

Willoughby Seago

Notes from

Conformal Field Theory and Vertex Operator Algebras

October 8th, 2024

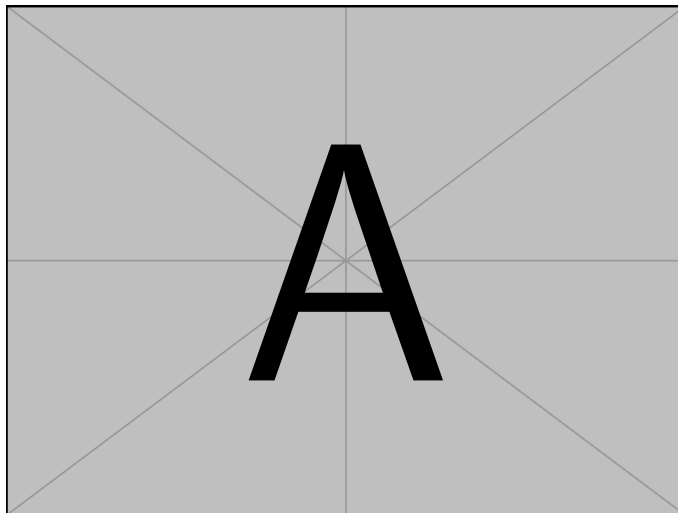
SMSTC

Conformal Field Theory and Vertex Operator Algebras

Willoughby Seago

October 8th, 2024

These are my notes from the SMSTC course *Conformal Field Theories and Vertex Operator Algebras* taught by Dr Anatoly Konechny. These notes were last updated at 22:35 on October 16, 2024.



Chapters

	Page
Chapters	ii
Contents	iii
1 Conformal Geometry	1
2 Quantum Field Theory	17
Appendices	33
A Differential Geometry	34

Contents

	Page
Chapters	ii
Contents	iii
1 Conformal Geometry	1
1.1 Local Conformal Maps	1
1.1.1 Conformal Maps Preserve Angles	2
1.2 Conformal Transformations on \mathbb{R}^2	3
1.2.1 Analytic Structure of Conformal Transformations	4
1.2.2 Conformal Transformations Examples	5
1.2.3 Global Conformal Transformations	8
1.3 Infinitesimal Conformal Transformations	13
2 Quantum Field Theory	17
2.1 Correlation Functions	17
2.2 Distributions	19
2.3 Functional Integration	20
2.4 Path Integrals	22
2.4.1 Computing Correlators for the Free Boson	23
2.5 Continuous Symmetries	30
Appendices	33
A Differential Geometry	34
A.1 Tangent Space	34
A.2 Pushforward and Pullback	34
A.3 Riemannian Manifolds	35

One

Conformal Geometry

1.1 Local Conformal Maps

Intuitively, we want conformal maps to preserve angles, but not necessarily distances. To do so consider how the angle between two vectors in, say, \mathbb{R}^3 is computed,

$$\cos(\theta_{\mathbf{u},\mathbf{v}}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}. \quad (1.1.1)$$

We see here that if each vector were made longer by some positive constant, ρ , then we have

$$\cos(\theta_{\rho\mathbf{u},\rho\mathbf{v}}) \frac{\rho\mathbf{u} \cdot \rho\mathbf{v}}{\|\rho\mathbf{u}\| \|\rho\mathbf{v}\|} = \cos(\theta_{\mathbf{u},\mathbf{v}}). \quad (1.1.2)$$

We can think of scaling all of the vectors here by ρ as the same as scaling the metric by $1/\rho$. It's $1/\rho$ because if we make the units of a measurement smaller then the number we measure gets bigger (hence, *contravariant* vectors).

The following is really just fancy differential-geometry-speak for this rescaling of the metric, and we also allow the scaling of the metric to depend on position.

