**. If $[\cdot,\cdot]$ is associative, then $\mathfrak{g}$ is called an **associative algebra**. If there is an element $1 \in \mathfrak{g}$ such that for all $X \in \mathfrak{g}$ we have $[1, X] = [X, 1] = X$, then $\mathfrak{g}$ is called **unital** and e is called a **unit**.*

As is common practice in algebra, we define the relevant morphisms between algebras as the maps which in some way respect the algebraic structure.

Definition 2.2. *Let $(\mathfrak{g}, [\cdot,\cdot]_\mathfrak{g}), (\mathfrak{h}, [\cdot,\cdot]_\mathfrak{h})$ be algebras. An **algebra homomorphism** is a linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ such that for all $X, Y \in \mathfrak{g}$*

$$\phi([X,Y]_\mathfrak{g}) = [\phi(X), \phi(Y)]_\mathfrak{h}. \tag{2.1}$$

*If ϕ is a bijection, we call it an **algebra isomorphism**.*

Just as in the case of group or ring theory, we will also look at quotient spaces and formulate an isomorphism theorem.

Definition 2.3. *Let $(\mathfrak{g}, [\cdot,\cdot])$ be an algebra. A linear subspace $\mathfrak{i} \subseteq \mathfrak{g}$ is called*

- *a **subalgebra** if $[\mathfrak{i},\mathfrak{i}] \subseteq \mathfrak{i}$ (i.e. for all $E, F \in \mathfrak{i}$ we have $[E,F] \in \mathfrak{i}$);*
- *a (two-sided) **ideal** if $[\mathfrak{g},\mathfrak{i}] \subseteq \mathfrak{i}$ and $[\mathfrak{i},\mathfrak{g}] \subseteq \mathfrak{i}$.*

The group theoretic analogues of subalgebras and ideals are subgroups and normal subgroups respectively. Accordingly, an ideal of a Lie algebra allow us to define the corresponding quotient space, which is also a Lie algebra.

Proposition 2.4. *Let $(\mathfrak{g}, [\cdot,\cdot])$ be an algebra, and $\mathfrak{i} \subseteq \mathfrak{g}$ an ideal. Let $\mathfrak{q} := \mathfrak{g}/\mathfrak{i}$ be the quotient vector space, and denote its elements as $\lfloor X \rceil := X + \mathfrak{i}$. Then $\mathfrak{q}$ is also an algebra with bracket defined by $[\lfloor X \rceil, \lfloor Y \rceil] = \lfloor [X,Y] \rceil$.*

Proof. We only need to show the definition of this bracket is representative independent. So let $X, X', Y \in \mathfrak{g}$, such that $E = X - X' \in \mathfrak{i}$. Note $[E,Y] \in \mathfrak{i}$, so $\lfloor [E,Y] \rceil = 0$. Hence we have

$$\lfloor [X,Y] \rceil = \lfloor [(X - X') + X', Y] \rceil = \lfloor [E,Y] \rceil + \lfloor [X',Y] \rceil = \lfloor [X',Y] \rceil . \tag{2.2}$$

The proof for representative independence of the second coordinate is done in a similar fashion. □

We now have all the necessary tools to formulate the first isomorphism theorem for Lie algebras.

Theorem 2.5. *Let $(\mathfrak{g}, [\cdot,\cdot]_\mathfrak{g})$ and $(\mathfrak{h}, [\cdot,\cdot]_\mathfrak{h})$ be algebras, and $\phi : \mathfrak{g} \to \mathfrak{h}$ an algebra homomorphism. Then:*

1. *$\ker(\phi)$ is an ideal of $\mathfrak{g}$;*
2. *$\mathrm{im}(\phi)$ is a subalgebra of $\mathfrak{h}$;*
3. *there exists a unique algebra isomorphism $\overline{\phi} : \mathfrak{g}/\ker(\phi) \to \mathrm{im}(\phi)$ such that for all $X \in \mathfrak{g}$ we have $\phi(g) = \overline{\phi}(\lfloor X \rceil)$.*

Proof. Note the first isomorphism theorem for vector spaces already tells us $\ker(\phi)$ and $\mathrm{im}(\phi)$ are subspaces, and that there exists a $\overline{\phi} : \mathfrak{g}/\ker(\phi) \to \mathrm{im}(\phi)$ that is a vector space isomorphism. We therefore only need to proof the properties involving the bracket.

1. For $E \in \ker(\phi)$ and $X \in \mathfrak{g}$ we have $\phi([E,X]_\mathfrak{g}) = [\phi(E),\phi(X)]_\mathfrak{h} = [0,\phi(X)]_\mathfrak{h} = 0$, so $[E,X]_\mathfrak{g} \in \ker(\phi)$. A similar argument gives $[X,E]_\mathfrak{g} \in \ker(\phi)$, and hence $\ker(\phi)$ is an ideal.
2. For $Z,W \in \mathrm{im}(\phi)$, we may write $Z = \phi(X)$ and $W = \phi(Y)$ for some $X,Y \in \mathfrak{g}$. Hence $[Z,W]_\mathfrak{h} = [\phi(X),\phi(Y)]_\mathfrak{h} = \phi([X,Y]_\mathfrak{g}) \in \mathrm{im}(\phi)$, so $\mathrm{im}(\phi)$ is a subalgebra.
3. Let $X,Y \in \mathfrak{g}$. Then $\overline{\phi}([\lfloor X \rceil, \lfloor Y \rceil]_{\mathfrak{g}/\mathfrak{i}}) = \overline{\phi}(\lfloor [X,Y]_\mathfrak{g} \rceil) = \phi([X,Y]_\mathfrak{g}) = [\phi(X),\phi(Y)]_\mathfrak{h} = \left[\overline{\phi}(\lfloor X \rceil), \overline{\phi}(\lfloor Y \rceil)\right]_\mathfrak{h}$. Thus $\overline{\phi}$ is an algebra homomorphism, and since it's an isomorphism of vector spaces, it is an algebra isomorphism.

□

We are specifically interested in a certain class of algebras, namely the Lie algebras.

Definition 2.6. *A **Lie algebra** is an algebra $(\mathfrak{g},[\cdot,\cdot])$, such that $[\cdot,\cdot]$ is alternating*(i.e. $[X,X] = 0$ for all $X \in \mathfrak{g}$), and for all $X,Y,Z \in \mathfrak{g}$ we have the so called Jacobi identity*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0. \tag{2.3}$$

*The bracket of a Lie algebra is called a **Lie bracket**, and a subspace $\mathfrak{h} \subseteq \mathfrak{g}$ that is closed under the Lie bracket is called a **Lie subalgebra**.*

We first give a couple of examples of Lie algebras.

Example 2.7. *(i) Take $\mathfrak{g} = \mathbb{F}^3$, and take the Lie bracket equal to the cross-product, which is defined as*

$$\left[\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}, \begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}\right] = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \times \begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix} = \begin{pmatrix}x_2y_3 - x_3y_2\\x_3y_1 - x_1y_3\\x_1y_2 - x_2y_1\end{pmatrix}. \tag{2.4}$$

*(ii) Let A be an associative algebra, and define $[\cdot,\cdot] : A \times A \to A$ as the **commutator**: $[X,Y] = XY - YX$. Then $(A,[\cdot,\cdot])$ is a Lie algebra. A common example is to take $A = \mathrm{End}_\mathbb{F}(V)$ with V a vector space, with composition as multiplication.*

(iii) Let $n \in \mathbb{N}$, and let V an n-dimensional vector space. Define $\mathfrak{sl}(V) = \{X \in \mathrm{End}_\mathbb{F}(V) \,|\, \mathrm{Tr}(X) = 0\}$ as the set of traceless endomorphisms of V, and define the Lie bracket on $\mathfrak{sl}(V)$ again as the commutator. To see this is

*Any bilinear map which is alternating is also antisymmetric. The converse is also true, *unless* the characteristic of $\mathbb{F}$ equals 2.

well-defined, let $X, Y \in \mathrm{End}_{\mathbb{F}}(V)$, choose a basis of V and write X and Y in matrix form as $X = (X_{ij})_{1\leq i,j\leq n}$, $Y = (Y_{ij})_{1\leq i,j\leq n}$. Then

$$\begin{aligned}\mathrm{Tr}(XY) &= \sum_{i=1}^{n}(XY)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} X_{ij}Y_{ji} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} X_{ij}Y_{ji} = \sum_{i=1}^{n}(YX)_{i,i} = \mathrm{Tr}(YX).\end{aligned} \tag{2.5}$$

So in particular, for any $X, Y \in \mathfrak{sl}(V)$ we have $\mathrm{Tr}([X,Y]) = \mathrm{Tr}(XY - YX) = \mathrm{Tr}(XY) - \mathrm{Tr}(YX) = \mathrm{Tr}(XY) - \mathrm{Tr}(XY) = 0$, so $[X,Y] \in \mathfrak{sl}(V)$. Since in general for $X, Y \in \mathfrak{sl}(V)$ the product XY is not contained in $\mathfrak{sl}(V)$, $\mathfrak{sl}(V)$ is not an $\mathbb{F}$ algebra in the standard way (i.e. with composition as multiplication). Therefore, if one would want to study this space, it would be natural to view it as a Lie algebra.

2.2 Lie algebra representations and the universal enveloping algebra

One should recall the main reason we study Lie algebras (in this thesis) is because they are the natural structure formed by symmetries. As outlined in section 1.3.2, such a Lie algebra appears in the quantum theory as a set of operators on the state space. In mathematical terms, this means the state space being a *representation* of the Lie algebra. In this section we formally define Lie algebra representations, and we construct the *universal enveloping algebra*: an associative algebra which can be used to further study representations.

We first give the definition of a Lie algebra representation.

Definition 2.8. *Let $(\mathfrak{g}, [\cdot,\cdot])$ be an Lie algebra. A **Lie algebra representation** of $\mathfrak{g}$ is a pair (V, ρ), with V a vector space and $\rho : \mathfrak{g} \to \mathrm{End}_{\mathbb{F}}(V)$ a Lie algebra homomorphism. Here we view $\mathrm{End}_{\mathbb{F}}(V)$ as a Lie algebra with the commutator as Lie bracket.*

Although according to this definition the pair (V, ρ) is the representation, it is common to be sloppy and refer to either V or ρ as "the representation". However, both should always be kept in mind. Furthermore, the sentence "Let (V, ρ) be a representation of $\mathfrak{g}$" should be understood as "Let V be a vector space and let $\rho : \mathfrak{g} \to \mathrm{End}_{\mathbb{F}}(V)$ be a Lie algebra homomorphism".

Since in general V and ρ can be very complicated, it is natural to look

at the "smallest" possible nontrivial representations. These are so called irreducible representations.

Definition 2.9. *Let $(\mathfrak{g}, [\cdot,\cdot])$ be a Lie algebra, V a vector space, and $\rho : \mathfrak{g} \to \mathrm{End}_{\mathbb{F}}(V)$ a Lie algebra representation. Let $W \subseteq V$ be a subspace. Now define $\rho_W : \mathfrak{g} \to \mathrm{Hom}_{\mathbb{F}}(W, V)$ by $X \mapsto \rho(X)|_W$. If for all $X \in \mathfrak{g}$ the map $\rho_W(X)$ is an endomorphism of W, which means the image of $\rho_W(X)$ is contained in W, we may view ρ_W as a map from $\mathfrak{g}$ to $\mathrm{End}_{\mathbb{F}}(W)$. In this case, we call (W, ρ_W) a **subrepresentation** of (V, ρ). The subrepresentations (V, ρ) and $(0, 0)$ are called the **trivial subrepresentations** of (V, ρ). If (V, ρ) has exactly two subrepresentations, then it is called **irreducible**.*

Since in general the image of ρ is not closed under composition, one may try to extend $\mathfrak{g}$ to a larger structure on which ρ does have this property. The most obvious choice for this larger structure would be an unital associative algebra, since $\mathrm{End}_{\mathbb{F}}(V)$ already has this structure. More concrete, we will embed $\mathfrak{g}$ into an unital associative algebra $\mathcal{U}(\mathfrak{g})$ with a map $i : \mathfrak{g} \hookrightarrow \mathcal{U}(\mathfrak{g})$, such that every representation $\rho : \mathfrak{g} \to \mathrm{End}_{\mathbb{F}}(V)$ has a natural unique extension to a $\mathbb{F}$-linear map $\tilde{\rho} : \mathcal{U}(\mathfrak{g}) \to \mathrm{End}_{\mathbb{F}}(V)$. This $\mathcal{U}(\mathfrak{g})$ can be constructed fairly easily, which we will do now.

Definition 2.10. *Let $(\mathfrak{g}, [\cdot,\cdot])$ be a Lie algebra. For all $n \in \mathbb{N}$, let $\mathfrak{g}^{\otimes n}$ be the tensor product of n copies of $\mathfrak{g}$ (in particular $\mathfrak{g}^{\otimes 0} \cong \mathbb{F}$), and define*

$$\mathcal{T}(\mathfrak{g}) := \bigoplus_{n \in \mathbb{N}} \mathfrak{g}^{\otimes n}. \tag{2.6}$$

We will now define a natural multiplication on $\mathcal{T}(\mathfrak{g})$. Let $p, q \in \mathbb{N}$, and let $X = \bigotimes_{i=1}^{p} X_i \in \mathfrak{g}^{\otimes p}$ and $Y = \bigotimes_{j=1}^{q} Y_j \in \mathfrak{g}^{\otimes q}$. We then define a multiplication $\cdot : \mathfrak{g}^{\otimes p} \times \mathfrak{g}^{\otimes q} \to \mathfrak{g}^{\otimes (p+q)}$ by:

$$X \cdot Y = \left(\bigotimes_{i=1}^{p} X_i \right) \cdot \left(\bigotimes_{j=1}^{q} Y_j \right) := \bigotimes_{k=1}^{p+q} Z_k, \tag{2.7}$$

$$Z_k = \begin{cases} X_k & \text{for } 1 \leq k \leq p \\ Y_{k-p} & \text{for } p < k \leq p+q. \end{cases} \tag{2.8}$$

In other words, we concatenate X and Y. Note that for all $n \in \mathbb{N}$, we may view $\mathfrak{g}^{\otimes n}$ as a subset of $\mathcal{T}(\mathfrak{g})$. This allows us to extend our newly defined multiplication $\cdot$ bilinearly to $\mathcal{T}(\mathfrak{g})$. Since $\bigcup_{n \in \mathbb{N}} \mathfrak{g}^{\otimes n}$ generates $\bigoplus_{n \in \mathbb{N}} \mathfrak{g}^{\otimes n} = \mathcal{T}(\mathfrak{g})$, we have now defined $\cdot$ on the entirety of $\mathcal{T}(\mathfrak{g})$. Note $\cdot$ is associative, and has unit

$1 \in \mathfrak{g}^{\otimes 0}$. *This makes* $\mathcal{T}(\mathfrak{g})$ *into a unital associative algebra, which we will call the* ***tensor algebra*** *of* $\mathfrak{g}$. *We now define* S *to be the ideal generated by the set* $\{X \otimes Y - Y \otimes X - [X,Y] : X, Y \in \mathfrak{g} = \mathfrak{g}^{\otimes 1}\}$, *and we define the* ***universal enveloping algebra*** $\mathcal{U}(\mathfrak{g})$ *as*

$$\mathcal{U}(\mathfrak{g}) := \mathcal{T}(\mathfrak{g})/S. \tag{2.9}$$

We write the elements of $\mathcal{U}(\mathfrak{g})$ *as*

$$\left\lfloor \bigotimes_{i=1}^{p} X_i \right\rceil =: X_1 X_2 \dots X_p. \tag{2.10}$$

This definition might seem very complicated and arbitrary at first glance, but the constructed algebra $\mathcal{U}(\mathfrak{g})$ has exactly the universal property mentioned before. This is summarised in the following theorem.

Theorem 2.11. *Let* $(\mathfrak{g}, [\cdot,\cdot])$ *be a Lie algebra and* $i : \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ *the natural embedding. Then* $\mathcal{U}(\mathfrak{g})$ *has the following universal property: for every unital associative algebra* A *and every Lie algebra homomorphism*[†] $\phi : \mathfrak{g} \to A$ *there exists a unique algebra homomorphism* $\overline{\phi} : \mathcal{U}(\mathfrak{g}) \to A$ *which sends* $1 \in \mathcal{U}(\mathfrak{g})$ *to* $1 \in A$ *and satisfies* $\overline{\phi} \circ i = \phi$.

Proof. Let A be an unital associative algebra, and let $\phi : \mathfrak{g} \to A$ be a Lie algebra homomorphism. Let $q : \mathcal{T}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$ denote the quotient map. First of all, we define $j : \mathfrak{g} \to \mathcal{T}(\mathfrak{g})$ by $X \mapsto X \in \mathfrak{g}^{\otimes 1}$, where we view $\mathfrak{g}^{\otimes 1}$ as a subset of $\mathcal{T}(\mathfrak{g})$. Note j is injective, and since no non-trivial element of $\mathfrak{g} = j(\mathfrak{g})$ is contained in S, so is $i := q \circ j$. Furthermore, since both q and j are linear, so is i. Also note $\mathfrak{g}^{\otimes 1}$ generates $\mathcal{T}(\mathfrak{g})$ in a multiplicative sense (allowing addition, multiplication and scalar multiplication). This allows us to define the natural extension of ϕ, namely $\tilde{\phi} : \mathcal{T}(\mathfrak{g}) \to A$ which is determined by the identity

$$\tilde{\phi}\left(\bigotimes_{i=1}^{p} X_i\right) = \prod_{i=1}^{p} \phi(X_i), \tag{2.11}$$

and extended linearly. In particular it maps $1 \in \mathcal{T}(\mathfrak{g})$ to $1 \in A$. It is easy to see that $\tilde{\phi} \circ j = \phi$. Now for any $X, Y \in \mathfrak{g}$ we have:

$$\begin{aligned} &\tilde{\phi}(X \otimes Y - Y \otimes X - [X,Y]) \\ &= \tilde{\phi}(X \otimes Y) - \tilde{\phi}(Y \otimes X) - \tilde{\phi}([X,Y]) \\ &= \tilde{\phi}(X)\tilde{\phi}(Y) - \tilde{\phi}(Y)\tilde{\phi}X) - (\tilde{\phi}(X)\tilde{\phi}(Y) - \tilde{\phi}(Y)\tilde{\phi}X)) = 0. \end{aligned} \tag{2.12}$$

[†]Here we view A as a Lie algebra as shown in example 2.7.(ii)

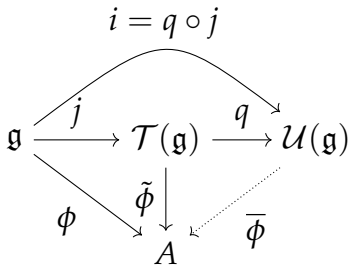

Figure 2.1: *Commutative diagram of relevant maps.*

Hence the set $\{X \otimes Y - Y \otimes X - [X,Y] : X, Y \in \mathfrak{g} = \mathfrak{g}^{\otimes 1}\}$ is contained in $\ker \tilde{\phi}$, so S is as well. Hence, by the isomorphism theorem (theorem 2.5), there exists a unique algebra homomorphism $\overline{\phi} : \mathcal{U}(\mathfrak{g}) \to A$ satisfying $\overline{\phi} \circ q = \tilde{\phi}$. Therefore we have

$$\bar{\phi} \circ i = \bar{\phi} \circ q \circ j = \tilde{\phi} \circ j = \phi. \tag{2.13}$$

Finally, to see the unicity of $\bar{\phi}$, note $\tilde{\phi}$ is the unique map such that $\tilde{\phi} \circ j = \phi$, as there is no choice for the behaviour of $\tilde{\phi}$ on $\mathfrak{g}$. Any $\bar{\phi}'$ satisfying $\bar{\phi}' \circ i = \phi$ would give us $\bar{\phi}' \circ q \circ j = \phi$, and hence $\tilde{\phi} = \bar{\phi}' \circ q$. $\bar{\phi}$ is the unique map satisfying this last property, and is therefore unique entirely.