Definition 1.1.3 — Local Conformal Map Let (\mathcal{M}_1, g_1) and (\mathcal{M}_2, g_2) be n -dimensional Riemannian manifolds. Let $U_1 \subseteq \mathcal{M}_1$ and $U_2 \subseteq \mathcal{M}_2$ be open subsets. A (local) **conformal transformation** is a smooth, injective map, $\varphi : U_1 \rightarrow U_2$, satisfying the pullback condition

$$\varphi^* g_2 = \Lambda g_1 \quad (1.1.4)$$

for some function $\Lambda : U_1 \rightarrow \mathbb{R}_{>0}$.

A conformal map defined on all of \mathcal{M}_1 is a global conformal map.

Remark 1.1.5 It is possible to relax the conditions on φ , and require only that it is differentiable. However, requiring smoothness and injectivity is common when it comes to applications, so we make it a basic requirement. It's also common to further restrict to orientation-preserving maps, but we won't do that just yet.

We can express the pullback condition, $\varphi^*g_2 = \Lambda g_1$, in local coordinates. Let $x = (x^1, \dots, x^n)$ be coordinates covering U_1 , and $y = (y^1, \dots, y^n)$ coordinates covering U_2 . Then φ is fully specified by the functions φ^i which are defined such that $y^i = \varphi^i(x)$. The metrics, g_i , may be specified by their components, $(g_i)_{jk} : U_i \rightarrow \mathbb{R}$. In these coordinates the pullback condition becomes

$$\sum_{k,l} \frac{\partial \varphi^k}{\partial x^i} \frac{\partial \varphi^l}{\partial x^j} (g_2)_{kl}(\varphi(x)) = \Lambda(x) (g_1)_{ij}(x). \quad (1.1.6)$$

Taking determinants of either side of this equation we have

$$\det\left(\frac{\partial \varphi^k}{\partial x^i}\right) \det(g_2) \det\left(\frac{\partial \varphi^l}{\partial x^j}\right) = \Lambda^n \det(g_1). \quad (1.1.7)$$

Now, $\Lambda^n \neq 0$ and $\det(g_i) \neq 0$, so it follows that $\det(\partial \varphi^k / \partial x^i) \neq 0$, meaning that the matrix with components $\partial \varphi^k / \partial x^i$ is invertible. This means that a conformal map is always **locally invertible**. That is, for any $p \in U_1$ we have a neighbourhood $V_1 \subseteq U_1$ with $p \in V_1$ such that φ restricted to V_1 is a bijection.

Note that it's possible to be locally invertible but not fully invertible. There may be a point in $U_1 \setminus V_1$ which maps to the same point as a point in V_1 , so the function will not be injective.

1.1.1 Conformal Maps Preserve Angles

Consider two vectors $u, v \in T_p \mathcal{M}_1$ and some $p \in \mathcal{M}_1$. Taking some open neighbourhood of p , $U_1 \subseteq \mathcal{M}_1$, we can also take coordinates $x = (x^1, \dots, x^n)$ covering U_1 . This gives a basis $\{\partial / \partial x^i|_p\}$ for $T_p \mathcal{M}_1$. The angle between these vectors is given, as in \mathbb{R}^3 , by

$$\cos(\theta_{u,v}) = \frac{(u, v)}{\|u\| \|v\|} = \frac{u^i (g_1)_{ij} v^j}{\sqrt{u^l (g_1)_{lk} u^k v^p (g_1)_{pq} v^q}}. \quad (1.1.8)$$

Here we've started to employ the Einstein summation convention, and we shall do so from now on. Let $\varphi : U_1 \rightarrow U_2$ be a conformal transformation and suppose that U_2 is covered by coordinates $y = (y^1, \dots, y^n)$. Consider the pushforward

$$d\varphi_p : T_p \mathcal{M}_1 \rightarrow T_{\varphi(p)} \mathcal{M}_2. \quad (1.1.9)$$

Under this the vectors u and v map to the vectors

$$\tilde{u} = d\varphi_p(u), \quad \text{and} \quad \tilde{v} = d\varphi_p(v), \quad (1.1.10)$$

which have coordinates

$$\tilde{u}^i = \frac{\partial \varphi^i}{\partial x^j} u^j, \quad \text{and} \quad \tilde{v}^i = \frac{\partial \varphi^i}{\partial x^j} v^j. \quad (1.1.11)$$

We can now calculate the angle between these vectors as follows:

$$\cos(\theta_{\tilde{u}, \tilde{v}}) = \frac{(\tilde{u}, \tilde{v})}{\|\tilde{u}\| \|\tilde{v}\|} \quad (1.1.12)$$

$$= \frac{\tilde{u}^a(g_2)_{ab} \tilde{v}^b}{\sqrt{\tilde{u}^c(g_2)_{cd} \tilde{u}^d \tilde{v}^e(g_2)_{ef} \tilde{v}^f}} \quad (1.1.13)$$

$$= \frac{\frac{\partial \varphi^a}{\partial x^i} u^i(g_2)_{ab} \frac{\partial \varphi^b}{\partial x^j} u^j}{\sqrt{\frac{\partial \varphi^c}{\partial x^l} u^l(g_2)_{cd} \frac{\partial \varphi^d}{\partial x^k} u^k \frac{\partial \varphi^e}{\partial x^p} v^p(g_2)_{ef} \frac{\partial \varphi^f}{\partial x^q} v^q}} \quad (1.1.14)$$

$$= \frac{u^i \Lambda(g_1)_{ij} v^j}{\sqrt{u^l \Lambda(g_1)_{lk} u^k v^p \Lambda(g_1)_{pq} v^q}} \quad (1.1.15)$$

$$= \frac{u^i(g_1)_{ij} v^j}{\sqrt{u^l(g_1)_{lk} u^k v^p(g_1)_{pq} v^q}} \quad (1.1.16)$$

$$= \cos(\theta_{u,v}) \quad (1.1.17)$$

where we've used the pullback condition

$$\frac{\partial \varphi^a}{\partial x^i}(g_2)_{ab} \frac{\partial \varphi^b}{\partial x^j} = \Lambda(g_1)_{ij}. \quad (1.1.18)$$

This shows that conformal transformations really do preserve angles as we were looking for.

1.2 Conformal Transformations on \mathbb{R}^2

We now restrict ourselves to the case $\mathcal{M}_1 = \mathcal{M}_2 = \mathbb{R}^2$ with the standard Euclidean metric. In Cartesian coordinates, (x^1, x^2) , the metric corresponds to the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (1.2.1)$$

It turns out that it's useful to identify \mathbb{R}^2 with \mathbb{C} using the complex coordinates

$$z = x^1 + ix^2, \quad \text{and} \quad \bar{z} = x^1 - ix^2. \quad (1.2.2)$$

In the coordinates (z, \bar{z}) the metric corresponds to the matrix

$$\frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (1.2.3)$$

To see this notice that

$$x^1 = \frac{1}{2}(z + \bar{z}), \quad \text{and} \quad x^2 = \frac{1}{2i}(z - \bar{z}). \quad (1.2.4)$$

Thus, by the chain rule, we have

$$dx^1 = \frac{\partial x^1}{\partial z} dz + \frac{\partial x^1}{\partial \bar{z}} d\bar{z}, \quad \text{and} \quad dx^2 = \frac{\partial x^2}{\partial z} dz + \frac{\partial x^2}{\partial \bar{z}} d\bar{z}. \quad (1.2.5)$$

Inserting this into the usual Euclidean metric we have

$$g = (dx^1)^2 + (dx^2)^2 \quad (1.2.6)$$

$$= \left(\frac{\partial x^1}{\partial z} dz + \frac{\partial x^1}{\partial \bar{z}} d\bar{z} \right)^2 + \left(\frac{\partial x^2}{\partial z} dz + \frac{\partial x^2}{\partial \bar{z}} d\bar{z} \right)^2 \quad (1.2.7)$$

$$= \left(\frac{1}{2} dz + \frac{1}{2} d\bar{z} \right)^2 + \left(\frac{1}{2i} dz - \frac{1}{2i} d\bar{z} \right)^2 \quad (1.2.8)$$

$$= dz d\bar{z}. \quad (1.2.9)$$

Thus, we must have that

$$g_{zz} dz dz = g_{\bar{z}\bar{z}} d\bar{z} d\bar{z} = 0, \quad \text{and} \quad g_{z\bar{z}} dz d\bar{z} + g_{\bar{z}z} d\bar{z} dz = 1. \quad (1.2.10)$$

The requirement that g is symmetric implies the matrix form must be symmetric, which is how we end up with the 1 evenly split between the off-diagonal elements.