We conclude $\bar{\phi}$ both exists and is unique.

□

The most important consequence of theorem 2.11 is the fact that any representation (V, ρ) of a Lie algebra $(\mathfrak{g}, [\cdot,\cdot])$ automatically yields a map $\bar{\rho} : \mathcal{U}(\mathfrak{g}) \to \mathrm{End}(V)$. This allows us to transition between $\mathfrak{g}$ representations and $\mathcal{U}(\mathfrak{g})$ representations without any problems. We will now state a useful theorem concerning the universal enveloping algebra, and partially proof it. It is the Poincaré-Birkhoff-Witt theorem, which gives an easy way to construct a basis for such an algebra.

Theorem 2.12. *(The Poincaré-Birkhoff-Witt theorem) Let $(\mathfrak{g}, [\cdot,\cdot])$ be a Lie algebra, and let $B = (b_i)_{i \in I}$ be a totally ordered basis of $\mathfrak{g}$ (that is, B is totally ordered by some total ordering $\preceq$). Then the set*

$$S(B) := \{b_{n_1} b_{n_2} \dots b_{n_m} : m \in \mathbb{N}, n_1, \dots, n_m \in I, b_{n_1} \preceq b_{n_2} \preceq \dots \preceq b_{n_m}\} \tag{2.14}$$

forms a basis of $\mathcal{U}(\mathfrak{g})$.

Proof. We will show $S(B)$ generates $\mathcal{U}(B)$. A full proof which also shows all elements of $S(B)$ are also linearly independent can be found in [Hum12]. By definition of $\mathcal{T}(\mathfrak{g})$, the set

$$U(B) := \left\{ \bigotimes_{i=1}^{m} b_{n_i} : m \in \mathbb{N}, \text{ and } n_1, \ldots, n_m \in I \right\} \subseteq \mathcal{T}(\mathfrak{g}) \tag{2.15}$$

of unordered products of basis vectors generates $\mathcal{T}(\mathfrak{g})$. Hence, the images of these products under the quotient map generate $\mathcal{U}(\mathfrak{g})$. It is therefore sufficient to write an element of $U(B)$ as a linear combination of elements of $S(B)$. So let $m \in \mathbb{N}$, and let $n_1, \ldots, n_m \in I$. We continue the proof by induction on m.

- For $m = 0$ or $m = 1$ there is nothing to prove, as the corresponding empty product or single element is certainly ordered.
- Now assume the statement holds for $m - 1$. That is, the class of every (not necessarily ordered) product in $U(B)$ of length $m - 1$ can be written in terms of the ordered products of $S(B)$. Our task is now to write our arbitrary product $b := b_{n_1} b_{n_2} \ldots b_{n_m}$ of length m as a linear combination of elements of $S(B)$. Note that if the Lie bracket on $\mathfrak{g}$ would be the zero map, then this certainly would not be a problem. Indeed, because for all $X, Y \in \mathfrak{g}$ we have $X \otimes Y - Y \otimes X \in S$ (with S defined as above (2.9)). Hence, $XY = YX \in \mathcal{U}(\mathfrak{g})$, so $\mathcal{U}(\mathfrak{g})$ is commutative and we can simply swap around all b_i until we are done. In general, the Lie bracket is not zero and the process is more complicated, since swapping two adjacent terms results in a commutator term. But still, since we are dealing with finite products, we will only need a finite amount of swaps (of adjecent terms) to arrive at a fully ordered product. Let N denote the least amount of swaps we need. We continue with induction on N.
  - For $N = 0$ we do not need any swaps, so we are done.
  - Assume any product of length m for which less than N swaps will suffice to order it can be written in terms of elements of $S(B)$. Now we look back at the product $b = b_{n_1} b_{n_2} \ldots b_{n_m}$ we started with. Since we can order this product by N swaps, there is an index $j \in \{1, 2, \ldots, m-1\}$ such that the product $b_{n_1} \ldots b_{n_{j+1}} b_{n_j} \ldots b_{n_m}$ only requires $N - 1$ swaps. Now note we have

$$\begin{aligned} b &= b_{n_1} \ldots b_{n_j} b_{n_{j+1}} \ldots b_{n_m} \\ &= b_{n_1} \ldots b_{n_{j-1}} \left(b_{n_{j+1}} b_{n_j} + [b_{n_j}, b_{n_{j+1}}] \right) b_{n_{j+2}} \ldots b_{n_m}. \end{aligned} \tag{2.16}$$

Since $[b_{n_j}, b_{n_{j+1}}] \in \mathfrak{g}$, we can write it in terms of the basis $B = (b_i)_{i\in I}$ as

$$[b_{n_i}, b_{n_{i+1}}] = \sum_{i\in I} \lambda_i b_i, \tag{2.17}$$

with all $\lambda_i \in \mathbb{F}$, all but finitely many equal to zero. Thus it follows

$$\begin{aligned} b &= b_{n_1} \dots b_{n_{j-1}} \left(b_{n_{j+1}} b_{n_j} + [b_{n_j}, b_{n_{j+1}}] \right) b_{n_{j+2}} \dots b_{n_m} \\ &= b_{n_1} \dots b_{n_{j-1}} \left(b_{n_{j+1}} b_{n_j} + \sum_{i\in I} \lambda_i b_i \right) b_{n_{j+2}} \dots b_{n_m} \\ &= \underbrace{b_{n_1} \dots b_{n_{j-1}} b_{n_{j+1}} b_{n_j} b_{n_{j+2}} \dots b_{n_m}}_{\text{only } N-1 \text{ swaps required}} + \sum_{i\in I} \lambda_i \underbrace{b_{n_1} \dots b_{n_{j-1}} b_i b_{n_{j+2}} \dots b_{n_m}}_{\text{only } m-1 \text{ terms}}. \end{aligned} \tag{2.18}$$

Now the first term only requires $N-1$ swaps to be fully ordered, and can thus be written as a linear combination of elements of $S(B)$, by our second induction hypothesis. All other terms (the ones grouped within the summation) are products of only $m-1$ terms, as b_{n_j} and $b_{n_{j+1}}$ have been swapped out in favour of some b_i. Hence all these terms can also be written in terms of elements of $S(B)$, and consequently b can as well.

This concludes both our induction proofs, and therefore the proof as a whole. □

To wind up this section, we will prove one more lemma from representation theory of Lie algebras that will come into use later. It is known as Schur's lemma ([Hum12]).

Lemma 2.13. *(Schur's lemma) Let $(\mathfrak{g}, [\cdot,\cdot])$ be a Lie algebra, and let (V,ρ),(W,σ) be irreducible representations of $\mathfrak{g}$. Let $\phi : V \to W$ be a homomorphism of representations, that is, ϕ is a linear map such that for all $X \in \mathfrak{g}$ and for all $v \in V$ we have $\phi(\rho(X)(v)) = \sigma(X)(\phi(v))$. Then the following statements hold:*

1. *Either $\phi = 0$ or ϕ is an isomorphism of representations (that is, it is a bijective representation homomorphism).*
2. *If $(V,\rho) = (W,\sigma)$, and if ϕ (which is now an endomorphism) has at least one eigenvalue $\lambda \in \mathbb{F}$, then $\rho(\phi) = \lambda \,\mathrm{Id}$.*

Proof. 1. The most important point for our proof is the fact that both V and W are assumed to be irreducible representations. This is important, since both the image and the kernel of ϕ are in fact subrepresentations. To see this, let $v \in \ker \phi$. Then for all $X \in \mathfrak{g}$:

$$\phi(\rho(X)(v)) = \sigma(X)(\phi(v)) = \sigma(X)(0) = 0 \tag{2.19}$$

Hence $\rho(X)(v) \in \ker \phi$. This means the linear subspace $\ker \phi$ is closed under the action of $\mathfrak{g}$, and it is therefore a subrepresentation. Since V is assumed to be irreducible it is therefore either equal to V or to 0, and ϕ is therefore either 0 or injective.

In a similar fashion, any element w in the image of ϕ can be written as $w = \phi(v)$ for some $v \in V$. Now for $X \in \mathfrak{g}$:

$$\sigma(X)(w) = \sigma(X)(\phi(v)) = \phi(\rho(X)(v)) \tag{2.20}$$

And hence $\sigma(X)(w) \in \operatorname{im} \phi$, and we can again conclude $\operatorname{im} \phi$ is a subrepresentation. Since W is irreducible, this means $\operatorname{im} \phi$ is either equal to 0 or to W, and hence ϕ is either 0 or surjective.

Note that if ϕ is zero, it cannot be injective or surjective, since this would imply either $V = 0$ or $W = 0$, but this is ruled out by the fact V and W are irreducible (the zero space is not irreducible since it has only one subrepresentation). Hence, ϕ is either zero or both injective and surjective, and thus an isomorphism.

2. Assume $(V, \rho) = (W, \sigma)$, and assume ϕ has at least one eigenvalue λ. From part 1 of this lemma, ϕ is either zero or a bijection. If ϕ is the zero map, it is a multiple of the identity, so here there is nothing to prove. Now let $v \in V$ be an eigenvector of ϕ with eigenvalue λ. Then for all $X \in \mathfrak{g}$ we have:

$$\phi(\rho(X)(v)) = \rho(X)(\phi(v)) = \rho(X)(\lambda v) = \lambda \rho(X)(v) \tag{2.21}$$

Hence $\rho(X)(v)$ is also an eigenvector of ϕ, with the same eigenvalue. In other words, the eigenspace of ϕ associated with λ is closed under the action of $\mathfrak{g}$, and it is therefore a subrepresentation of V! Since we assumed ϕ is not the zero map, this means this eigenspace is whole of V, which directly proves $\phi = \lambda \operatorname{Id}$.

□

It is worth noting that the condition that ϕ has at least one eigenvalue is automatically satisfied in the case that V is finite dimensional and $\mathbb{F}$ is

algebraically closed. This is because if V is finite dimensional, then ϕ has a characteristic polynomial. The fact that $\mathbb{F}$ is algebraically closed then immediately implies this characteristic polynomial has a solution, and this solution is exactly an eigenvalue.

Remark. Just like in the case of group representations, the notation is often shortened. When working with a certain fixed representation, the ρ is suppressed and for $v \in V$ we simply write Xv instead of $\rho(X)(v)$, as if X acted directly on V.

This finishes our discussion on Lie algebras in general. From here on out, we will focus on the Lie algebras corresponding to conformal symmetries.

2.3 The Witt and Virasoro algebras

To find the Lie algebra of conformal transformations, we have to find the infinitesimal generators of such transformations. Just as the quantum Hamiltonian generates time translation through $e^{iaH}|\phi(t)\rangle = |\phi(t+a)\rangle$, the infinitesimal generator T of some family $\{h_s\}$ of conformal transformations should satisfy $e^{sT}f(x) = f(h_s(x))$. Recall conformal transformations are exactly the holomorphic and anti-holomorphic maps on $\mathbb{C}$. We now focus on the holomorphic ones. A holomorphic map can always locally be expressed as a (convergent) Laurent series in z, which can be approximated arbitrarily well by a polynomial $p(z, z^{-1})$ as $z \mapsto z + p(z, z^{-1})$. We therefore look for the generators of transformations of the form $z \mapsto z + az^n$ (with $n \in \mathbb{Z}$), as these transformations in turn generate such polynomials. We claim the generators are simply given by $l_n := -z^{n+1}\partial_z$. Namely, for some smooth test function f with a locally convergent Taylor series, we have

$$\begin{aligned} f(z - az^{n+1}) &= \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(z)(-az^{n+1})^k = \sum_{k=0}^{\infty} \frac{(-az^{n+1}\partial_z)^k}{k!} f(z) \\ &= \exp(-az^{n+1}\partial_z) f(z) = \exp(al_n) f(z). \end{aligned}$$

This shows the l_n indeed generate holomorphic transformations. Moreover, we can explicitly calculate the commutation relations of the l_n. For

$n, m \in \mathbb{Z}$ and a test function f we namely have

$$\begin{aligned}
[\ell_n, \ell_m]f = (\ell_n\ell_m - \ell_m\ell_n)f &= z^{n+1}\frac{\partial}{\partial z}\left(z^{m+1}\frac{\partial f}{\partial z}\right) - z^{m+1}\frac{\partial}{\partial z}\left(z^{n+1}\frac{\partial f}{\partial z}\right) \\
&= z^{n+1}\left((m+1)z^m\frac{\partial f}{\partial z} + z^{m+1}\frac{\partial^2 f}{\partial z^2}\right) \\
&\qquad - z^{m+1}\left((n+1)z^m\frac{\partial f}{\partial z} + z^{n+1}\frac{\partial^2 f}{\partial z^2}\right) \\
&= ((m+1)-(n+1))z^{n+m+1}\frac{\partial f}{\partial z} + z^{n+m+2}\frac{\partial^2 f}{\partial z^2} - z^{n+m+2}\frac{\partial^2 f}{\partial z^2} \\
&= -(n-m)z^{n+m+1}\frac{\partial f}{\partial z} = (n-m)\ell_{n+m}f,
\end{aligned} \tag{2.22}$$

and hence $[l_n, l_m] = (n-m)l_{n+m}$. This shows the set of infinitesimal generators is indeed closed under the commutator. Since the commutator is also clearly alternating and since it obeys the Jacobi identity (commutators defined by associative products in general do), it forms a Lie algebra. The complexification[‡] of this Lie algebra is known as the Witt algebra.

Definition 2.14. *The **Witt algebra** is the complex Lie algebra whose vector space $\mathfrak{W}$ is generated by the set $\{\ell_n\}_{n\in\mathbb{Z}}$, with Lie bracket given by*

$$[l_n, l_m] = (n-m)l_{n+m}.$$

Similarly, from the antiholomorphic conformal transformations we get another set of infinitesimal generators, which are given by $\bar{l}_n := -\bar{z}^{n+1}\partial_{\bar{z}}$. These generators have the same commutation relations as the l_n, and therefore also form a copy of the Witt algebra. However, since these two copies commute with each other and thereby act independently, it is often sufficient to study only the holomorphic part.

When transitioning to a quantum theory, it is not the Witt algebra but a central extension of the Witt algebra that is encountered. We first state what we mean by a central extension [Sch08].

Definition 2.15. *Let $\mathfrak{g}$ be a Lie algebra. A **Lie algebra extension** of $\mathfrak{g}$ by a Lie algebra $\mathfrak{c}$ is a Lie algebra $\mathfrak{h}$ with a short exact sequence* $0 \to \mathfrak{c} \overset{i}{\hookrightarrow} \mathfrak{h} \overset{p}{\twoheadrightarrow} \mathfrak{g} \to 0$,

[‡]Formally, the complexification of a real Lie algebra $\mathbb{R}$ is the complex Lie algebra whose vector space is given by $\mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}$, with Lie bracket $[\lambda \otimes X, \mu \otimes Y] = \lambda\mu \otimes [X, Y]$. If one is given a basis of $\mathfrak{g}$, the complexification can simply be understood as allowing complex coefficients instead of only real ones.

*with i and p Lie algebra homomorphisms. If $i(\mathfrak{c})$ is **central** in $\mathfrak{h}$, i.e. for all $X \in \mathfrak{h}$ and $C \in i(\mathfrak{c})$ we have $[X, C] = 0$, then $\mathfrak{h}$ is called a **central extension** of $\mathfrak{g}$ by $\mathfrak{c}$. If $\mathfrak{h} \cong \mathfrak{c} \oplus \mathfrak{g}$ as Lie algebras, then this extension is called **trivial**.*

To gain some intuition as to why we would look at a central extension, assume we have some set of symmetries G of a quantum system, which is represented by a Hilbert space $\mathbb{H}$. One might think this means there is some unitary action of G on $\mathbb{H}$. However, the relevant objects within this Hilbert space are not simply its elements, but more so the *rays* of $\mathbb{H}$, which are one dimensional subspaces of $\mathbb{H}$ and by definition precisely the elements of the projective space $\mathbb{P}(\mathbb{H})$. This space comes with an additional structure, namely a way to calculate transition probabilities. Given two rays $|\phi\rangle, |\phi\rangle$ their transition probability namely is given by

$$\text{chance}(|\phi\rangle \to |\psi\rangle) = \frac{|\langle\phi|\psi\rangle|}{\sqrt{\langle\phi|\phi\rangle\langle\psi|\psi\rangle}}, \tag{2.23}$$

which is indeed a well-defined operation on $\mathbb{P}(\mathbb{H})$. If we want G to act as a set of symmetries on $\mathbb{P}(\mathbb{H})$, its elements should respect this transition probability, so each $f \in G$ acts as a map $f : \mathbb{P}(\mathbb{H}) \to \mathbb{P}(\mathbb{H})$ such that for all $|\phi\rangle, |\phi\rangle$ we have

$$\text{chance}(f\,|\phi\rangle \to f\,|\psi\rangle) = \text{chance}(|\phi\rangle \to |\psi\rangle). \tag{2.24}$$

By Wigner's theorem ([Bar64]), such a map has a lift to a map $\hat{f} : \mathbb{H} \to \mathbb{H}$, which is unique up to multiplication with some complex phase factor. This lift has to behave nicely under composition, which means the following: Given two general f, f' acting on $\mathbb{P}(\mathbb{H})$, we can either take the composition $f \circ f'$, or we can choose to first lift them to $\hat{f}, \hat{f}'$, then take the composition $\hat{f} \circ \hat{f}'$ of these two and then let this new map act on $\mathbb{P}(\mathbb{H})$. These two maps on $\mathbb{P}(\mathbb{H})$ (which are $f \circ f'$ and the projection of $\hat{f} \circ \hat{f}'$) should be identical, so $\widehat{f \circ f'}$ and $\hat{f} \circ \hat{f}'$ can at most differ by a phase factor. The effect of this on the symmetry Lie algebra is that we can *lift* the Lie algebra of infinitesimal transformations $\mathfrak{g}$ to a slightly bigger algebra $\hat{\mathfrak{g}} := \mathfrak{g} \oplus \mathbb{C}$, which satisfies $[\hat{X}, \hat{Y}] = \widehat{[X, Y]} + \lambda(X, Y)$ for all $X, Y \in \mathfrak{g}$, where $\lambda(X, Y) \in \mathbb{C}$ [Wened]. Hence the sequence

$$0 \to \mathbb{C} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0 \tag{2.25}$$

is exact. To summarise, if a quantum system has a set of infinitesimal symmetries $\mathfrak{g}$, then these symmetries form a projective representation, which is equivalent to an ordinary representation of a central extension $\hat{\mathfrak{g}}$ of $\mathfrak{g}$. This then finally brings us to the Lie algebra which takes the central role in this thesis; the Virasoro algebra.