1.2.1 Analytic Structure of Conformal Transformations

In the complex coordinates, (z, \bar{z}) , we may consider a conformal transformation, $\varphi: U \rightarrow V$, as consisting of two maps,

$$z \mapsto \varphi(z, \bar{z}), \quad \text{and} \quad \bar{z} \mapsto \bar{\varphi}(z, \bar{z}) = \overline{\varphi(z, \bar{z})}. \quad (1.2.11)$$

Since the second is just the conjugate of the first we need only specify one of these maps.

We can now write out the pullback condition and check what this imposes on φ :

$$\varphi^*(dz, d\bar{z}) = d\varphi \overline{d\varphi} \quad (1.2.12)$$

$$= \left(\frac{\partial \varphi}{\partial z} dz + \frac{\partial \varphi}{\partial \bar{z}} d\bar{z} \right) \left(\frac{\partial \bar{\varphi}}{\partial z} dz + \frac{\partial \bar{\varphi}}{\partial \bar{z}} d\bar{z} \right) \quad (1.2.13)$$

$$= \frac{\partial \varphi}{\partial z} \frac{\partial \bar{\varphi}}{\partial z} (dz)^2 + \frac{\partial \varphi}{\partial z} \frac{\partial \bar{\varphi}}{\partial \bar{z}} dz d\bar{z} + \frac{\partial \varphi}{\partial \bar{z}} \frac{\partial \bar{\varphi}}{\partial z} d\bar{z} dz + \frac{\partial \varphi}{\partial \bar{z}} \frac{\partial \bar{\varphi}}{\partial \bar{z}} (d\bar{z})^2$$

$$= \left(\frac{\partial \varphi}{\partial z} \frac{\partial \bar{\varphi}}{\partial \bar{z}} + \frac{\partial \varphi}{\partial \bar{z}} \frac{\partial \bar{\varphi}}{\partial z} \right) dz d\bar{z} + \frac{\partial \varphi}{\partial z} \frac{\partial \bar{\varphi}}{\partial z} (dz)^2 + \frac{\partial \varphi}{\partial \bar{z}} \frac{\partial \bar{\varphi}}{\partial \bar{z}} (d\bar{z})^2. \quad (1.2.14)$$

Introducing the notation $\partial_w f = \partial f / \partial w$ we can now use

$$\overline{\partial_w f} = \frac{\partial \bar{f}}{\partial \bar{w}} = \frac{\partial \bar{f}}{\partial \bar{w}} = \partial_{\bar{w}} \bar{f} \quad (1.2.15)$$

to rewrite this in terms of $\partial \varphi / \partial z$ and $\partial \varphi / \partial \bar{z}$:

$$\varphi^*(dz, d\bar{z}) = ((\partial_z \varphi)(\overline{\partial_z \varphi}) + (\partial_{\bar{z}} \varphi)(\overline{\partial_{\bar{z}} \varphi})) dz d\bar{z} + (\partial_z \varphi)(\overline{\partial_{\bar{z}} \varphi})(dz)^2 + (\partial_{\bar{z}} \varphi)(\overline{\partial_z \varphi})(d\bar{z})^2. \quad (1.2.16)$$

Now, if the pullback condition is to hold we must have

$$\varphi^*(dz, d\bar{z}) = \Lambda(z, \bar{z}) dz d\bar{z}. \quad (1.2.17)$$

So, we require that the $(dz)^2$ and $(d\bar{z})^2$ terms vanish, which imposes

$$(\partial_z \varphi)(\overline{\partial_{\bar{z}} \varphi}) = 0. \quad (1.2.18)$$

Thus, we must have

$$\partial_z \varphi = 0, \quad \text{or} \quad \partial_{\bar{z}} \varphi = 0. \quad (1.2.19)$$

The second of these says that φ has no \bar{z} dependence, and we know that φ is (at least) real differentiable, and thus φ is complex differentiable (with respect to z) and so φ is holomorphic. The first of these says that φ has no z dependence, and we know that φ is (at least) real differentiable, and thus φ is complex differentiable with respect to \bar{z} , and so φ is antiholomorphic (that is, it's holomorphic as a function of \bar{z}). We therefore have two cases

- $\varphi = \varphi(z)$ is holomorphic; or
- $\varphi = \varphi(\bar{z})$ is antiholomorphic.

We are then left with the condition

$$(\partial_z \varphi)(\overline{\partial_z \varphi}) + (\partial_{\bar{z}} \varphi)(\overline{\partial_{\bar{z}} \varphi}) = |\partial_z \varphi|^2 + |\partial_{\bar{z}} \varphi|^2 = \Lambda(z, \bar{z}) > 0, \quad (1.2.20)$$

which must hold for all $z \in U$.

The differential, $d\varphi$, of a map $\varphi : U \rightarrow V$ in complex coordinates has the matrix representation

$$d\varphi = \begin{pmatrix} \partial_z \varphi & \partial_{\bar{z}} \varphi \\ \partial_z \bar{\varphi} & \partial_{\bar{z}} \bar{\varphi} \end{pmatrix}. \quad (1.2.21)$$

Thus, in the holomorphic case we have

$$d\varphi = \begin{pmatrix} \partial_z \varphi & 0 \\ 0 & \partial_{\bar{z}} \bar{\varphi} \end{pmatrix}, \quad \text{and} \quad \det(d\varphi) = |\partial_z \varphi|^2 > 0, \quad (1.2.22)$$

assuming that the derivative doesn't vanish as required to have a positive scaling of the metric. In the antiholomorphic case, making the same non-vanishing derivative assumption, we have

$$d\varphi = \begin{pmatrix} 0 & \partial_{\bar{z}} \varphi \\ \partial_z \bar{\varphi} & 0 \end{pmatrix}, \quad \text{and} \quad \det(d\varphi) = -|\partial_{\bar{z}} \varphi|^2 < 0. \quad (1.2.23)$$

The interpretation of these is that holomorphic conformal maps preserve orientation, whereas antiholomorphic conformal maps reverse orientation. This shouldn't be surprising, the most basic antiholomorphic map is $z \mapsto \bar{z}$, which is a reflection in the real axis, so reverses orientation.

We state these results now as a theorem.

Theorem 1.2.24. Any local conformal orientation preserving (reversing) transformation $\varphi : U \rightarrow \mathbb{C}$ is specified by a holomorphic (antiholomorphic) function $z \mapsto \varphi(z)$ ($z \mapsto \varphi(\bar{z})$) which is invertible everywhere on $\varphi(U)$.

1.2.2 Conformal Transformations Examples

We can now ask what conformal transformations $\varphi : \mathbb{C} \rightarrow \mathbb{C}$ exist. We'll also consider conformal transformations defined only on $\mathbb{C} \setminus \{\text{pt}\}$ where pt is some point where the transformation need not be defined.

1.2.2.1 Rigid Transformations

The rigid transformations, or **isometries** of \mathbb{C} are the distance preserving transformations. As such they preserve the metric, and thus are conformal with $\Lambda = 1$. It is well known that the isometries of the plane fall into three different classes.

Translation. A translation, $T_b : \mathbb{C} \rightarrow \mathbb{C}$, by $b \in \mathbb{C}$ is defined by

$$T_b(z) = z + b. \quad (1.2.25)$$

This clearly preserves the metric since the distance between z and w is $|z - w|$, and the distance between $T_b(z)$ and $T_b(w)$ is $|T_b(z) - T_b(w)| = |z + b - w - b| = |z - w|$. Clearly T_b is bijective and smooth also, and it's defined everywhere, so it's a global conformal transformation.

Translations form a group, which is simply the additive group $(\mathbb{C}, +)$.