Definition 2.16. *The **Virasoro algebra with central charge** $\mathbf{c} \in \mathbb{C}^*$ is the Lie algebra whose vector space is given by*

$$\mathfrak{V}_c := \mathrm{Span}_{\mathbb{C}}(\{L_n\}_{n\in\mathbb{Z}} \cup \{I\}),$$

where the Lie bracket is defined by

$$[L_n, L_m] = (n-m)L_{n+m} - \frac{c}{12}(m^3 - m)\delta_{n+m,0} I \quad \textit{and} \quad [L_n, I] = 0, \tag{2.26}$$

and extended linearly.

Note we have not shown (2.26) indeed defines a Lie algebra, and it is a priori not obvious it satisfies the Jacobi identity. The calculation however is more tedious than interesting, so we will not perform it here. Moreover, although we have defined a Virasoro algebra for every $c \in \mathbb{C}^*$, they are in fact all isomorphic. This can be seen by explicitly defining an isomorphism $\phi_c : \mathfrak{V}_c \to \mathfrak{V}_1$ by

$$\phi_c(L_n) = L_n \qquad \text{and} \qquad \phi_c(I) = \frac{I}{c},$$

and extending linearly. ϕ_c is clearly linear and bijective. Moreover, if we let $[\cdot,\cdot]_c$ denote the bracket on $\mathfrak{V}_c$, for $n, m \in \mathbb{Z}$ we have

$$\begin{aligned}
\phi_c([L_n, L_m]_c) &= \phi_c\left((n-m)L_{n+m} - \frac{c}{12}(m^3-m)\delta_{n+m,0} I\right) \\
&= (n-m)L_{n+m} - \frac{c}{12}(m^3-m)\delta_{n+m,0}\phi_c(I) \\
&= (n-m)L_{n+m} - \frac{1}{12}(m^3-m)\delta_{n+m,0} I \\
&= \phi_c([L_n, L_m]_1).
\end{aligned}$$

Hence ϕ_c is a Lie algebra isomorphism. One could argue there is no need to include a central charge in the definition of the Virasoro algebra, and simply always work with $\mathfrak{V}_1$ instead. We will however not do this, as in representations of the Virasoro algebra it is customary to let I act as the identity. The central charge allows us to without loss of generality to always assume this is the case.

As promised, the Virasoro algebra is a central extension of the Witt algebra, which we will now prove. Furthermore, it can be shown it is up to isomorphism, the only nontrivial central extension of the Witt algebra.

Proposition 2.17. *For every $c \in \mathbb{C}$, the Virasoro algebra with central charge c is a nontrivial central extension of the Witt algebra, and is up to isomorphism the only one.*

Proof. We here only proof the first statement, a proof for uniqueness can be found in [Wened].

Let $c \in \mathbb{C}$. First note $\mathbb{C}$ is a Lie algebra if we define a Lie bracket on $\mathbb{C}$ as the zero map, so for all $z, w \in \mathbb{C}$ we have $[z, w] = 0$. We now construct a short exact sequence $0 \to \mathbb{C} \overset{i}{\hookrightarrow} \mathfrak{V}_c \overset{p}{\twoheadrightarrow} \mathfrak{W} \to 0$ of Lie algebras. For this, we define $i : \mathbb{C} \to \mathfrak{V}_c$ by $z \mapsto zI$, and $p : \mathfrak{V}_c \to \mathfrak{W}$ by $p(I) = 0$ and $p(L_n) = \ell_n$ for all $n \in \mathbb{Z}$, and extent linearly. It is then clear that i is injective, p is surjective, and $p \circ i = 0$. Hence, if i and p are indeed Lie algebra homomorphisms, the given sequence defines a Lie algebra extension. Since both i and p are linear, it suffices to check they satisfy (2.1). Since $\mathbb{C}$ is a 1-dimensional Lie algebra, (2.1) is satisfied automatically, by the alternating property of the Lie bracket, so i is indeed a Lie algebra homomorphism. To check (2.1) for p, it suffices to show it holds for its basis vectors. For all $n \in \mathbb{Z}$ clearly $p([I, L_n]) = p(0) = 0 = [0, \ell_n] = [p(I), p(L_n)]$. Furthermore, for $n, m \in \mathbb{Z}$ we have

$$\begin{aligned} p([L_n, L_m]) &= p\left((n-m)L_{n+m} - \frac{c}{12}(m^3 - m)\delta_{n+m,0} I\right) \\ &= (n-m)p(L_{n+m}) - \frac{c}{12}(m^3 - m)\delta_{n+m,0} p(I) \\ &= (n-m)\ell_{n+m} = [\ell_n, \ell_m] = [p(L_n), p(L_m)]. \end{aligned} \tag{2.27}$$

We conclude p satisfies (2.1), and is therefore a Lie algebra homomorphism. Hence the given sequence is indeed a short exact sequence of Lie algebra homomorphisms, and therefor $\mathfrak{V}_c$ is a Lie algebra extension of $\mathfrak{W}$ by $\mathbb{C}$. Furthermore, because the image of i is just the subspace spanned by I, the image of i is central, so $\mathfrak{V}_c$ is a central extension of $\mathfrak{W}$ by $\mathbb{C}$. □

As stated before, the quantum theory of a system with conformal symmetry is a representation of a central extension of the Witt algebra, and by the previous proposition, we can assume it is the Virasoro algebra. In the next chapter we will visit a concrete example of this principle, namely the theory of bosonic string theory.

Chapter 3

Free bosonic string theory

In this chapter we study the free bosonic string, and we will find the quantum theory indeed forms a representation of the Virasoro algebra. We first discuss the theory of the classical string, and subsequently its quantum theory-. Most of the discussion in this chapter is based of the work of [Zwi04].

3.1 The free string

In line with the the previously given summary of Lagrangian mechanics, we will use a Lagrangian to study the behaviour of a string. Here we will use parameterisations, which in the string context take the following form.

Definition 3.1. *A **string parameterisation** is a differentiable and injective map $X : \mathbb{R} \times [0, \pi] \to \mathbb{R}^{(1,d)}$ such that:*

- *$\partial_\tau X(\tau, \sigma)$ is either timelike or lightlike for all $\tau \in \mathbb{R}$ and $\sigma \in [0, \pi]$ (meaning our the directions in our parameter space indeed represent time and space directions in the target space);*
- *and for all $\tau \in \mathbb{R}$ the image of X intersected with the plane $\{\tau\} \times \mathbb{R}^{d-1}$ is homeomorphic with C (meaning the string does exist at all times, and the string never self-intersects or changes from an open string to a closed string). We will call this set the **string at time** τ.*

The image of such a string parameterisation is called a **world sheet**. In order to realise our string as a physical object, we have to choose a proper Lagrangian. Inspired by the relativistic point particle, whose action was

just the proper time along the worldline, we choose the action of the string to be the proper surface of the world sheet. Introducing the notation $\dot{X} = \frac{\partial X}{\partial \tau}$ and $X' = \frac{\partial X}{\partial \sigma}$, this means we set

$$\begin{aligned} S &:= -\frac{1}{2\pi\alpha'} \int_{\tau \in T} d\tau \int_{\sigma \in C} d\sigma \sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2} \\ &= -\frac{1}{2\pi\alpha'} \int_{\tau \in T} d\tau \int_{\sigma \in C} d\sigma \mathcal{L} \end{aligned} \tag{3.1}$$

with

$$\mathcal{L} := -\frac{1}{2\pi\alpha'} \sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}. \tag{3.2}$$

Here the constant α' called the *slope parameter* has units of inverse energy squared, such that the action is dimensionless in natural units. It will be useful to introduce the canonical momenta $\mathcal{P}^\tau$ and $\mathcal{P}^\sigma$ as

$$\mathcal{P}^\tau := \frac{\partial \mathcal{L}}{\partial \dot{X}} \qquad \text{and} \qquad \mathcal{P}^\sigma := \frac{\partial \mathcal{L}}{\partial X'}. \tag{3.3}$$

These momenta can be explicitly calculated and are equal to

$$\mathcal{P}^\tau = -\frac{1}{2\pi\alpha'} \frac{(\dot{X} \cdot X') X' - (X')^2 \dot{X}}{\sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}} \tag{3.4}$$

and

$$\mathcal{P}^\sigma = -\frac{1}{2\pi\alpha'} \frac{(\dot{X} \cdot X') \dot{X} - (\dot{X})^2 X'}{\sqrt{(\dot{X} \cdot X')^2 - (\dot{X})^2 (X')^2}}. \tag{3.5}$$

Moreover, we will impose *free endpoint boundary conditions*, which say $\mathcal{P}^\sigma$ vanishes on the endpoints of the string. This condition allows the string endpoints to move freely, which makes sure the momentum of the string is conserved, as we will show shortly. Now any parameterisation X for which this action is stationary satisfies the Euler-Lagrange equation (1.22), which in our case of the free string is given by

$$\frac{\partial \mathcal{P}^\tau}{\partial \tau} + \frac{\partial \mathcal{P}^\sigma}{\partial \sigma} = 0. \tag{3.6}$$

One could say we we are now done, as we have found equations of motions. However, as becomes obvious from the complexity of (3.4) and (3.5)

these equations are in their current form near unsolvable and do not actually tell us a lot in their current form. The only point of light in this situation is that these equations of motion hold for *every* parameterisation of the string, which implies we can choose a specific one that helps us simplify the this computational mess. In the following section we will choose a specific parameterisation, and show how it helps solve the equations of motion. To this end, we will first define the string momentum, and prove it is independent of how it is measured. This can be done by writing (3.6) in the more suggestive form of

$$\begin{pmatrix} \frac{\partial}{\partial \tau} \\ \frac{\partial}{\partial \sigma} \end{pmatrix} \times \begin{pmatrix} \mathcal{P}^\sigma \\ -\mathcal{P}^\tau \end{pmatrix} = 0. \tag{3.7}$$

This tells us $(\mathcal{P}^\tau, -\mathcal{P}^\sigma)$ is a rotationless vector field on the parameter space, and therefore for any path γ the line integral

$$p(\gamma) := \int_\gamma \begin{pmatrix} \mathcal{P}^\sigma \\ -\mathcal{P}^\tau \end{pmatrix} \cdot dr = \int_\gamma \mathcal{P}^\sigma d\tau - \int_\gamma \mathcal{P}^\tau d\sigma \tag{3.8}$$

is independent of the precise curve γ, and only dependent on the beginning and endpoints of γ. Moreover, since $\mathcal{P}^\sigma$ vanishes on the world sheet edges, the above integral is zero for any curve lying on such an edge. As a consequence of these two properties, for any two curves γ, δ reaching from one edge of the world sheet to the other, $p(\gamma) = p(\delta)$. This allows us to define the *momentum of the string* as

$$p := \int_\gamma \begin{pmatrix} \mathcal{P}^\sigma \\ -\mathcal{P}^\tau \end{pmatrix} \cdot dr = \int_\gamma \mathcal{P}^\sigma d\tau - \int_\gamma \mathcal{P}^\tau d\sigma, \tag{3.9}$$

where γ is any curve starting on the world sheet edge described by $\sigma = 0$ and reaching to the other edge. We can use this momentum to choose a neat parameterisation of the string, which we will do now.

3.1.1 Useful parameterisations and solving the (classical) equations of motion

We will explicitly work in the reference frame of some Lorentz observer. That is, we fix some timelike vector n such that $n \cdot p \neq 0$, and for every point X on the world sheet we set τ to satisfy

$$n \cdot X = 2\alpha'(n \cdot p)\tau. \tag{3.10}$$

Moreover, we are free to choose the σ parameterisation such that $n \cdot \mathcal{P}^\tau$ is independent of σ. To see this can be done, given some parameterisation X, let $\Gamma : [0, \pi] \to [0, \pi]$ be some reparameterization (meaning it is bijective and differentiable). Then we have

$$\begin{aligned} \frac{\partial}{\partial\sigma} n \cdot \mathcal{P}^\tau(\Gamma(\sigma), \tau) &= n \cdot \frac{\partial}{\partial\sigma} \mathcal{P}^\tau(\Gamma(\sigma), \tau) \\ &= n \cdot \left(\frac{\partial}{\partial\Gamma(\sigma)} \mathcal{P}^\tau(\Gamma(\sigma), \tau) \right) \frac{\partial\Gamma(\sigma)}{\partial\sigma}. \end{aligned} \tag{3.11}$$

Hence by properly choosing such a reparameterisation for every τ we can use $\frac{\partial\Gamma(\sigma)}{\partial\sigma}$ to make sure the last term is constant. Consequently, it is also independent of τ, which can be seen by integrating it over a path of constant τ, as we have

$$\pi n \cdot \mathcal{P}^\tau(\tau, \sigma) = \int_{\sigma=0}^{\pi} n \cdot \mathcal{P}^\tau(\tau, \sigma) = n \cdot \int_{\sigma=0}^{\pi} \mathcal{P}^\tau(\tau, \sigma) = n \cdot p. \tag{3.12}$$

Hence $n \cdot \mathcal{P}^\tau = \frac{n \cdot p}{\pi}$, which is constant. Moreover, by (3.6) we have

$$\frac{\partial}{\partial\sigma} n \cdot \mathcal{P}^\sigma = n \cdot \frac{\partial}{\partial\sigma} \mathcal{P}^\sigma = -\frac{\partial}{\partial\tau} n \cdot \mathcal{P}^\tau = 0, \tag{3.13}$$

and hence $n \cdot \mathcal{P}^\sigma$ is independent of σ. But by assumption of free endpoints, $\mathcal{P}^\sigma$ is zero in both the endpoints, so in particular $n \cdot \mathcal{P}^\sigma = 0$ on the endpoints, and therefore $n \cdot \mathcal{P}^\sigma = 0$ everywhere. By (3.5) this tells us

$$\left(\dot{X} \cdot X'\right) n \cdot \dot{X} - \left(\dot{X}\right)^2 n \cdot X' = 0, \tag{3.14}$$

but by the paramaterisation condition (3.10), we have $n \cdot X' = n \cdot \frac{\partial}{\partial\sigma} X = 0$. Now since we assumed $n \cdot p \neq 0$, (3.10) implies $n \cdot \dot{X}$ is nonzero, and therefore we can conclude

$$\dot{X} \cdot X' = \frac{(\dot{X})^2 n \cdot X'}{n \cdot \dot{X}} = 0. \tag{3.15}$$

This allows us to simplify the expression for $\mathcal{P}^\tau$ to

$$\mathcal{P}^\tau = -\frac{1}{2\pi\alpha'} \frac{(X')^2 \dot{X}}{\sqrt{-(\dot{X})^2 (X')^2}} \tag{3.16}$$

Version of July 11, 2020– Created July 11, 2020 - 11:00

and by (3.12) this tells us

$$\begin{aligned}\frac{n\cdot p}{\pi} &= -\frac{1}{2\pi\alpha'}\frac{(X')^2}{\sqrt{-(\dot{X})^2(X')^2}}n\cdot\dot{X}\\ &= -\frac{1}{2\pi\alpha'}\frac{(X')^2}{\sqrt{-(\dot{X})^2(X')^2}}2\alpha'(n\cdot p),\end{aligned} \tag{3.17}$$

where the second equality follows from our initial parameteristion condition from (3.10). Cancelling common factors, we conclude

$$\sqrt{-(\dot{X})^2(X')^2} = -(X')^2 \tag{3.18}$$

and equivalently

$$(\dot{X})^2 + (X')^2 = 0. \tag{3.19}$$

This allows us to again simplify $\mathcal{P}^\sigma$ and $\mathcal{P}^\tau$, namely as

$$\mathcal{P}^\tau = -\frac{1}{2\pi\alpha'}\frac{(X')^2\dot{X}}{\sqrt{-(\dot{X})^2(X')^2}} = \frac{1}{2\pi\alpha'}\dot{X} \tag{3.20}$$

and

$$\mathcal{P}^\sigma = -\frac{1}{2\pi\alpha'}\frac{(\dot{X})^2X'}{\sqrt{-(\dot{X})^2(X')^2}} = -\frac{1}{2\pi\alpha'}\frac{(\dot{X})^2X'}{\sqrt{((X')^2)^2}} = -\frac{1}{2\pi\alpha'}X'. \tag{3.21}$$