Rotation. An anticlockwise rotation, $R_{\alpha,0} : \mathbb{C} \rightarrow \mathbb{C}$, by an angle α about the origin, 0, is defined by

$$R_{\alpha,0}(z) = e^{i\alpha}z. \quad (1.2.26)$$

This preserves the metric, since $|R_{\alpha,0}(z) - R_{\alpha,0}(w)| = |e^{i\alpha}z - e^{i\alpha}w| = |e^{i\alpha}(z - w)| = |z - w|$. Again, this is bijective, smooth, and defined everywhere, so it is also a global conformal transformation.

Rotations about the origin form a group, which is simply $SO(2, \mathbb{R}) \cong U(1)$.

A rotation about some point, $c \in \mathbb{C}$, by an angle α is simply given by

$$R_{\alpha,c} = T_c \circ R_{\alpha,0} \circ T_{-c}. \quad (1.2.27)$$

That is, we first translate the centre of rotation to the origin, then rotate, then translate back. This is again a conformal transformation, as all composites of conformal transformations are.

Rotations and translations form a group.

Reflection. A reflection, $C_{0,0} : \mathbb{C} \rightarrow \mathbb{C}$, in the real axis, that is $x^2 = 0$, is given by conjugation

$$C_{0,0}(z) = \bar{z}. \quad (1.2.28)$$

This preserves the metric, since $|C_{0,0}(z) - C_{0,0}(w)| = |\bar{z} - \bar{w}| = |\overline{z - w}| = |z - w|$. Again, this is bijective, smooth, and defined everywhere, so it is also a global conformal transformation.

Reflections in the real axis form a group isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

A reflection in an arbitrary line can similarly be given as a composition of previously defined conformal transformations. To do so we parametrise the line by a point, $c \in \mathbb{C}$, through which it passes, and the angle, α , that the line makes to the positive real axis. Then a reflection in this line is given by

$$C_{\alpha,c} = T_b \circ R_{-\alpha,0} \circ C_{0,0} \circ R_{\alpha,0} \circ T_{-b}. \quad (1.2.29)$$

This is again a conformal transformation.

1.2.2.2 Dilation

A **dilation** or **dilatation** is a global conformal transformation, $D_\rho : \mathbb{C} \rightarrow \mathbb{C}$, for $\rho \in \mathbb{R}_{>0}$, defined by

$$D_\rho(z) = \rho z. \quad (1.2.30)$$

Intuitively, for $\rho > 1$ this corresponds to stretching space out evenly in all directions, and for $\rho < 1$ it's a contraction. Clearly this is bijective and smooth, and it is conformal with $\Lambda = \rho^2$, since $d(\rho z) d(\rho \bar{z}) = \rho^2 dz d\bar{z}$.

Dilations form a group, which is just the multiplicative group $(\mathbb{C}^\times, \cdot)$.

Just looking at the formulae we see that rotations about the origin and dilations are very similar, and in fact, we can combine them into a single transformation $D_a : \mathbb{C} \rightarrow \mathbb{C}$ with $a \in \mathbb{C}$ and $a = \rho e^{i\alpha}$ for $\rho \in \mathbb{R}_{>0}$ and $\alpha \in \mathbb{R}$ and

$$D_a(z) = az = \rho e^{i\alpha} z, \quad (1.2.31)$$

which is a dilation by ρ and a rotation by α (the order of which is not important). That is, $D_a = D_\rho \circ R_{\alpha,0} = R_{\alpha,0} \circ D_\rho$, and $R_{\alpha,0} = D_{e^{i\alpha}}$.

The combination of all translations, dilations, and rotations forms a group, $\text{Aff}(\mathbb{C})$, the affine group.

1.2.2.3 Circle Inversion

One example of an antiholomorphic conformal transformation is the map $\mathcal{J}_{0,R} : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C} \setminus \{0\}$, where $R \in \mathbb{R}_{>0}$, defined by

$$z \mapsto \frac{R^2}{\bar{z}}. \quad (1.2.32)$$

We can show that this is conformal by computing the metric after the transformation:

$$d\left(\frac{R^2}{\bar{z}}\right) d\left(\frac{R^2}{z}\right) = \frac{R^4}{|z|^4} dz d\bar{z}. \quad (1.2.33)$$

So, this is conformal with $\Lambda = R^4/|z|^4$.

This transformation has a nice geometric picture, as seen in [Figure 1.1](#). A point inside the unit circle is mapped to a point outside the unit circle, and that point is mapped to the original point inside the unit circle. We call this **inversion**, it swaps the inside and outside of the unit circle (removing 0 from the interior for the map to be defined). This swapping can even be done geometrically. Given a point, P , in the unit circle draw a line through the origin, O , and P . Then this map swaps P with the point P' which is such that $|OP||OP'| = R^2$.

Notice that $\mathcal{J}_{0,R} \circ \mathcal{J}_{0,R} = \text{id}_{\mathbb{C} \setminus \{0\}}$, so this is invertible everywhere it's defined.

It's possible to do an inversion in a circle centred elsewhere by translation, define $\mathcal{J}_{c,R} : \mathbb{C} \setminus \{c\} \rightarrow \mathbb{C} \setminus \{c\}$ by $\mathcal{J}_{c,R} = T_c \circ \mathcal{J}_{0,R} \circ T_{-c}$.

We can get a holomorphic version of this transformation by composing with C , since $C \circ \mathcal{J}_{0,R}$ acts as $z \mapsto R^2/z$ which is holomorphic on $\mathbb{C} \setminus \{0\}$.

1.2.2.4 Special Conformal Transformations

We have covered almost all conformal transformations of the plane now. It turns out that there's just one more type, known as **special conformal transformations**. To arrive at this one may consider the composite

$$\mathcal{J}_{0,R} \circ T_a \circ \mathcal{J}_{0,R}(z) = \mathcal{J}_{0,R} \circ T_a \left(\frac{R^2}{\bar{z}} \right) = \mathcal{J}_{0,R} \left(\frac{R^2}{\bar{z}} + a \right) = \frac{R^2}{\frac{R^2}{\bar{z}} + a} = \frac{z}{\frac{a}{R^2}z + 1}. \quad (1.2.34)$$

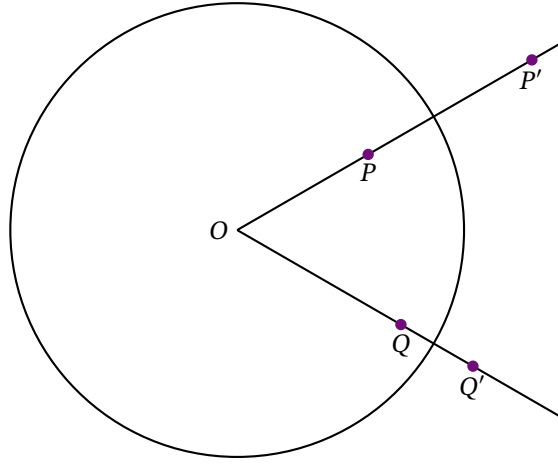


Figure 1.1: Circle inversion, the interior and exterior of the circle are swapped. Here, for example, $R = 3$, $|OP| = 2$, and so P' was chosen with $|OP'| = 9/2$ so that $|OP||OP'| = 9$. Similarly, $|OQ| = 2.5$ and so Q' was chosen with $|OQ'| = 9/2.5$. Notice that points closer to the origin map to points further away from the circle and points near the circle map to points near the circle.

Since \bar{a}/R^2 can take on any complex value as a and R vary we see that a special conformal transformation, $S_c : \mathbb{C} \setminus \{-1/c\} \rightarrow \mathbb{C}$ with $c \in \mathbb{C} \setminus \{0\}$ is given by

$$S_c(z) = \frac{z}{cz + 1}. \quad (1.2.35)$$

For $c = 0$ $S_0(z) = z/(0z + 1) = z$, so $S_0 = \text{id}_{\mathbb{C}}$ is defined everywhere.

Note that when considering the composite $\mathcal{I}_{0,R} \circ T_a \circ \mathcal{I}_{0,R}$ we should, *a priori*, remove two points from the domain, 0 and $-R^2/\bar{a}$. However, after going through the calculation we see that actually the resulting expression is defined at 0. Thus, we may allow 0 after all. We will, in general, always take conformal transformations to be defined on the largest possible domain. In this case we can think of this as an analytic continuation of the special conformal transformation to the point 0. This is a hint that we need to better handle these undefined points, which we do in .