Together with equation (3.6) this tells us

$$\ddot{X} = 2\pi\alpha'\frac{\partial}{\partial\tau}\mathcal{P}^\tau = -2\pi\alpha'\frac{\partial}{\partial\sigma}\mathcal{P}^\sigma = X'', \tag{3.22}$$

and moreover the vanishing of $\mathcal{P}^\sigma$ on the endpoints gives us that for all $\tau\in\mathbb{R}$

$$X'(\tau,0) = X'(\tau,\pi) = 0. \tag{3.23}$$

We conclude the equations of motion can be transformed into the following four equations: the wave equation (from (3.22))

$$X'' - \ddot{X} = 0, \tag{3.24}$$

our two parameterisation conditions (respectively (3.15) and (3.19))

$$\dot{X} \cdot X' = 0 \tag{3.25}$$

and

$$(\dot{X})^2 + (X')^2 = 0, \tag{3.26}$$

and the boundary condition

$$X'(\tau, 0) = X'(\tau, \pi) = 0. \tag{3.27}$$

3.1.2 Solving the equations

In the following section, we will solve the four equations mentioned above, and completely determine the set of all possible solutions. To this end, we begin with the wave equation (3.24). If we perform a change of variables from σ, τ to $\tau + \sigma, \tau - \sigma$, we find

$$\begin{aligned}
\frac{\partial^2 X}{\partial(\tau+\sigma)\partial(\tau-\sigma)} &= \frac{\partial}{\partial(\tau+\sigma)}\left(\frac{\partial X}{\partial \tau}\frac{\partial \tau}{\partial(\tau-\sigma)} + \frac{\partial X}{\partial \sigma}\frac{\partial \sigma}{\partial(\tau-\sigma)}\right) \\
&= \frac{\partial}{\partial(\tau+\sigma)}\left(\frac{\partial X}{\partial \tau} - \frac{\partial X}{\partial \sigma}\right) \\
&= \frac{\partial\left(\frac{\partial X}{\partial \tau} - \frac{\partial X}{\partial \sigma}\right)}{\partial \tau}\frac{\partial \tau}{\partial(\tau+\sigma)} + \frac{\partial\left(\frac{\partial X}{\partial \tau} - \frac{\partial X}{\partial \sigma}\right)}{\partial \sigma}\frac{\partial \sigma}{\partial(\tau+\sigma)} \\
&= \frac{\partial^2 X}{\partial \tau^2} - \frac{\partial^2 X}{\partial \tau \partial \sigma} + \frac{\partial^2 X}{\partial \sigma \partial \tau} - \frac{\partial^2 X}{\partial \sigma^2} = \frac{\partial^2 X}{\partial \tau^2} - \frac{\partial^2 X}{\partial \sigma^2} = 0,
\end{aligned} \tag{3.28}$$

where we used the chain rule twice, and the inverse function theorem. This tells us that $\frac{\partial X}{\partial(\tau-\sigma)}$ is independent of $\tau + \sigma$, so we can write $\frac{\partial X}{\partial(\tau-\sigma)} = G_1(\tau - \sigma)$, for some function $G_1 : \mathbb{R} \times [0, \pi] \to \mathbb{R}^{(1,d)}$. Integrating now gives

$$X(\tau - \sigma, \tau + \sigma) = \int G_1(\tau - \sigma) d(\tau - \sigma) + F(\tau + \sigma), \tag{3.29}$$

and since $\int G_1(\tau - \sigma) d(\tau - \sigma)$ only depends on $\tau - \sigma$, we can define $G(\tau - \sigma) := \int G_1(\tau - \sigma) d(\tau - \sigma)$, and we can therefore write

$$X = F(\tau + \sigma) + G(\tau - \sigma). \tag{3.30}$$

Version of July 11, 2020– Created July 11, 2020 - 11:00

From (3.27) we know $\frac{\partial X}{\partial \sigma}\Big|_{\sigma=0} = 0$, which implies that for all $\tau \in \mathbb{R}$

$$F'(\tau) - G'(\tau) = \frac{\partial X}{\partial \sigma}\Big|_{\sigma=0} = 0. \tag{3.31}$$

So $F' = G'$, so $F = G + a_0$ for some constant vector a_0. If we thus define $f = 2(G + a_0/2)$, then for all σ, τ we find

$$\begin{aligned} X(\tau,\sigma) &= F(\tau+\sigma) + G(\tau-\sigma) = G(\tau+\sigma) + \frac{a_0}{2} + G(\tau-\sigma) + \frac{a_0}{2} \\ &= \frac{1}{2}(f(\tau+\sigma) + f(\tau-\sigma)). \end{aligned} \tag{3.32}$$

The factor 2 in the definition of f is simply a matter of convention that will clean up our work a bit later on.

Moreover, from $\frac{\partial X}{\partial \sigma}\Big|_{\sigma=\pi} = 0$ we find that for all τ

$$f'(\tau+\pi) - f'(\tau-\pi) = 0, \tag{3.33}$$

so f' is periodic with period 2π.

Since f' is periodic with period 2π and smooth, it follows from Fourier's theorem ([Sne95]) we can write

$$f'(x) = c_0 + \sum_{n=1}^{\infty} \left(c_n \cos(nx) + d_n \sin(nx)\right), \tag{3.34}$$

for constants $a_n, b_n \in \mathbb{R}$. Integrating now gives

$$f(x) = f_0 + c_0 x + \sum_{n=1}^{\infty} \left(A_n \cos(nx) + B_n \sin(nx)\right) \tag{3.35}$$

for some other constants f_0 and A_n, B_n. Now using the goniometric formulas

$$\begin{aligned} \cos(A+B) &= \cos(A)\cos(B) - \sin(A)\sin(B), \\ \sin(A+B) &= \sin(A)\cos(B) + \sin(B)\cos(A), \end{aligned}$$

we can find that for every $n \in \mathbb{N}$ we have

$$\begin{aligned} \cos(n(\tau+\sigma)) + \cos(n(\tau-\sigma)) &= 2\cos(n\tau)\cos(n\sigma), \\ \sin(n(\tau+\sigma)) + \sin(n(\tau-\sigma)) &= 2\sin(n\tau)\cos(n\sigma). \end{aligned}$$

Using this, we can substitute (3.35) into (3.32), which yields

$$\begin{aligned} X(\tau,\sigma) &= \frac{1}{2}\left(f(\tau+\sigma)+f(\tau-\sigma)\right) \\ &= f_0 + c_0\tau + \sum_{n=1}^{\infty}\left(A_n\cos(n\tau)+B_n\sin(n\tau)\right)\cos(n\sigma). \end{aligned} \tag{3.36}$$

Here we can determine c_0 already. Namely, by (3.21) $\mathcal{P}^\tau = \frac{1}{2\pi\alpha'}\dot{X}$, so

$$p = \int_{\sigma=0}^{\pi}\mathcal{P}^\tau = \int_{\sigma=0}^{\pi}\frac{1}{2\pi\alpha'}\dot{X} = \frac{1}{2\alpha'}c_0, \tag{3.37}$$

where we used that the integral over any term $\sin(n\sigma)$ or $\cos(n\sigma)$ vanishes. Hence $c_0 = 2\alpha' p$, and we find

$$X(\tau,\sigma) = f_0 + 2\alpha' p\tau + \sum_{n=1}^{\infty}\left(A_n\cos(n\tau)+B_n\sin(n\tau)\right)\cos(n\sigma). \tag{3.38}$$

We will rewrite this into a more common form

$$\begin{aligned} X(\tau,\sigma) &= f_0 + 2\alpha' p\tau + \sum_{n=1}^{\infty}\left(A_n\cos(n\tau)+B_n\sin(n\tau)\right)\cos(n\sigma) \\ &= f_0 + 2\alpha' p\tau + \frac{1}{2}\sum_{n=1}^{\infty}\left(A_n(e^{in\tau}+e^{-in\tau}) + \frac{B_n}{i}(e^{in\tau}-e^{-in\tau})\right)\cos(n\sigma) \\ &= f_0 + 2\alpha' p\tau - \frac{i}{2}\sum_{n=1}^{\infty}\left((B_n+iA_n)e^{in\tau} - (B_n-iA_n)e^{-in\tau}\right)\cos(n\sigma) \\ &= f_0 + 2\alpha' p\tau - \frac{i\sqrt{2\alpha'}}{\sqrt{n}}\sum_{n=1}^{\infty}\left(a_n^* e^{in\tau} - a_n e^{-in\tau}\right)\cos(n\sigma), \end{aligned} \tag{3.39}$$

where the constants a_n are defined appropriately. Lastly, for each $n \in \mathbb{Z}$ we will introduce a constant α_n, defined by

$$\alpha_n = \begin{cases} a_n\sqrt{n} & \text{for } n > 0 \\ \sqrt{2\alpha' p} & \text{for } n = 0 \\ a_{-n}^*\sqrt{-n} & \text{for } n < 0. \end{cases} \tag{3.40}$$

Note this implies $\alpha_n^* = \alpha_{-n}$. We can rewrite $X(\tau,\sigma)$ in terms of these α_n, namely

$$X(\tau,\sigma) = f_0 + \sqrt{2\alpha'}\alpha_0\tau + i\sqrt{2\alpha'}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{1}{n}\alpha_n e^{-in\tau}\cos(n\sigma), \tag{3.41}$$

Version of July 11, 2020– Created July 11, 2020 - 11:00

and moreover, we have

$$\dot{X} \pm X' = \sqrt{2\alpha'} \sum_{n \in \mathbb{Z}} \alpha_n e^{-in(\tau \pm \sigma)}. \tag{3.42}$$

This is the most general solution to the wave equation (3.24). However, it generally does not satisfy the constraints (3.25) and (3.26). In order to find the solutions which also satisfy these constraints, we will work in the so called *light-cone gauge*. This means we make a specific choice for the vector n.

3.1.3 The light-cone gauge

The light-cone gauge is defined by fixing the vector n to be the light-like vector $(1/\sqrt{2}, 1/\sqrt{2}, 0, \ldots, 0)$. To make our lives easier, we will start working in light-cone coordinates. This means that instead of writing a vector as $(x^0, x^1, ..., x^d)$, we write it as $(x^+, x^-, x^2, ..., x^d)$, where $x^+ = (x^0 + x^1)/\sqrt{2}$ and $x^0 = (x^0 - x^1)/\sqrt{2}$. To avoid cluttering notation we will introduce the notation $x^{\mathcal{I}} = (x^2, ..., x^d)$ for the tail of this vector, which means $x = (x^+, x^-, x^{\mathcal{I}})$. We call $x^{\mathcal{I}}$ the *transverse* part of x. It can be easily seen that for two spacetime vectors $x, y \in \mathbb{R}^{(1,d)}$ we have

$$x \cdot y = -x^+ y^- - x^- y^+ + \sum_{i=2}^{d} x^i y^i = -x^+ y^- - x^- y^+ + x^{\mathcal{I}} \cdot y^{\mathcal{I}}, \tag{3.43}$$

where the inner product between $x^{\mathcal{I}}$ and $y^{\mathcal{I}}$ is just the regular Euclidean inner product. In particular, for any vector $x = (x^+, x^-, x^{\mathcal{I}})$ we have

$$x^2 = -2x^+ x^- + (x^{\mathcal{I}})^2. \tag{3.44}$$

Returning to the task at hand, note that since $X^+ = n \cdot X$, our first parameterisation condition (3.10) implies $X^+ = \alpha' p^+ \tau$. From (3.25) and (3.26) we can now deduce that

$$\left(X' \pm \dot{X}\right)^2 = 0. \tag{3.45}$$

In lightcone coordinates, this means

$$0 = (\dot{X} \pm X')^2 = -2(\dot{X}^+ \pm X'^+)(\dot{X}^- \pm X'^-) + (\dot{X}^{\mathcal{I}} \pm X^{\mathcal{I}\prime})^2 \tag{3.46}$$

and since $X'^+ = 0$ and $2\dot{X}^+ = \alpha' p^+$, this tells us that *

$$\dot{X}^- \pm X'^- = \frac{(\dot{X}^{\mathcal{I}} \pm X^{\mathcal{I}\prime})^2}{2p^+\alpha'}. \tag{3.47}$$

*Note we can safely assume the momentum of our string is not light-like, and hence $p^+ \neq 0$.

Hence, in order to find X^- (up to some constant), it suffices to know how $X^\mathcal{I}$ behaves. We can therefore conclude that the function $X^\mathcal{I}(\tau,\sigma)$, together with the constants p^+ and x_0^-, fully determines X! This can explicitly be done by using the previous expression for $\dot{X}^- \pm X'^-$ from (3.42):

$$\begin{aligned}\sqrt{2\alpha'}\sum_{n\in\mathbb{Z}}\alpha_n^- e^{-in(\tau\pm\sigma)} = \dot{X}^- \pm X'^- &= \frac{(\dot{X}^\mathcal{I} \pm X^{\mathcal{I}\prime})^2}{2p^+\alpha'} \\ &= \frac{1}{p^+}\left(\sum_{n\in\mathbb{Z}}\alpha_n^\mathcal{I} e^{-in(\tau\pm\sigma)}\right)^2 \\ &= \frac{1}{p^+}\sum_{n\in\mathbb{Z}}\left(\sum_{m\in\mathbb{Z}}\alpha_{n-m}^\mathcal{I}\alpha_m^\mathcal{I}\right)e^{-in(\tau\pm\sigma)}.\end{aligned} \tag{3.48}$$

Hence we find that for all $n \in \mathbb{Z}$

$$\alpha_n^- = \frac{1}{\sqrt{2\alpha'}p^+}\sum_{m\in\mathbb{Z}}\alpha_{n-m}^\mathcal{I}\alpha_m^\mathcal{I}. \tag{3.49}$$

To summarise, in order to find a solution to the equations of motion it suffices to fix values for all $\alpha_n^\mathcal{I}$, a value for p^+, and a value for x_0^-. The $\alpha_n^\mathcal{I}$ fully determine $X^\mathcal{I}$, and together with p^+ they determine all α_n^-. Together with x_0^- this fixes X^- entirely, and since p^+ fixes X^+ by our parameterisation condition, we then know X.

Moreover, infinite sums in (3.49) are quite special, and will be very important later on, so we introduce special notation for them; they will be called the transverse Virasoro modes, and are denoted as

$$L_n = \frac{1}{2}\sum_{m\in\mathbb{Z}}\alpha_{n-m}^\mathcal{I}\alpha_m^\mathcal{I}. \tag{3.50}$$

It is no coincidence the notation we introduce for these modes coincides with the notation of the generators of the Virasoro algebra in chapter 2. In fact, in the quantum theory, the operators corresponding to these modes generate the Virasoro algebra.

3.2 Quantisation of the string

In the previous section we found out a classical string can be fully described by fixing the constants $\alpha_n^\mathcal{I}$, p^+, and x_0^-. In a quantum theory of the string, we have to turn these constants into operators, and establish their

commutation relations. In order to do this, we will first write down a new action for the transverse coordinates. This action is given by

$$S = \frac{1}{4\pi\alpha'} \int d\tau \int d\sigma \left((\dot{X}^{\mathcal{I}})^2 - (X^{\mathcal{I}\prime})^2 \right). \tag{3.51}$$

Note this action gives $\mathcal{P}^{\tau\mathcal{I}}$ and $\mathcal{P}^{\sigma\mathcal{I}}$ as the momenta conjugate to X, and the Euler-Lagrange equations are exactly the wave equation. Hence we can use this action to construct a quantum theory of the string. The advantage of this action over the Nambu-Goto action is that it does not contain a square root, which would be problematic, as we have to interpret the variables as operators. In order to solve this action, we once again expand $X^{\mathcal{I}}$ into its Fourier modes as

$$X^{\mathcal{I}} = q_0^{\mathcal{I}} + 2\sqrt{\alpha'} \sum_{n=1}^{\infty} q_n^{\mathcal{I}} \frac{\cos(n\sigma)}{\sqrt{n}}. \tag{3.52}$$

Substituting this back into our action gives

$$S = \int \left[\frac{1}{4\alpha'} \dot{q}_0^{\mathcal{I}} \dot{q}_0^{\mathcal{I}} + \sum_{n=1}^{\infty} \left(\frac{1}{2n} \dot{q}_n^{\mathcal{I}} \dot{q}_n^{\mathcal{I}} - \frac{n}{2} q_n^{\mathcal{I}} q_n^{\mathcal{I}} \right) \right] d\tau, \tag{3.53}$$

where we used that for all $n, m \in \mathbb{Z}_{\geq 1}$ we have the elementary formulas

$$\int_0^{\pi} \cos(nx)\cos(mx)\, dx = \int_0^{\pi} \sin(nx)\sin(mx)\, dx = \begin{cases} 0 & \text{if } n \neq m \\ \pi/2 & \text{if } n = m. \end{cases}$$

This action is now neatly the sum of the action of a free particle, which only has kinetic energy, and an infinite set of actions of harmonic oscillators. The free part simply has a linear solution, which we can write as $q_0^{\mathcal{I}} = x_0^{\mathcal{I}} + 2\alpha' p^{\mathcal{I}} \tau$. In order to solve for the oscillators, we first substitute $p_n^{\mathcal{I}}$ for $\dot{q}_n^{\mathcal{I}}$ to get

$$H = \alpha' p_0^{\mathcal{I}} p_0^{\mathcal{I}} + \sum_{n=1}^{\infty} \frac{n}{2} (p_n^{\mathcal{I}} p_n^{\mathcal{I}} + q_n^{\mathcal{I}} q_n^{\mathcal{I}}). \tag{3.54}$$

As illustrated in section 1.3.1, we expand these $q_n^{\mathcal{I}}$ as

$$q_n^{\mathcal{I}} = \frac{i}{\sqrt{2}} (a_n^{\mathcal{I}} e^{-in\tau} - a_n^{\mathcal{I}\dagger} e^{in\tau}), \tag{3.55}$$

which gives us

$$X^{\mathcal{I}} = x_0^{\mathcal{I}} + 2\alpha' p^{\mathcal{I}} \tau + i\sqrt{2\alpha'} \sum_{n=1}^{\infty} \left(a_n^{\mathcal{I}} e^{-in\tau} - a_n^{\mathcal{I}\dagger} e^{in\tau} \right) \frac{\cos(n\sigma)}{\sqrt{n}}. \tag{3.56}$$

Here we introduce operators $\alpha_n^{\mathcal{I}}$ as

$$\alpha_n^{\mathcal{I}} = \begin{cases} a_n^{\mathcal{I}}\sqrt{n} & \text{for } n > 0 \\ \sqrt{2\alpha'}p_0 & \text{for } n = 0 \\ (a_{-n}^{\mathcal{I}})^\dagger\sqrt{-n} & \text{for } n < 0 \end{cases} \tag{3.57}$$

completely analogous to (3.40) such that (3.56) becomes

$$X^{\mathcal{I}}(\tau,\sigma) = x_0^{\mathcal{I}} + \sqrt{2\alpha'}\alpha_0^{\mathcal{I}}\tau + i\sqrt{2\alpha'}\sum_{n\in\mathbb{Z}\setminus\{0\}}\frac{1}{n}\alpha_n^{\mathcal{I}}e^{-in\tau}\cos(n\sigma). \tag{3.58}$$

This concludes the work necessary for the transverse directions. What remains are the $+$ and $-$ directions. As before, the $+$ direction is very simple, since we can define the operators p^+ and α_0^+ to obey

$$X^+ = 2\alpha' p^+\tau = \sqrt{2\alpha'}\alpha_0^+\tau. \tag{3.59}$$

Since we can express X^+ solely in α_0^+, we are free to set $x_0^+ = 0$ and $\alpha_n^+ = 0$ for all $n \neq 0$. Moreover, X^- has a familiar expansion

$$X^- = x_0^- + \sqrt{2\alpha'}\alpha_0^-\tau + i\sqrt{2\alpha'}\sum_{n\neq 0}\frac{1}{n}\alpha_n^- e^{-in\tau}\cos(n\sigma), \tag{3.60}$$

where the α_n^- can be expressed in terms of the $\alpha_n^{\mathcal{I}}$, namely through the Virasoro operators

$$L_n := \frac{1}{2}\sum_{p\in\mathbb{Z}}\alpha_{n-p}^{\mathcal{I}}\alpha_p^{\mathcal{I}} \tag{3.61}$$

as

$$\alpha_n^- = \frac{1}{\sqrt{2\alpha' p^+}}L_n. \tag{3.62}$$

We now have all operators we need, but of course this is not enough to quantise the system; we need to know how these operators interact with each other. We thus have to construct commutation relations among all operators, which we will now do.

3.2.1 Commutation relations and normal ordering

As illustrated in the example of the single quantum harmonic oscillator, for each $n \in \mathbb{N}$ and $I \in \{2, ..., D\}$ we have $[a_n^I, (a_n^I)^\dagger] = 1$. Hence, by definition $[\alpha_n^I, \alpha_{-n}^I] = n$. Moreover, since different a_n^I and a_m^J act independently, two α_n^I and α_m^J only fail to commute when $I = J$ and $n = -m$. We thus conclude

$$[\alpha_n^I, \alpha_m^J] = n\delta_{I,J}\delta_{n,-m}. \tag{3.63}$$

With these commutation relations in mind, we can express the Hamiltonian in terms of L_0, namely via

$$\begin{aligned} H &= \alpha' p_0^\mathcal{I} p_0^\mathcal{I} + \sum_{n=1}^{\infty} \frac{n}{2}(p_n^\mathcal{I} p_n^\mathcal{I} + q_n^\mathcal{I} q_n^\mathcal{I}) = \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \sum_{n=1}^{\infty} n\left(a_n^\mathcal{I} a_n^{\mathcal{I}\dagger} - \frac{d-1}{2}\right) \\ &= \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \sum_{n=1}^{\infty}\left(\alpha_n^\mathcal{I}\alpha_{-n}^\mathcal{I} - \frac{n(d-1)}{2}\right) \\ &= \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \sum_{n=1}^{\infty}\left(\frac{1}{2}\alpha_n^\mathcal{I}\alpha_{-n}^\mathcal{I} - \frac{n(d-1)}{2}\right) + \sum_{n=1}^{\infty}\left(\frac{1}{2}\alpha_n^\mathcal{I} a_{-n}^\mathcal{I}\right) \\ &= \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \sum_{n=1}^{\infty}\left(\frac{1}{2}a_{-n}^\mathcal{I}\alpha_n^\mathcal{I}\right) + \sum_{n=1}^{\infty}\left(\frac{1}{2}\alpha_n^\mathcal{I} a_{-n}^\mathcal{I}\right) = L_0. \end{aligned}$$

However, this calculation contains one major flaw; the second and third line contain infinite sums over positive integers. These sums are ill defined to say the least. The problem we encounter here becomes even more apparent when we try to *normal order* L_n. To normal order an operator means to write it as a sum over products of creation and annihilation operators, such that all annihilation operators appear left of the creation operators. This is desirable, since such a normal ordered operator gives a clear result when acting on the vacuum state. As a consequence of (3.63), every L_n for $n \neq 0$ has no ordering ambiguity. However, there is a problem for L_0, since α_n^I and α_{-n}^I do not commute for $n \neq 0$. If we were to take (3.61) as definition for L_0, then in order to normal order this L_0 we would find

$$\begin{aligned} L_0 &\overset{?}{=} \frac{1}{2}\sum_{p\in\mathbb{Z}} \alpha_{-p}^\mathcal{I}\alpha_p^\mathcal{I} = \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \frac{1}{2}\sum_{p>0}\alpha_{-p}^\mathcal{I}\alpha_p^\mathcal{I} + \frac{1}{2}\sum_{p>0}\alpha_p^\mathcal{I}\alpha_{-p}^\mathcal{I} \\ &= \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \frac{1}{2}\sum_{p>0}\alpha_{-p}^\mathcal{I}\alpha_p^\mathcal{I} + \frac{1}{2}\sum_{p>0}\left(\alpha_{-p}^\mathcal{I}\alpha_p^\mathcal{I} + [\alpha_p^\mathcal{I}, \alpha_{-p}^\mathcal{I}]\right) \\ &= \frac{1}{2}\alpha_0^\mathcal{I}\alpha_0^\mathcal{I} + \sum_{p>0}\alpha_{-p}^\mathcal{I}\alpha_p^\mathcal{I} + \frac{1}{2}\sum_{p>0}(d-1)p. \end{aligned} \tag{3.64}$$

This is problematic, since this last expression contains a clearly divergent part. However, we can avoid these problems by defining L_0 to be normal ordered without this divergent part, as

$$L_0 := \frac{1}{2}\alpha_0^{\mathcal{I}}\alpha_0^{\mathcal{I}} + \sum_{p>0}\alpha_{-p}^{\mathcal{I}}\alpha_p^{\mathcal{I}}. \tag{3.65}$$

Consequently, the correct expression for α_0^- becomes

$$\alpha_0^- := \frac{\sqrt{2\alpha'}}{p^+}(L_0 + a) \tag{3.66}$$

for some yet to be fixed constant a, and the Hamiltonian equals

$$H = L_0 + a. \tag{3.67}$$

By direct calculation we can now find the commutator of some arbitrary α_n^I and L_m. In the case where $m \neq 0$ this can be done by substituting (3.61) and using the fact that α_n^I and α_m^J always commute if $I \neq J$. Hence, $[\alpha_n^I, L_m]$ can be simplified to

$$[\alpha_n^I, L_m] = \left[\alpha_n^I, \frac{1}{2}\sum_{p\in\mathbb{Z}}\alpha_{m-p}^{\mathcal{I}}\alpha_p^{\mathcal{I}}\right] = \frac{1}{2}\sum_{p\in\mathbb{Z}}[\alpha_n^I, \alpha_{m-p}^I\alpha_p^I]. \tag{3.68}$$

This can further be simplified by using the general identity $[X, YZ] = [X, Y]Z + Y[X, Z]$, to

$$\begin{aligned}
[\alpha_n^I, L_m] &= \frac{1}{2}\sum_{p\in\mathbb{Z}}[\alpha_n^I, \alpha_{m-p}^I\alpha_p^I] \\
&= \frac{1}{2}\sum_{p\in\mathbb{Z}}\left([\alpha_n^I, \alpha_{m-p}^I]\alpha_p^I + \alpha_{m-p}^I[\alpha_n^I, \alpha_p^I]\right) \\
&= \frac{1}{2}\sum_{p\in\mathbb{Z}}\left(n\delta_{n+m-p,0}\alpha_p^I + \alpha_{m-p}^I n\delta_{n+p,0}\right) \\
&= \frac{1}{2}\left(n\alpha_{n+m}^I + \alpha_{m+n}^I n\right) = n\alpha_{n+m}^I.
\end{aligned} \tag{3.69}$$

A similar calculation gives the same result when $m = 0$. We can now finally determine $[L_n, L_m]$. In order to do this properly, we first rewrite L_n as

$$L_n = \frac{1}{2}\sum_{p\geq 0}\alpha_{n-p}^{\mathcal{I}}\alpha_p^{\mathcal{I}} + \frac{1}{2}\sum_{p<0}\alpha_p^{\mathcal{I}}\alpha_{n-p}^{\mathcal{I}}, \tag{3.70}$$

which also holds for $n = 0$. We then find for $n, m \in \mathbb{Z}$ that

$$\begin{aligned}
[L_n, L_m] &= \frac{1}{2}\sum_{p\geq 0}[\alpha^{\mathcal{I}}_{n-p}\alpha^{\mathcal{I}}_p, L_m] + \frac{1}{2}\sum_{p<0}[\alpha^{\mathcal{I}}_p\alpha^{\mathcal{I}}_{n-p}, L_m] \\
&= \frac{1}{2}\sum_{p\geq 0}[\alpha^{\mathcal{I}}_{n-p}, L_m]\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p\geq 0}\alpha^{\mathcal{I}}_{n-p}[\alpha^{\mathcal{I}}_p, L_m] \\
&+ \frac{1}{2}\sum_{p<0}[\alpha^{\mathcal{I}}_p, L_m]\alpha^{\mathcal{I}}_{n-p} + \frac{1}{2}\sum_{p<0}\alpha^{\mathcal{I}}_p[\alpha^{\mathcal{I}}_{n-p}, L_m] \\
&= \frac{1}{2}\sum_{p\geq 0}(n-p)\alpha^{\mathcal{I}}_{n-p+m}\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p\geq 0}p\alpha^{\mathcal{I}}_{n-p}\alpha^{\mathcal{I}}_{p+m} \\
&+ \frac{1}{2}\sum_{p<0}p\alpha^{\mathcal{I}}_{p+m}\alpha^{\mathcal{I}}_{n-p} + \frac{1}{2}\sum_{p<0}(n-p)\alpha^{\mathcal{I}}_p\alpha^{\mathcal{I}}_{n-p+m}.
\end{aligned} \tag{3.71}$$

Recall two α_n and α_m only fail to commute if $n + m = 0$. Hence, assuming this is not the case (so $n + m \neq 0$), the thirst and fourth and the second and third terms can be directly combined, which gives us

$$\begin{aligned}
[L_n, L_m] &= \frac{1}{2}\sum_{p\in\mathbb{Z}}(n-p)\alpha^{\mathcal{I}}_{n-p+m}\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p\in\mathbb{Z}}p\alpha^{\mathcal{I}}_{n-p}\alpha^{\mathcal{I}}_{p+m} \\
&= \frac{1}{2}\sum_{p\in\mathbb{Z}}(n-p)\alpha^{\mathcal{I}}_{n+m-p}\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p\in\mathbb{Z}}(p-m)\alpha^{\mathcal{I}}_{n+m-p}\alpha^{\mathcal{I}}_p \\
&= \frac{1}{2}\sum_{p\in\mathbb{Z}}(n-m)\alpha^{\mathcal{I}}_{n+m-p}\alpha^{\mathcal{I}}_p \\
&= (n-m)L_{n+m}.
\end{aligned} \tag{3.72}$$

For the case where $m = -n$, we have to proceed slightly differently. Namely, we first rewrite (3.71) such that the operator occurring on the right is $\alpha^{\mathcal{I}}_p$, which becomes

$$\begin{aligned}
[L_n, L_{-n}] &= \frac{1}{2}\sum_{p\geq 0}(n-p)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p\geq 0}p\alpha^{\mathcal{I}}_{n-p}\alpha^{\mathcal{I}}_{p-n} \\
&+ \frac{1}{2}\sum_{p<0}p\alpha^{\mathcal{I}}_{p-n}\alpha^{\mathcal{I}}_{n-p} + \frac{1}{2}\sum_{p<0}(n-p)\alpha^{\mathcal{I}}_p\alpha^{\mathcal{I}}_{-p} \\
&= \frac{1}{2}\sum_{p\geq 0}(n-p)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p\geq -n}(p+n)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_p \\
&+ \frac{1}{2}\sum_{p>n}(n-p)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_p + \frac{1}{2}\sum_{p>0}(n+p)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_p.
\end{aligned} \tag{3.73}$$

If this expression is organised slightly differently, it becomes

$$\begin{aligned}[L_n, L_{-n}] = &\frac{1}{2}\sum_{p>n} 4n\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p} + \frac{1}{2}\sum_{p=1}^{n}(3n+p)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p} \\ &+ \frac{1}{2}2n\alpha^{\mathcal{I}}_{0}\alpha^{\mathcal{I}}_{0} + \frac{1}{2}\sum_{p=-n}^{-1}(p+n)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p}.\end{aligned} \tag{3.74}$$

By now using $[\alpha^I_n, \alpha^J_m] = n\delta_{I,J}\delta_{n,-m}$ we can rewrite the last term as

$$\frac{1}{2}\sum_{p=-n}^{-1}(p+n)\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p} = \frac{1}{2}\sum_{p=1}^{n}(n-p)\left(\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p} + [\alpha^{\mathcal{I}}_{p}, \alpha^{\mathcal{I}}_{-p}]\right) \tag{3.75}$$

$$= \frac{1}{2}\sum_{p=1}^{n}(n-p)\left(\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p} + p(d-1)\right). \tag{3.76}$$

Substituting this back into (3.74) we get

$$[L_n, L_{-n}] = 2n\left(\frac{1}{2}\alpha^{\mathcal{I}}_{0}\alpha^{\mathcal{I}}_{0} + \sum_{p>0}\alpha^{\mathcal{I}}_{-p}\alpha^{\mathcal{I}}_{p}\right) + \frac{1}{2}\sum_{p=1}^{n}(n-p)p(d-1). \tag{3.77}$$

By induction, one can proof the identities[†]

$$\sum_{p=1}^{n} p = \frac{1}{2}n(n+1) \quad \text{and} \quad \sum_{p=1}^{n} p^2 = \frac{1}{6}(n+3n^2+2n^3), \tag{3.78}$$

which we can use to simplify (3.77) to

$$[L_n, L_{-n}] = 2nL_0 + \frac{d-1}{12}(n^3 - n). \tag{3.79}$$

Looking back at (3.72), we can therefore conclude that for *all* $n, m \in \mathbb{Z}$ we have

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{d-1}{12}(m^3 - m)\delta_{n,-m}. \tag{3.80}$$

As promised, the L_n indeed satisfy the Virasoro commutation relations. This hints at conformal symmetry of the string. In section 3.4, we will explain why the string is indeed has this conformal symmetry. First however, we briefly discuss Lorentz invariance of the string. As we will see, this will put an interesting restriction on the possible values of d. Namely, the only possibility turns out to be $d = 25$! [‡]

[†]These two formulas are the simplest examples of the so called *Faulhaber's polynomials*. More on these can be found in [Knu93].

[‡]This is not a factorial.

Version of July 11, 2020– Created July 11, 2020 - 11:00

3.3 Lorentz symmetry and conserved quantities

Just as discussed for the single harmonic oscillator, a relativistic quantum string theory needs to contain Lorentz generators, which must satisfy the commutation relations from (1.47). To recall, these relations are given by

$$[M^{\mu\nu}, M^{\rho\sigma}] = i\eta^{\mu\rho}M^{\nu\sigma} - i\eta^{\nu\rho}M^{\mu\sigma} + i\eta^{\mu\sigma}M^{\rho\nu} - i\eta^{\nu\sigma}M^{\rho\mu}. \tag{3.81}$$

Moreover, want these generators to be Hermitian and normal ordered. The classically conserved charge is

$$M^{\mu\nu} = \frac{1}{2\pi\alpha'}\int_{\sigma=0}^{\pi}(X^\mu \dot{X}^\nu - X^\nu \dot{X}^\mu)\, d\sigma, \tag{3.82}$$

which upon mode expansion and evaluation of the integral gives

$$M^{\mu\nu} = x_0^\mu p_0^\nu - x_0^\nu p_0^\mu - i\sum_{n=1}^{\infty}\frac{1}{n}(\alpha_{-n}^\mu \alpha_n^\nu - \alpha_{-n}^\nu \alpha_n^\mu). \tag{3.83}$$

Although it is tempting to interpret this expression in terms of quantum operators, this will not work entirely. As long as $\mu, \nu \neq -$, everything goes well. To see this, note the last term in (3.83) is always Hermitian, since

$$\begin{aligned}\left(-i\sum_{n=1}^{\infty}\frac{1}{n}(\alpha_{-n}^\mu \alpha_n^\nu - \alpha_{-n}^\nu \alpha_n^\mu)\right)^\dagger &= +i\sum_{n=1}^{\infty}\frac{1}{n}\left((\alpha_n^\nu)^\dagger(\alpha_{-n}^\mu)^\dagger - (\alpha_n^\mu)^\dagger(\alpha_{-n}^\nu)^\dagger\right)\\ &= +i\sum_{n=1}^{\infty}\frac{1}{n}\left(\alpha_{-n}^\nu \alpha_n^\mu - \alpha_{-n}^\mu \alpha_n^\nu\right)\\ &= -i\sum_{n=1}^{\infty}\frac{1}{n}(\alpha_{-n}^\mu \alpha_n^\nu - \alpha_{-n}^\nu \alpha_n^\mu).\end{aligned}$$

Moreover, in the case where $\mu, \nu \neq -$, x_0^μ and p_0^ν commute[§]. As x_0^μ and p_0^ν are assumed to commute, this implies (3.83) is Hermitian. Moreover, the commutation relations of different $M^{\mu\nu}, M^{\rho\sigma}$ (where $\mu, \nu, \rho, \sigma \neq -$) are easily checked to satisfy (3.81). For example, assuming $\mu, \nu, \rho, \sigma \in \{2, \ldots, d\}$, the only relevant relations to be checked are those of the form $[M^{IJ}, M^{IK}]$, for $I, J, K \in \{2, \ldots, d\}$ distinct. Now since

$$\left[-i\frac{1}{n}\alpha_{-n}^I \alpha_n^J, i\frac{1}{n}\alpha_{-n}^K \alpha_n^I\right] = -\frac{1}{n}\alpha_n^J \alpha_{-n}^K \quad \text{and} \quad [x_0^I p_0^J, -x_0^K p_0^I] = -ix_0^K p_0^J,$$

[§]We assume $\mu \neq \nu$, as $\mu = \nu$ implies $M^{\mu\nu} = 0$ and there is nothing to show.

Version of July 11, 2020– Created July 11, 2020 - 11:00

it follows that

$$\begin{aligned}[M^{IJ}, M^{IK}] &= -ix_0^K p_0^J + ix_0^J p_0^K + \sum_{n=1}^{\infty} \left(-\frac{1}{n}\alpha_n^J \alpha_{-n}^K + \frac{1}{n}\alpha_n^K \alpha_{-n}^J \right) \\ &= -iM^{KJ} = iM^{JK},\end{aligned} \tag{3.84}$$

which indeed coincides with (3.81).

In the situation where one of μ, ν, ρ, σ equals $+$, life becomes even easier, as all α_n^+ are zero for nonzero n. However, the last set of operators, those of the form M^{-I} for $I \in \{2, \ldots, d\}$, are more involved. First of all, (3.83) is not Hermitian for $\mu = -$ and $\nu = I$, since x_0^I and p_0^- do not commute. The correct Hermitian expression can be found by averaging (3.83) with its Hermitian adjoint, which for $\mu = -$ and $\nu = I$ gives

$$M^{-I} = x_0^- p_0^I - \frac{1}{2}(x_0^I p_0^- + p_0^- x_0^I) - i \sum_{n=1}^{\infty} \frac{1}{n}(\alpha_{-n}^- \alpha_n^I - \alpha_{-n}^I \alpha_n^-). \tag{3.85}$$

It can be shown that this operator indeed generates the correct Lorentz transformations, and thereby has to satisfy the commutation relations from (3.81). In particular, we would have to have

$$[M^{-I}, M^{-J}] = 0. \tag{3.86}$$

Directly calculating this commutator is cumbersome, to say the least. This becomes apparent when one recalls that the α_n^- can be expressed in terms of the L_n (c.f. (3.62)), which are quadratic in the α_n^I. The expression in (3.85) is therefore *cubic* in the α_n^I, which makes confirming (3.86) that much harder. However, utilising all commutation relations stated in this chapter, it is possible to calculate that

$$[M^{-I}, M^{-J}] = -\frac{1}{\alpha' p^{+2}} \sum_{n=1}^{\infty} K_m (\alpha_{-n}^I \alpha_n^J - \alpha_{-n}^J \alpha_n^I), \tag{3.87}$$

where

$$K_m = m\left(1 - \frac{d-1}{24}\right) + \frac{1}{m}\left(\frac{d-1}{24} + a\right). \tag{3.88}$$

A step by step handguide of this calculation is given in [Wei11]. Here a is the ordering constant introduced in (3.66). Since the $\alpha_{-n}^I \alpha_n^J - \alpha_{-n}^J \alpha_n^I$ are independent, in order for $[M^{-I}, M^{-J}]$ to vanish K_m has to be zero for every $m \in \mathbb{Z}_{\geq 1}$. This implies

$$1 - \frac{d-1}{24} = 0 \qquad \text{and} \qquad \frac{d-1}{24} + a = 0, \tag{3.89}$$

from which we can conclude

$$d = 25 \qquad \text{and} \qquad a = -1. \tag{3.90}$$

We thus conclude this free bosonic string theory can only exist in 26 spacetime dimensions! Although it might seem reasonable to therefore discard the theory all together and deem it unrealistic, such measures are not necessary. The theory can still be saved by *compactifying* certain dimensions. This means that instead of strings living in $\mathbb{R} \times \mathbb{R}^{25}$, we consider strings living in $\mathbb{R} \times \mathbb{R}^3 \times (S^1)^{22}$. If the metric on the latter space is chosen such that all S^1 are small enough, this space will resemble $\mathbb{R} \times \mathbb{R}^3$ to the observers living in this universe. Moreover, more realistic superstring theories has been developed, which also allow fermionic particles. A similar argument as in the bosonic case shows such theories are only consistent in either ten or eleven spacetime dimensions. Hence, it will be up to experimentalists to either find evidence or counterevidence for these extra dimensions, in order to either strengthen the case of string theory, or reject it.

3.4 Conformal properties of the world sheet

Apart from Lorentz invariance, bosonic string theory also has the promised conformal symmetry, which was hinted at by the occurrence of the Virasoro algebra in the quantum theory. The best way to exhibit the conformal invariance of the string is by introducing a new action for the string, namely the *Polyakov action*. This action depends on a metric on the world sheet, so we will let $h(\sigma, \tau)$ be a metric on the world sheet of signature $(-,+)$. The Polyakov action is given by

$$S_h = -\frac{1}{4\pi\alpha'} \int d\tau \int d\sigma \sqrt{-\det(h^{-1})} h^{\alpha\beta} \partial_\alpha X^\mu \partial_\beta X^\nu \eta_{\mu\nu} \tag{3.91}$$

written in Einstein notation (cf. (1.5)). Here we view h as a 2×2 matrix with components $h^{\alpha\beta}$, and η is the 2×2 matrix $\eta = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ mentioned earlier. The reason the Polyakov action is interesting, is for two reasons. First of all, it is equivalent to the Nambu-Goto action with which we started this chapter. This can be shown by calculating the equations of motion corresponding to this action, which is done in [Zwi04]. The second reason is because it is invariant under conformal transformations in a very clear way. Namely, if we have a conformal transformation that takes h to $\tilde{h} = \lambda h$

for some $\lambda \in \mathbb{R}_{>0}$, then

$$\begin{aligned}\sqrt{-\det(\tilde{h}^{-1})}\tilde{h}^{\alpha\beta} &= \sqrt{-\det((\lambda h)^{-1})}\lambda h^{\alpha\beta} \\ &= \sqrt{-\lambda^{-2}\det(h^{-1})}\lambda h^{\alpha\beta} \\ &= \sqrt{-\det(h^{-1})}h^{\alpha\beta}.\end{aligned} \tag{3.92}$$

This directly implies $S_h = S_{\tilde{h}}$, and hence the Polyakov action is invariant under conformal transformations. This explains the existence of a copy of the Virasoro algebra in the quantum theory.

We have hereby concluded our discussion of the bosonic string. One way to learn more about string theories is by looking at other, possibly simpler, theories that share properties with string theory. One of the most important properties are of course its set of symmetries. We have discussed both Lorentz invariance and conformal symmetry of the string. Whereas Lorentz invariance occurs in any relativistic theory, (two-dimensional) conformal symmetry is specifically found in string theories, as the world lines of particles are here replaced by two-dimensional world sheets. This inspires us to look more abstractly into theories that have two-dimensional conformal invariance. Since such theories are always representations of the Virasoro algebra, we will specifically look at the simplest among these representations.

Version of July 11, 2020– Created July 11, 2020 - 11:00

Chapter 4

Representations of the Virasoro algebra

In this chapter we study representations of the Virasoro algebra. However, not every such representation is physically relevant, so we will introduce certain restrictions on these representations. Just as in string theory, we will always think of L_0 as the Hamiltonian. Another way to see why this is a sensible choice, is by considering the *closed* string instead of the open one. A closed string can, just as the open string, be parameterised by two coordinates τ, σ, representing the time and space directions respectively. The difference with the open string is that we take $\sigma \in \mathbb{R}$ instead of $\sigma \in [0, \pi]$, and the parameterisation X has to satisfy

$$X(\tau, \sigma + 2\pi) = X(\tau, \sigma) \tag{4.1}$$

for all $\tau, \sigma \in \mathbb{R}$. We can also view such an X as a function of the complex coordinate $z = \tau + i\sigma$. The constraint (4.1) then implies X satisfies $X(z + 2\pi i) = X(z)$, and hence that it only depends on $\exp(z)$. Formally this means X factorises as a map $\tilde{X} : \mathbb{C}^* \to \mathbb{R}^{(1,d)}$, such that $X(z) = \tilde{X} \circ \exp$. On $\mathbb{C}^*$, the radial direction now corresponds to the time direction. We can now, just as in section 2.3 where we originally defined the Witt algebra, let act l_0 act on X. We then get

$$\begin{aligned} \exp(al_0)X(z) = \exp(-az\partial_z)X(z) &= \sum_{k=0}^{\infty} \frac{(-az\partial_z)^k}{k!} X(z) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} X^{(k)}(z)(-az)^k = X((1-a)z). \end{aligned}$$

We thus find l_0 to generate dilatations, which indeed correspond to time translations. Hence, it makes sense to view l_0, and consequently L_0, as the Hamiltonian. This will be the main inspiration behind the definition of so called lowest weight representation, which we will make shortly. First however, we will show some more general results. We start with a short lemma. Recall we denote the Virasoro algebra with central charge $c \in \mathbb{C}$ as $\mathfrak{V}_c$.

Lemma 4.1. *Let V be a $\mathfrak{V}_c$-representation, and let $v \in V$ with $L_0 v = \lambda v$ for some $\lambda \in \mathbb{C}$. Then for all $v_k := L_{n_k} L_{n_{k-1}} \dots L_{n_1} v$ with $n_1, \dots, n_m \in \mathbb{Z}$ we have $L_0 v_k = (\lambda - \sum_{i=1}^{k} n_i) v_k$.*

Proof. We prove this by induction on k. For $k = 0$, there is nothing to prove. So now assume the statement holds for $v_{k-1} := L_{n_{k-1}} \dots L_{n_1} v$, i.e. $L_0 v_{k-1} = (\lambda - \sum_{i=1}^{k-1} n_i) v_{k-1}$ (IH). We then have

$$\begin{aligned} L_0 v_k &= L_0 (L_{n_k} L_{n_{k-1}} \dots L_{n_1} v) = (L_0 L_{n_k}) L_{n_{k-1}} \dots L_{n_1} v \\ &= (L_{n_k} L_0 + [L_0, L_{n_k}]) L_{n_{k-1}} \dots L_{n_1} v \\ &\overset{IH}{=} L_{n_k} \left(\lambda - \sum_{i=1}^{k-1} n_i \right) L_{n_{k-1}} \dots L_{n_1} v - n_k L_{n_k} L_{n_{k-1}} \dots L_{n_1} v \\ &= \left(\lambda - \sum_{i=1}^{k} n_i \right) v_k. \end{aligned}$$

□

In particular, this either means v_k' is an eigenvector of L_0, or it is zero. In the first case, this means that applying any L_i to an eigenvector of L_0 just lowers the L_0-eigenvalue of this eigenvector.

The second result we need to mention is simply an application of the Poincaré-Birkhoff-Witt theorem (or PBW theorem for short, c.f. theorem 2.12) on the Virasoro algebra. In order to do this, we need to give an ordered basis of $\mathfrak{V}_c$. By definition, the set $\{L_n\}_{n \in \mathbb{Z}} \cup \{I\}$ is a basis, so we only need to come up with a total order on this set. A natural choice is the following: for all $n, m \in \mathbb{Z}$ with $n \leq m$ we define $L_n \preceq L_m$ and $L_n \preceq I$. The PBW theorem than gives the following result.

Corollary 4.2. *The set*

$$\mathcal{B} := \left\{ L_{n_1} L_{n_2} \dots L_{n_m} I^k : \begin{array}{l} k \in \mathbb{N}, n_1, \dots, n_m \in \mathbb{Z}, \\ n_1 \leq n_2 \leq \dots \leq n_m \end{array} \right\} \tag{4.2}$$

forms a basis of $\mathcal{U}(\mathfrak{V}_c)$.

Moreover, we make some last definitions that will come in handy in the remainder of this chapter.

Definition 4.3. *We define*

$$\mathcal{B}^- := \left\{ L_{n_1} L_{n_2} \dots L_{n_m} : \begin{array}{l} m \in \mathbb{N}, n_1, \dots, n_m \in \mathbb{Z}, \\ n_1 \leq n_2 \leq \dots \leq n_m \leq 0 \end{array} \right\} \subseteq \mathcal{B}, \tag{4.3}$$

and we define the ***creation operator algebra*** *as*[*]

$$\mathcal{U}(\mathfrak{V}_c)^{(-)} := \langle \mathcal{B}^- \rangle \subseteq \mathcal{U}(\mathfrak{V}_c). \tag{4.4}$$

Furthermore, for all $N \in \mathbb{N}$ we define

$$\mathcal{B}^{(N)} := \left\{ L_{n_1} L_{n_2} \dots L_{n_m} \in \mathcal{B}^- : \sum_{i=1}^{m} n_i = -N \right\} \tag{4.5}$$

and

$$\mathcal{U}(\mathfrak{V}_c)^{(N)} := \left\langle \mathcal{B}^{(N)} \right\rangle \subseteq \mathcal{U}(\mathfrak{V}_c). \tag{4.6}$$

Elements of both $\mathcal{U}(\mathfrak{V}_c)^{(N)}$ are said to have ***level*** *N.*

Note that by the PBW theorem (theorem 2.12), it follows directly that

$$\mathcal{U}(\mathfrak{V}_c)^{(-)} = \bigoplus_{N \in \mathbb{N}} \mathcal{U}(\mathfrak{V}_c)^{(N)}. \tag{4.7}$$

4.1 Verma modules

When L_0 is seen as the Hamiltonian, we immediately get an extra condition on the L_0-eigenvalues. In any realistic physical system, the energy is bounded from below. Hence, the eigenvalues of L_0, which represent the energies, should also be bounded from below. This motivates us to define the so called *lowest weight representations*.

Definition 4.4. *A* ***lowest weight representation*** *(LWR) is a representation R of the Virasoro algebra, such that the following hold:*

1. *L_0 has an eigenvector $v_\Delta \in R$ with an eigenvalue $\Delta \in \mathbb{C}$ that has the smallest real part*[†]*. That is, $\mathrm{Re}(\Delta)$ is the minimum of $\{\mathrm{Re}(\lambda) : \lambda$ eigenvalue of $L_0\}$.*

[*]Here $\langle \mathcal{B}^- \rangle$ can both be seen as the vector space or the algebra generated by $\mathcal{B}^-$, as these are the same as sets.

[†]In particular, $v_\Delta \neq 0$.

2. $Iv_\Delta = v_\Delta$.

*We call Δ the **conformal dimension**.*

From these simple extra requirements we can already conclude something more about the action of $\mathfrak{V}_c$ on R. We will summarise this in the following lemma.

Lemma 4.5. *Let (R, ρ) be a lowest weight representation, with $\Delta \in \mathbb{C}$ and v_Δ as in definition 4.4. Then the following holds.*

1. *v_Δ is a **primary vector**. That is, it is nonzero and for all $n > 0$, we have $L_n v_\Delta = 0$.*
2. *If R is irreducible, for all $v \in R$ we have $Iv = v$. In other words, I acts as the identity.*

Proof. 1. From lemma 4.1 it follows directly that $L_0 L_n v_\Delta = (\Delta - n) L_n v_\Delta$. However, since $n > 0$, the real part of $\Delta - n$ is strictly smaller than the real part of Δ. Therefore, $L_n v_\Delta$ cannot be an eigenvector of L_0, so it has to equal 0.

2. Note that since I is central in $\mathfrak{V}_c$ (as defined in definition 2.15), for all $X \in \mathfrak{V}_c$ and $v \in R$ we have $\rho(I)(\rho(X)v) = \rho(X)(\rho(I)v)$. Hence, $\rho(I)$ is a homomorphism of representations (as defined in lemma 2.13). Since by the second assumption in definition 4.4 $\rho(I)$ has at least one eigenvalue (namely the eigenvalue 1, since $\rho(I)v_\Delta = v_\Delta$), by the second part of Schur's lemma (lemma 2.13) it acts as the identity on whole of R. □

From the first statement of this lemma we can directly learn more about the action of $\mathcal{U}(\mathfrak{V}_c)$ on an irreducible LWR R. Recall that by theorem 2.11 we get an action of $\mathcal{U}(\mathfrak{V}_c)$ on R for free, and hence we can look at the subspace $\mathcal{U}(\mathfrak{V}_c)v_\Delta := \{Uv_\Delta : U \in \mathcal{U}(\mathfrak{V}_c)\}$. This space is clearly closed under the action of $\mathcal{U}(\mathfrak{V}_c)$, and hence it is closed under the action of $\mathfrak{V}_c$ itself. It is therefore a subrepresentation. Since R was assumed to be irreducible, we conclude either $\mathcal{U}(\mathfrak{V}_c)v_\Delta = 0$ or $\mathcal{U}(\mathfrak{V}_c)v_\Delta = R$. By the second assumption of definition 4.4 the first option is ruled out, so we conclude $\mathcal{U}(\mathfrak{V}_c)v_\Delta = R$. By corollary 4.2, $\mathcal{U}(\mathfrak{V}_c)$ is generated by the set $\mathcal{B}$, from which we can conclude that R is generated (as a vector space) by

$$\begin{aligned} &\left\{ L_{n_1} L_{n_2} \dots L_{n_m} I^k v_\Delta : \begin{array}{l} m, k \in \mathbb{N}, n_1, \dots, n_m \in \mathbb{Z}, \\ n_1 \leq n_2 \leq \dots \leq n_m \end{array} \right\} \\ = &\left\{ L_{n_1} L_{n_2} \dots L_{n_m} v_\Delta : \begin{array}{l} m \in \mathbb{N}, n_1, \dots, n_m \in \mathbb{Z} \\ n_1 \leq n_2 \leq \dots \leq n_m \end{array} \right\}, \end{aligned} \tag{4.8}$$

where we used the second assumption of definition 4.4 to ignore the I^k-terms. Taking a closer look at an arbitrary element $L_{n_1}L_{n_2}\dots L_{n_m}v_\Delta$ of this set, the first part of lemma 4.5 tells us that if $n_m > 0$, then $L_{n_1}L_{n_2}\dots L_{n_m}v_\Delta = 0$. Now definition 4.3 seems very natural to make, as we can straightaway conclude that

$$R = \mathcal{U}(\mathfrak{V}_c)v_\Delta = \mathcal{U}(\mathfrak{V}_c)^{(-)}v_\Delta. \tag{4.9}$$

In other words, the linear map

$$\phi_R : \mathcal{U}(\mathfrak{V}_c)^{(-)} \to R \quad \text{defined by} \quad U \mapsto Uv_\Delta \tag{4.10}$$

is surjective. The case where ϕ_R is also injective warrants a separate definition.

Definition 4.6. *Let R be a lowest weight representation, with primary vector v_Δ and conformal dimension Δ (as defined in definition 4.4). If ϕ_R defined in (4.10) is both injective and surjective, then R is called a **Verma module**, which is denoted as $\mathcal{V}_\Delta$.*

By definition, $\mathcal{V}_\Delta = \mathcal{U}(\mathfrak{V}_c)^{(-)}v_\Delta$ for some primary vector v_Δ. From (4.7) it follows directly that

$$\mathcal{V}_\Delta = \mathcal{U}(\mathfrak{V}_c)^{(-)}v_\Delta = \bigoplus_{N\in\mathbb{N}} \left(\mathcal{U}(\mathfrak{V}_c)^{(N)}v_\Delta\right). \tag{4.11}$$

We have now written $\mathcal{V}_\Delta$ nicely as a direct sum of smaller subspaces (note that these subspaces are *not* subrepresentations, as they are not closed under the action of $\mathfrak{V}_c$), and these subspaces actually have have a significant meaning. That is, lemma 4.1 directly implies that $\mathcal{U}(\mathfrak{V}_c)^{(N)}v_\Delta$ is exactly the eigenspace of L_0 corresponding to the eigenvalue $\Delta + N$. The elements of $\mathcal{V}_\Delta$ can thus be arranged as in fig. 4.1.

It is important to note that the notation $\mathcal{V}_\Delta$ is not ambiguous, in the sense for that in general two Verma modules with equal conformal dimension are isomorphic. This will follow from the following proposition.

Proposition 4.7. *Let $\Delta \in \mathbb{C}$, let V be a Verma module and R a lowest weight representation, both with conformal dimension Δ. Then the map $\phi_R \circ {\phi_V}^{-1}$ is a homomorphism of representations, as defined in lemma 2.13.*

Proof. Let ${v_\Delta}^V$ and ${v_\Delta}^R$ be primary states in V respectively R, and let $L \in \mathfrak{V}_c$. Then any $v \in V$ can be uniquely written as $U{v_\Delta}^V$ for some $U \in$

‡This figure is inspired by figure 2.16 in [Rib14].

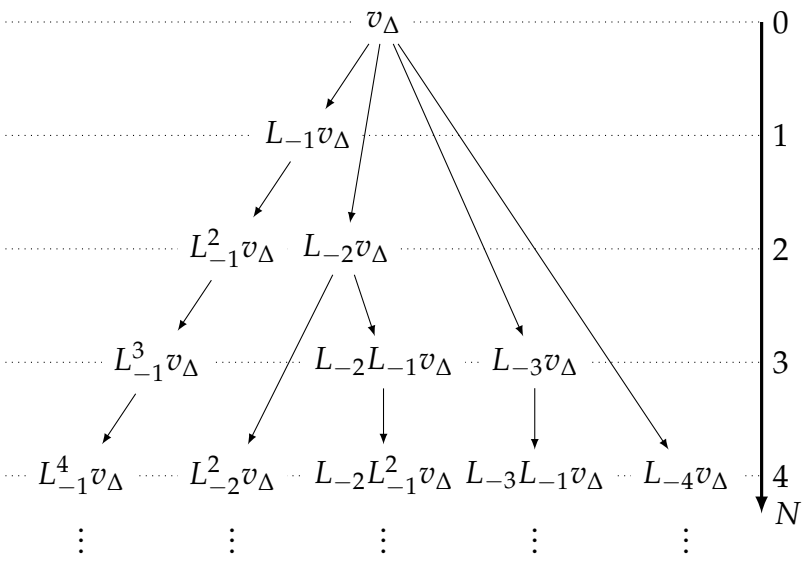

Figure 4.1: *The elements of a Verma module with primary vector v_Δ, ordered by their level N.* ‡

$\mathcal{U}(\mathfrak{V}_c)^{(-)}$. Hence we have

$$\begin{aligned} L((\phi_R \circ {\phi_V}^{-1})(v)) &= L((\phi_R \circ {\phi_V}^{-1})(U{v_\Delta}^V)) = L(\phi_R(U)) \\ &= L(Uv_{\Delta R}) = (LU)v_{\Delta R} \\ &= (\phi_R \circ {\phi_V}^{-1})((LU){v_\Delta}^V) \\ &= (\phi_R \circ {\phi_V}^{-1})(Lv). \end{aligned} \tag{4.12}$$

We thus conclude $\phi_R \circ {\phi_V}^{-1}$ is indeed a homomorphism of representations. □

Now assume we have two Verma modules $\mathcal{V}_\Delta$ and $\mathcal{V}'_\Delta$, that both have conformal dimension $\Delta \in \mathbb{C}$. Proposition 4.7 then states both $\phi_{\mathcal{V}_\Delta}$ and $\phi_{\mathcal{V}'_\Delta}$ are bijections, and hence $\phi_{\mathcal{V}_\Delta} \circ {\phi_{\mathcal{V}'_\Delta}}^{-1}$ is as well. This directly implies $\mathcal{V}_\Delta$ and $\mathcal{V}'_\Delta$ are isomorphic.

Furthermore, Verma modules are in general not irreducible, as they may contain nontrivial subrepresentations. Since the L_0-eigenvalues are discrete, such subrepresentations are also lowest weight representations, which are generated by so called *singular vectors*. Finding these singular hence gives a way to construct subrepresentations. We will hence spend the remainder of this chapter searching for them.

4.2 Finding singular vectors

In this section, $\mathcal{V}_\Delta$ is a Verma module with primary state v_Δ with conformal dimension Δ. We will first start with introducing some terminology .

Definition 4.8. *Let R be a representation of the Virasoro algebra, and let v_Δ be a primary state (cf. definition 4.4). Then for each $U \in \mathcal{U}(\mathfrak{V}_c)^{(-)}$, the state Uv_Δ, if it is linearly independent from v_Δ, is called a **descendent state** or simply **descendent**. If it is also a primary state, then it is called **singular vector**.*

Assume $\mathcal{V}_\Delta$ has a singular vector s. From (4.11) we can write $s = \sum_{i=0}^{M} s_i$ for some $M \in \mathbb{N}$ and with $s_i \in \mathcal{U}(\mathfrak{V}_c)^{(i)}$. Then for all $n > 0$ and each s_j we have

$$0 = L_n s = L_n \left(\sum_{i=0}^{M} s_i \right) = \sum_{i=0}^{M} L_n s_i, \tag{4.13}$$

and hence

$$L_n s_j = - \sum_{\substack{i=0 \\ i \neq j}}^{M} L_n s_i. \tag{4.14}$$

However, by lemma 4.1, if $L_n s_j$ is nonzero it is an eigenvector of L_0 with eigenvalue $\Delta + j - n$, and $-\sum_{\substack{i=0 \\ i \neq j}}^{M} L_n s_i$ clearly is not! Hence $L_n s_j = 0$ for each n, and therefore each s_j is a singular vector. We can conclude that if there is a singular vector v, we can assume it is an eigenvector of L_0. We will discuss two examples of these singular vectors, namely those at the low levels (recall the definition of level in definition 4.3).

Example 4.9. *Assume we have a singular vector $s \in \mathcal{V}_\Delta$ at level 1 (in other words, $L_0 s = \Delta + 1$). Then $s \in \mathcal{U}(\mathfrak{V}_c)^{(1)} v_\Delta = \langle L_{-1} \rangle v_\Delta$, so $s = a L_{-1} v_\Delta$ for some $a \in \mathbb{C}^\times$. Now by assumption for all $n > 0$ we have $L_n s = a L_n L_{-1} v_\Delta = 0$. By definition of a lowest weight representations this is automatically true for all $n > 1$. We remain to demand $a L_1 L_{-1} v_\Delta = 0$, and hence $L_1 L_{-1} v_\Delta = 0$. To actually get a concrete condition from this expression, we can simplify this expression using the commutator rules we used to define the Virasoro algebra in definition 2.16, namely*

$$\begin{aligned} L_1 L_{-1} v_\Delta &= (L_{-1} L_1 + [L_1, L_{-1}]) v_\Delta = L_{-1} L_1 v_\Delta + a[L_1, L_{-1}] v_\Delta \\ &= 0 + 2 L_0 v_\Delta = 2\Delta v_\Delta. \end{aligned} \tag{4.15}$$

Thus there is a singular vector at level 1 if and only if $\Delta = 0$.

Version of July 11, 2020– Created July 11, 2020 - 11:00

Example 4.10. *Assume we have a singular vector s at level 2. Then there are $a, b \in \mathbb{C}$, not both zero, such that $s = aL_{-2}v_\Delta + bL_{-1}^2 v_\Delta$. Similarly to the previous example, we do not have to worry about $L_n s$ for $n > 2$, since these are zero by the previous assumptions we have made. We thus only need to require $L_1 s = L_2 s = 0$. This can be directly written out as*

$$\begin{aligned} L_1 s &= aL_1 L_{-2} v_\Delta + bL_1 L_{-1}^2 v_\Delta \\ &= a(L_{-2}L_1 + [L_1, L_{-2}])v_\Delta + b(L_{-1}L_1 + [L_1, L_{-1}])L_{-1}v_\Delta \\ &= aL_{-2}L_1 v_\Delta + 3aL_{-1}v_\Delta + bL_{-1}L_1 L_{-1}v_\Delta + 2bL_0 L_{-1}v_\Delta \\ &= 0 + 3aL_{-1}v_\Delta + bL_{-1}^2 L_1 v_\Delta + 2bL_{-1}L_0 v_\Delta + 2bL_0 L_{-1}v_\Delta \\ &= 3aL_{-1}v_\Delta + 2\Delta bL_{-1}v_\Delta + 2(\Delta+1)bL_{-1}v_\Delta \\ &= (3a + (4\Delta+2)b)L_{-1}v_\Delta, \end{aligned}$$

and trough a similar simple (but tedious) computation we can also find

$$L_2 s = ((4\Delta + \tfrac{1}{2}c)a + 6\Delta b)v_\Delta.$$

We thus need to find for which c, Δ the two expressions above can be simultaneously zero for a proper choice of a and b. This is equivalent to finding c, Δ such that the matrix

$$K_2 := \begin{pmatrix} 3 & 4\Delta + 2 \\ 4\Delta + \frac{1}{2}c & 6\Delta \end{pmatrix} \tag{4.16}$$

has a non-trivial kernel element (a, b), which on its turn is equivalent to $\det(K_2) = 0$. Since

$$\det(K_2) = 18\Delta - (4\Delta + \tfrac{1}{2}c)(4\Delta + 2) = -16\Delta^2 + (10 - 2c)\Delta - c,$$

which is the case if and only if

$$\Delta = \frac{10 - 2c \pm \sqrt{(2c-10)^2 - 4 \cdot 16 \cdot c}}{2 \cdot 16} = \frac{5 - c \pm \sqrt{(c-25)(c-1)}}{16}.$$

Although the method for finding singular vectors used in these two examples certainly works, it is not very efficient, and it only gets more complicated for higher levels. Hence we will look at another way of finding these values, which will ultimately become theorem 4.14. In order to get there, we first introduce a complex inner product on $\mathcal{V}_\Delta$. From a physical perspective, this inner product is very natural. If we want to view $\mathcal{V}_\Delta$ as the state space of some quantum theory, it should be endowed with an inner product. Moreover, just as for the open string, we want L_{-n} to be the Hermitian adjoint of L_n for each $n \in \mathbb{Z}$.

Definition 4.11. *We define* $\langle \cdot, \cdot \rangle : \mathcal{V}_\Delta \times \mathcal{V}_\Delta \to \mathbb{C}$ *as the sesquilinear*[§] *map satisfying*

$$\langle v_\Delta, v_\Delta \rangle = 1, \tag{4.17}$$

and for all $n \in \mathbb{Z}$ *and* $v, w \in \mathcal{V}_\Delta$

$$\langle L_n v, w \rangle = \langle v, L_{-n} w \rangle. \tag{4.18}$$

Although (4.17) and (4.18) certainly seem like a loose requirement, they do in fact fully determine $\langle \cdot, \cdot \rangle$. To see this, first note that by assumption $\mathcal{U}(\mathfrak{V}_c)^{(-)} v_\Delta$ forms a basis of $\mathcal{V}_\Delta$, so by the assumed sesquilinearity it suffices to look at expressions of the form

$$\langle L_{n_1} \dots L_{n_k} v_\Delta, L_{m_1} \dots L_{m_\ell} v_\Delta \rangle,$$

which by (4.18) equals

$$\langle v_\Delta, L_{-n_k} \dots L_{-n_1} L_{m_1} \dots L_{m_\ell} v_\Delta \rangle.$$

By using the commutation relations of the L_i (from (2.26)) we can write the vector $L_{n_k} \dots L_{-n_1} L_{m_1} \dots L_{m_\ell} v_\Delta$ in terms of the basis elements of $\mathcal{U}(\mathfrak{V}_c)^{(-)} v_\Delta$, and by again using the assumed sesquilinearity we can reduce this further to determining expressions of the form $\langle v_\Delta, L v_\Delta \rangle$, where $L \in \mathcal{B}^-$. Hence $L \in \mathcal{B}^{(N)}$ for some $N \in \mathbb{N}_{>0}$, which implies that $L v_\Delta$ is an eigenvector of L_0 with eigenvalue $\Delta + N$. However, this in fact directly implies that v_Δ and $L v_\Delta$ are orthogonal, meaning that $\langle v_\Delta, L v_\Delta \rangle = 0$. To see this, note

$$\begin{aligned} \overline{\Delta} \langle v_\Delta, L v_\Delta \rangle &= \langle \Delta v_\Delta, L v_\Delta \rangle = \langle L_0 v_\Delta, L v_\Delta \rangle = \langle v_\Delta, L_0 L v_\Delta \rangle \\ &= \langle v_\Delta, (\Delta + N) L v_\Delta \rangle = (\Delta + N) \langle v_\Delta, L v_\Delta \rangle \end{aligned} \tag{4.19}$$

and therefore

$$0 = (N + \Delta - \overline{\Delta}) \langle v_\Delta, L v_\Delta \rangle = (N + 2 \operatorname{Im}(\Delta) i) \langle v_\Delta, L v_\Delta \rangle. \tag{4.20}$$

For all $N \neq 0$ however, $N + 2 \operatorname{Im}(\Delta) i$ is nonzero, so $\langle v_\Delta, L v_\Delta \rangle$ has to be zero. In a very similar fashion it can easily be shown that any two vectors of a different level are orthogonal, which can be used to simplify calculations considerably, as will be illustrated in the following example.

[§]By sesquilinearity we understand linearity in the *second* argument, and antilinearity in the first argument.

Example 4.12. *Suppose we wanted to calculate* $x := \langle (2L_0 + L_{-2})v_\Delta, (iL_{-1} + 2L_{-2} + (3+i)L_{-2}L_{-1})v_\Delta \rangle$. *Using the sesquilinearity we can write this as*

$$\begin{aligned} x &= \langle (2L_0 + L_{-2})v_\Delta, (iL_{-1} + L_{-2} + (3+i)L_{-2}L_{-1})v_\Delta \rangle \\ &= 2i\langle L_0 v_\Delta, L_{-1}v_\Delta \rangle + 2\langle L_0 v_\Delta, L_{-2}v_\Delta \rangle + 2(3+i)\langle L_0 v_\Delta, L_{-2}L_{-1}v_\Delta \rangle \\ &\quad + i\langle L_{-2}v_\Delta, L_{-1}v_\Delta \rangle + 2\langle L_{-2}v_\Delta, L_{-2}v_\Delta \rangle + (3+i)\langle L_{-2}v_\Delta, L_{-2}L_{-1}v_\Delta \rangle. \end{aligned}$$

By using the orthogonality property of the different basis vectors, only one of these six terms survives, and the above expression reduces to

$$x = 2\langle L_{-2}v_\Delta, L_{-2}v_\Delta \rangle,$$

which we can directly calculate, namely via

$$\begin{aligned} x &= 2\langle L_{-2}v_\Delta, L_{-2}v_\Delta \rangle = 2\langle v_\Delta, L_2 L_{-2}v_\Delta \rangle = 2\langle v_\Delta, (L_{-2}L_2 + [L_2, L_{-2}])v_\Delta \rangle \\ &= 2\langle v_\Delta, [L_2, L_{-2}]v_\Delta \rangle = 2\langle v_\Delta, (4L_0 - \tfrac{c}{12} \cdot 6 \cdot I)v_\Delta \rangle = 2\langle v_\Delta, (4\Delta - \tfrac{c}{2})v_\Delta \rangle \\ &= (8\Delta - c)\langle v_\Delta, v_\Delta \rangle = 8\Delta - c. \end{aligned}$$

Definition 4.13. *Let* $N \in \mathbb{N}$, *let* $M = \#\mathcal{B}^{(N)}$ [¶], *and let* $b_1, \dots, b_M \in \mathcal{B}^{(N)}$ *be distinct elements (so* $\mathcal{B}^{(N)} = \{b_i : 1 \leq i \leq M\}$*). We define* $G^{(N)}$ *to be the* ***Gramian matrix*** *of* $\langle \cdot, \cdot \rangle$ *on the subspace* $\mathcal{U}(\mathfrak{V}_c)^{(N)}v_\Delta$ *with respect to the basis* $b_1, \dots, b_M$. *In other words,* $G^{(N)}$ *is the* $M \times M$ *matrix with coefficients* $G^{(N)}_{ij} = \langle b_i, b_j \rangle$, *which looks like*[‖]

$$G^{(N)} = \begin{pmatrix} \langle b_1, b_1 \rangle & \langle b_1, b_2 \rangle & \dots & \langle b_1, b_M \rangle \\ \langle b_2, b_1 \rangle & \langle b_2, b_2 \rangle & \dots & \langle b_2, b_M \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle b_M, b_1 \rangle & \langle b_M, b_2 \rangle & \dots & \langle b_M, b_M \rangle \end{pmatrix}. \tag{4.21}$$

For example, $G^{(2)}$ is given by

$$G^{(2)} = \begin{pmatrix} \langle L_{-2}v_\Delta, L_{-2}v_\Delta \rangle & \langle L_{-2}v_\Delta, L_{-1}^2 v_\Delta \rangle \\ \langle L_{-1}^2 v_\Delta, L_{-2}v_\Delta \rangle & \langle L_{-1}^2 v_\Delta, L_{-1}^2 v_\Delta \rangle \end{pmatrix} = \begin{pmatrix} 4\Delta - \frac{c}{2} & 6\Delta \\ 6\Delta & 8\Delta + 2 \end{pmatrix}. \tag{4.22}$$

The matrices $G^{(N)}$ can now be used to find subrepresentations through the following theorem ([Rib14]).

[¶] In fact $M = P(N)$, with $P(N)$ being the *partition function* from number theory, which gives the amount of ways N can be written as the sum of non-negative integers.

[‖] Implicitly we have to choose an ordering on $\mathcal{B}^{(N)}$ here. For our purpose it does not matter which one is chosen, but to make this definition consistent it can be assumed to be lexicographically ordered.

Version of July 11, 2020– Created July 11, 2020 - 11:00

Theorem 4.14. *Let $N \in \mathbb{N}$. Then the following are equivalent:*

(1) $\det G^{(N)} = 0$*;*

(2) There is a level N vector $v \in \mathcal{V}_\Delta$ which is orthogonal to all states in $\mathcal{V}_\Delta$;

(3) $\mathcal{V}_\Delta$ has a non-trivial subrepresentation with a nonzero level N state;

(4) There is a non-trivial singular vector (i.e. linearly independent from v_Δ) at a level $N' \leq N$.

In the literature, $\det G^{(N)}$ is referred to as the *Kac determinant*. To prove this theorem, we first need a short lemma.

Lemma 4.15. *Let $v \in \mathcal{V}_\Delta \setminus \{0\}$, and $n \in \mathbb{Z}_{>0}$. Then $L_{-n}v \neq 0$.*

Proof. First assume that v is a level N vector for some $N \in \mathbb{N}$. We can then write $v = \sum_{b\in\mathcal{B}^{(N)}} \lambda_b b v_\Delta$, with all $\lambda_b \in \mathbb{C}$. For such a $b \in \mathcal{B}^{(N)}$, we can define the length of b as the number of L_i's it consists of (e.g. $L_{-4}L_{-2}$ has length 2, and L_{-1}^3 has length 3). We can immediately do some observations concerning this newly introduced length.

- If b has length ℓ, then $L_{-n}b$, when written in terms of the level $N+n$ basis $\mathcal{B}^{(N+n)}$, consists of a single term of length $\ell+1$, and possibly other terms of smaller length. This can easily be seen by looking at how the commutator rules were used in (2.18). To write $L_{-n}b$ (which is a product of L_i's) in terms of $\mathcal{B}^{(N+n)}$, it needs to be properly ordered, and to this end repetitive swaps of the L_i are used. Such a swap consists of a substitution given by $L_iL_j = L_jL_i + [L_i, L_j]$, and obviously the term containing L_jL_i still has the same length as the original product, and the terms containing $[L_i, L_j]$ have a smaller length.
- If we take two $b, b' \in \mathcal{B}^{(N)}$ with the same length ℓ, then the terms of length $\ell+1$ obtained when writing $L_{-n}b$ and $L_{-n}b'$ in terms of $\mathcal{B}^{(N)}$ are only equal if b and b' are equal. This follows directly from the previous observation in which we showed these terms simply are $L_{-n}b$ and $L_{-n}b'$, but properly ordered *without* being concerned with the non-commutativity of the L_i. Hence they both only contain the L_i from b and b', and one additional L_{-n}. But if these $\ell+1$ length terms are equal, they contain the same L_i, so b and b' contain the same L_i, and are thus equal.
- Returning to v, let ℓ be the highest number such that there is a nonzero $b \in \mathcal{B}^{(N)}$ with length ℓ, and let k be the amount of such

terms. Then $L_{-n}v$ (when written in terms of $\mathcal{B}^{(N+n)}$) contains exactly k terms of length $\ell+1$, and by the previous point, these are all different. This means they cannot cancel out, so $L_{-n}v$ is nonzero.

If we now allow v to be an arbitrary state, it will still be a linear combination of vectors with a defined level (this follows from equation (4.11)), so we can write $v = \sum_{N=0}^{\infty} v_N$, with all v_N of level N and all but finitely many zero. It follows that $L_{-n}v = \sum_{N=0}^{\infty} L_{-n}v_N$, and since all the $L_{-n}v_N$ are nonzero (proven above) and linearly independent (as they are L_0-eigenvectors with different eigenvalues), $L_{-n}v$ is nonzero. □

We are now ready to prove theorem 4.14.

Proof. We first show the equivalence of (1) and (2), and then the equivalence of (2), (3) and (4).

(1)⇔(2) Assume $\det G^{(N)} = 0$. Then the rows of $G^{(N)}$ are linearly dependent, hence there exist nontrivial $\lambda_1, \dots \lambda_M \in \mathbb{C}$ such that for all $j \in \{1, \dots, M\}$ we have $\sum_{i=1}^{M} \langle \lambda_i b_i, b_j \rangle = 0$. It follows directly that if we define $v = \sum_{i=1}^{M} \overline{\lambda_i} b_i$, then for all $j \in \{1, \dots, M\}$

$$\langle v, b_j \rangle = \left\langle \sum_{i=1}^{M} \overline{\lambda_i} b_i, b_j \right\rangle = \sum_{i=1}^{M} \lambda_i \langle b_i, b_j \rangle = 0 \tag{4.23}$$

Therefore v is orthogonal to a full set of basis vectors of the level N subspace, and thus to the level N subspace itself. Since v is itself a level N vector, it is orthogonal to all other elements of $\mathcal{V}_\Delta$ with a different level, so by (4.11) it is orthogonal to the entirety of $\mathcal{V}_\Delta$.

Moreover, given a non-trivial level N vector v orthogonal to every other state, it can be written in terms of the basis $(v_1, \dots, v_M)$, and the reversal of the previous argument then gives us the required result.

(2)⇒(3) Let $v \in \mathcal{V}_\Delta$ be a level N vector orthogonal to all other states, and define $R = \{w \in \mathcal{V}_\Delta : w \text{ is orthogonal to all states in } \mathcal{V}_\Delta\}$. Then R is clearly a subspace, and since $v \in R$ and $v_\Delta \notin R$, it is neither 0 nor $\mathcal{V}_\Delta$. We still need to prove that R is closed under the action of $\mathfrak{V}_c$, but this can be easily done: let $w \in R$ and $n \in \mathbb{Z}$, and let $w' \in \mathcal{V}_\Delta$ be any vector. Then

$$\langle L_n w, w' \rangle = \langle w, L_{-n} w' \rangle = 0 \tag{4.24}$$

so $L_n w$ is orthogonal to every state in $\mathcal{V}_\Delta$. Therefore R is a nontrivial subrepresentation with a nonzero level N vector (namely v).

(3)⇒(4) Let $R \subseteq \mathcal{V}_\Delta$ be a non-trivial subrepresentation with a nonzero level N state v. We now take $N' \in \mathbb{N}$ to be the smallest number such that R has a nonzero level N' vector, and we let w be such a vector. Then clearly $w \neq v_\Delta$, as this would directly imply that $R = \mathcal{V}_\Delta$. Since for any positive $n \in \mathbb{Z}_{>0}$ the vector $L_n w$ is of a lower level than N (namely $N-n$), it has to be zero. But since this holds for all $n \in \mathbb{Z}_{>0}$ we have proven that w is a singular vector.

(4)⇒(2) Let $w \in \mathcal{V}_\Delta$ be a non-trivial singular vector at some level $N' \leq N$. If we now look at the level N vector $v = L_{N'-N} w$, then lemma 4.15 implies $v \neq 0$. Now for all level N vectors w' we have

$$\langle v, w' \rangle = \langle L_{N'-N} w, w' \rangle = \langle w, L_{N-N'} w' \rangle.$$

We can then write $L_{N-N'} w'$ in terms of the basis $\mathcal{B}^{(N')} v_\Delta$ as $L_{N-N'} w' = \sum_{b \in \mathcal{B}^{(N')}} \lambda_b b v_\Delta$, and focus on $\langle w, b v_\Delta \rangle$ for such a $b \in \mathcal{B}^{(N')}$. By definition, $b = L_{n_1} \dots L_{N_{m-1}} L_{N_m}$ for a nonempty set of negative $n_1, \dots, n_m \in \mathbb{Z}_{<0}$. Hence we have

$$\begin{aligned} \langle w, b \rangle &= \langle w, L_{n_1} \dots L_{N_{m-1}} L_{N_m} v_\Delta \rangle = \langle L_{N-m} v, L_{n_1} \dots L_{N_{m-1}} v_\Delta \rangle \\ &= \langle 0, L_{n_1} \dots L_{N_{m-1}} v_\Delta \rangle = 0 \end{aligned} \tag{4.25}$$

and this directly implies that

$$\langle v, w' \rangle = \langle w, L_{N-N'} w' \rangle = \sum_{b \in b^{(N)}} \lambda_b \langle w, b \rangle = \sum_{b \in b^{(N)}} \lambda_b \cdot 0 = 0. \tag{4.26}$$

Hence the level N vector v is orthogonal to all other level N vectors, and thus to every vector in $\mathcal{V}_\Delta$.

□

Theorem 4.14 is very powerful, as it gives us a clear way to predict and construct subrepresentations of Verma modules, namely by calculating the Kac determinant $\det G^{(N)}$. Luckily, this has been done in the most general way, and the result is summarised in the next theorem.

Theorem 4.16. *Let $N \in \mathbb{N}$. Then*

$$\det G^{(N)} = \lambda \prod_{\substack{r,s \in \mathbb{N} \\ rs \leq N}} (\Delta - \Delta_{r,s})^{P(N-rs)}, \tag{4.27}$$

where λ is some positive constant, $P(N-rs)$ is the partition function mentioned in the footnote on page 70, and

$$\Delta_{r,s} = \frac{1}{48}\left((13-c)(r^2+s^2) + \sqrt{(c-1)(c-25)}(r^2-s^2) - 24rs - 2 + 2c \right). \tag{4.28}$$

A proof of this theorem can be found in [FF82].

4.2.1 Unitarity

At the beginning of section 4.1, we defined lowest weight representations as representations in which the L_0-eigenvalues are bounded from below. The argument therefor was that in any quantum theory the energy levels are bounded from below, and hence, if this representation is the state space of some quantum theory, the eigenvalues of L_0 should be bounded. There are however other restrictions we can make. For instance, the eigenvalues of L_0 should be real, which implies $\Delta \in \mathbb{R}$. Moreover, for any Verma module contained in our representation, the "inner product" introduced in the previous section should indeed be an inner product, meaning it should be positive definite. Representations for which this is the case are called *unitary*. We should therefore determine for which values of Δ and c a Verma module is unitary. In the following, we briefly illustrate the consequences of this assumption.

First of all, for the vector $L_{-n}v_\Delta$, we have

$$\begin{aligned}\langle L_{-n}v_\Delta, L_{-n}v_\Delta\rangle &= \langle v_\Delta, L_nL_{-n}v_\Delta\rangle = \langle v_\Delta, (2nL_0 - \frac{c}{12}(n^3-n))v_\Delta\rangle \\ &= 2n\Delta + \frac{c}{12}(n^3-n).\end{aligned}$$

If a Verma module is unitary, the above expression should be a a positive real number for each $n \in \mathbb{N}_{>0}$. For it to be real, c has to be real as well. Now for any $c < 0$, the above expression becomes negative for large enough n, and therefore unitary representations only exist for $c \in \mathbb{R}_{\geq 0}$. Moreover, for $n = 1$ the above expression equals 2Δ, and we can therefore also assume that $\Delta > 0$. To determine for which remaining values of Δ and c the Verma module is unitary, we can use the Gramian matrices. Namely, the inner product is positive definite if and only if every Gramian matrix is. In particular, a zero of a Kac determinant (viewed as function of Δ) corresponds to a zero norm vector, and a negative Kac determinant implies the existence of negative norm vectors.

For $c > 1$, it can be shown the representation is *always* unitary, as proven in [DFMS12]. A sketch of how to prove this is the following. First one shows there are no zeros of the Kac determinant in the region $c > 1, \Delta > 0$. This

can be done by rewriting (4.28) to

$$\Delta_{r,s} = \frac{1-c}{96}\left(\left[(r+s)+(r-s)\sqrt{\frac{25-c}{1-c}}\right]^2 - 4\right), \tag{4.29}$$

and considering the two cases $c > 24$ and $24 \leq c < 1$. Then, one notes the number of negative eigenvalues of the Gramian matrix in two points $(c, \Delta), (c', \Delta')$ can only be different if they are separated by a curve where the Kac determinant vanishes. One then shows that for "high enough" values of Δ the Gramian matrices are indeed positive definite. These facts together prove that in the $c > 1$ region the Gramian matrices are always positive definite, and hence the theory is unitary [DFMS12]. In order to construct a consistent conformal field theory from such representations, the Verma modules occurring in a theory need to be *closed under fusion*. Fusion is a concept originating from the correlation functions between the different fields in the theory, but this is beyond the scope in this thesis. For $c > 1$ however, this means a consistent theory always consists of an infinite sum of Verma modules.

For $c < 1$, things are different. For $c < 1$ and $\Delta > 0$, the Verma module contains negative norm states, and is therefore *not* unitary [DFMS12]. We can however still look at quotients of such a Verma module. Namely, given a subrepresentation $R \subseteq \mathcal{V}_\Delta$, the quotient $\mathcal{V}_\Delta / R$ is again a $\mathfrak{V}_c$-representation. Such a quotient is of course only possible if $\mathcal{V}_\Delta$ is reducible, and theorem 4.14 and theorem 4.16 tell us when this is the case. To obtain a unitary quotient, the subrepresentation needs to contain all negative norm states. Although this almost never occurs, in the special cases where

$$c = 1 - \frac{6(p-q)^2}{pq} \tag{4.30}$$

for some coprime $p, q \in \mathbb{Z}_{\geq 2}$, interesting things happen. In this case, we can rewrite (4.28) into an easier expression. We namely have

$$13 - c = 12 + \frac{6(p-q)^2}{pq} = \frac{12pq}{pq} + \frac{6p^2 - 12pq + 6q^2}{pq} = 6\frac{p^2+q^2}{pq},$$
$$-2 + 2c = -12\frac{(p-q)^2}{pq},$$

and

$$\begin{aligned}\sqrt{(c-1)(c-25)} &= \sqrt{\frac{6(p-q)^2}{pq}\left(\frac{24pq}{pq}+\frac{6p^2-12pq+6q^2}{pq}\right)}\\ &= \sqrt{\frac{6(p-q)^2}{pq}\cdot\frac{6(p+q)^2}{pq}} = 6\frac{p^2-q^2}{pq},\end{aligned}$$

which implies that

$$\begin{aligned}48pq\Delta_{r,s} &= 12(p^2+q^2)(r^2+s^2)+12(p^2-q^2)(r^2+s^2)-24rs-(p-q)^2\\ &= 12(p^2r^2+q^2s^2-pqrs-(p-q)^2),\end{aligned}$$

and hence

$$\Delta_{r,s} = \frac{(pr-qs)^2-(p-q)^2}{4pq}. \tag{4.31}$$

It can then easily be shown the zeros of the Kac determinant obey the equalities

$$\begin{aligned}\Delta_{r,s} = \Delta_{q+r,p-s}, \qquad \Delta_{r,s}+rs = \Delta_{q-r,p-s} = \Delta_{q-r,p+s},\\ \Delta_{r,s}+(q-s)(p-s) = \Delta_{r,2p-s} = \Delta_{2q-r,s}.\end{aligned} \tag{4.32}$$

This implies that for c as in (4.30), a Verma module with $\Delta = \Delta_{r,s}$ for some $r,s \in \mathbb{N}$ contains infinitely many singular vectors. These singular vectors in their turn also generate Verma modules that are again reducible. As before, we can take quotients, and if $|q-p| = 1$, these are unitary [Gin88]. Moreover, given this fixed value of c, (4.32) tells us the following: when one is given a small set of Verma modules corresponding to low values of r and s, *every* reducible Verma module for this value of c is already contained in one of the modules one started with. This allows for consistent theories (i.e. theories closed under fusion) built from only a finite number (of quotients) Verma modules. Such theories are known as *minimal models*. Important examples are the model of the Lee-Yang singularity, for $(p,q) = (5,2)$, and the two-dimensional Ising model, where $(p,p') = (4,3)$. The theory for $(p,q) = (4,3)$ is however also a model for the free Majorana fermion ([DFMS12]), and this equivalence was used by Lars Onsager to find a first exact solution of the two-dimensional Ising model.

Version of July 11, 2020– Created July 11, 2020 - 11:00

Conclusion

In physics, there is a close connection between continuous symmetries and Lie algebras, which is twofold. On the one hand, the infinitesimal generators of symmetries in the classical theory form a Lie algebra. On the other hand, Noether's theorem implies that those symmetries correspond to conserved charges, such as energy and momentum. In the quantum theory, these charges turn into operators, which also form a Lie algebra. For conformal symmetries, we derived the corresponding classical algebra, namely the Witt algebra, by looking at the infinitesimal generators. We then argued that in the quantum theory, it is not the Witt algebra itself but a central extension that should occur; for the Witt algebra, the Virasoro algebra is the unique such central extension. As shown in this thesis, one important example of a quantum theory in which a representation of the Virasoro algebra is found, is the theory of the free open bosonic string.

To develop the string theory, we started with the Nambu-Goto action in $(1+d)$-dimensional Minkowski space. We then solved the equations of motion in the light-cone gauge. Here, instead of the usual d spacelike dimension and a single timelike dimension, one transitions to two lightlike and $d-1$ 'transverse' spacelike directions. In this solution, the transverse directions obey a wave equation, and are solved in terms of an infinite set of independent oscillation modes $\alpha_n^{\mathcal{I}}$ (with $n \in \mathbb{Z}$). Moreover, when assuming the motion in one lightlike direction to be linear, the motion in the second lightlike direction is fully determined by the Virasoro modes, defined as $L_n = \sum_{k\in\mathbb{Z}} \alpha_{n-k}^{\mathcal{I}} \alpha_k^{\mathcal{I}}$. In the quantum theory, the L_n turn into quantum operators. As shown in thesis, the Virasoro modes indeed obey the Virasoro commutation relations, which indicates string theory indeed has conformal symmetry. This can be seen explicitly from the Polyakov action.

Version of July 11, 2020– Created July 11, 2020 - 11:00

Instead of starting with a physical system, representations of the Virasoro algebra can also be studied directly in an abstract manner. Inspired by the open string, in such representations we view L_0 as the Hamiltonian, and hence assume the L_0-eigenvalues are bounded from below. The building blocks of these so-called lowest weight representations are the Verma modules. As shown in this thesis, a Verma module is reducible if and only if it contains a singular vector. In theorem 4.14 we showed this occurs exactly when a Kac determinant vanishes. Theorem 4.16 then gave an explicit formula for the Kac determinant, hence completely solidifying for which choices of Δ and c a Verma module is reducible. With this theory of Verma modules at hand, the next step would be to construct actual physical theories from these representations, the simplest of which are known as the minimal models.

Version of July 11, 2020– Created July 11, 2020 - 11:00

Bibliography

[And72] P. W. Anderson. More is different. *Science*, 177(4047):393–396, 1972.

[Bar64] V. Bargmann. Note on wigner's theorem on symmetry operations. *Journal of Mathematical Physics*, 5(7):862–868, 1964.

[BK13] G. W. Bluman and S. Kumei. *Symmetries and differential equations*, volume 81. Springer Science & Business Media, 2013.

[DFMS12] P. Di Francesco, P. Mathieu, and D. Sénéchal. *Conformal field theory*. Springer Science & Business Media, 2012.

[FF82] B. L. Feigin and D. B. Fuks. Invariant skew-symmetric differential operators on the line and verma modules over the virasoro algebra. *Functional Analysis and Its Applications*, 16(2):114–126, 1982.

[Gin88] P. Ginsparg. Applied conformal field theory. *arXiv preprint hep-th/9108028*, 1988.

[Hum12] J. E. Humphreys. *Introduction to Lie algebras and representation theory*, volume 9. Springer Science & Business Media, 2012.

[Knu93] D. E. Knuth. Johann faulhaber and sums of powers. *Mathematics of Computation*, 61(203):277–294, 1993.

[Rib14] S. Ribault. *Conformal field theory on the plane*. CEA Saclay, Institut de Physique Théorique, 2014.

[Ric16] D. Rickles. *A brief history of string theory: From dual models to M-theory*. Springer, 2016.

[Sch08] M. Schottenloher. *A mathematical introduction to conformal field theory*. Springer, Berlin, 2nd ed edition, 2008.

[Sne95] I. N. Sneddon. *Fourier transforms*. Courier Corporation, 1995.

[Wei11] T. Weigand. *Introduction to String Theory*. Institut für Theoretische Physik, Universität Heidelberg, Germany, 2011.

[Wened] K. Wendland. *Conformal Field Theory for Mathematicians*. Mathematics Institute, University of Freiburg, Germany, Unpublished.

[Zwi04] B. Zwiebach. *A first course in string theory*. Cambridge University Press, New York, 2004.