1.2.3 Global Conformal Transformations

Restricting to orientation preserving (holomorphic) conformal transformations it turns out that the requirement that a global conformal transformation is defined everywhere is actually quite restrictive. It leaves us with just rotations, dilations, and translations. Together these form the affine group, $\text{Aff}(\mathbb{C})$. A general affine transformation $\varphi : \mathbb{C} \rightarrow \mathbb{C}$, is of the form

$$\varphi(z) = az + b \quad (1.2.36)$$

for some $a \in \mathbb{C}^\times$ and $b \in \mathbb{C}$. It is a result of complex analysis that these are the only orientation-preserving global conformal transformations.

Lemma 1.2.37 Affine transformations are the only orientation-preserving global conformal transformations of \mathbb{C} .

Proof. We know that an orientation-preserving global conformal transformation must be a holomorphic function of \mathbb{C} which is invertible everywhere. Such a function may not have branch cuts or singularities, as these are points at which it is undefined. In a neighbourhood of an essential singularity the image of the function is the entirety of \mathbb{C} , thus, if the function is to be defined outside of this neighbourhood as well, and also be injective there can be no such essential singularities. Thus, φ must be meromorphic, and so can be written as

$$\varphi(z) = \frac{P(z)}{Q(z)} \quad (1.2.38)$$

where $P, Q \in \mathbb{C}[z]$ are polynomials, and we may assume that they don't share any common factors. Since we aren't allowed any poles we must actually have that $Q(z) = c$ is just a constant polynomial, and we can rescale P to take $Q(z) = 1$. So, $\varphi(z) = P(z)$ is a polynomial. This polynomial cannot have more than one distinct zero, since this would violate the injectivity requirement. Hence, $\varphi(z) = P(z) = a(z - z_0)^n$ for some $a, z_0 \in \mathbb{C}$ and $n \in \mathbb{Z}_{>0}$. We can show that actually we must have $n \leq 1$, since if $n \geq 2$ then such a map is not injective, mapping both $z = z_0 + 1$ and $z = z_0 - 1$ to 1. Thus, $\varphi(z) = P(z) = a(z - z_0)^n$ for $n = 0, 1$, and thus $\varphi(z)$ is linear. One further application of injectivity tells us that $\varphi(z) = b$ ($b \in \mathbb{C}$) is not an acceptable solution, and thus $\varphi(z) = az + b$ for some $a \in \mathbb{C}^\times$ and $b \in \mathbb{C}$, so φ is an affine transformation. \square

If we instead allow global transformations to be defined on \mathbb{C} minus a finite set of points then we can expand the allowed transformations a little bit. As with the proof above branch cuts aren't allowed, they require infinitely many undefined points. We also don't allow essential singularities by the same argument on injectivity as in the proof. Thus, we're left with $\varphi(z) = P(z)/Q(z)$ where P and Q are polynomials. We require that the set of zeros of Q coincides with the set of removed points. Injectivity requires that P has a unique zero of order 1, so it must be linear. Injectivity also requires that Q has a unique zero, if it had more than one then a small circle around the corresponding poles would map to a large circle, and it would be possible to choose circles around two different poles which mapped to the same large circle. Further, the order of this zero of Q must be 1, else a small circle around the corresponding pole would map to a circle that winds around more than once, again, failing to be injective. Thus, P and Q must be linear, and we are left only with transformations

$$z \mapsto \frac{az + b}{cz + d} \quad \text{where } ad - bc \neq 0. \quad (1.2.39)$$

These are called **Möbius transformations**. The condition that $ad - bc \neq 0$ is required for this map to be invertible. Such a transformation is defined on $\mathbb{C} \setminus \{-d/c\}$ if $c \neq 0$, or \mathbb{C} if $c = 0$.

If we have Möbius transformations

$$z \mapsto \frac{az + b}{cz + d}, \quad \text{and} \quad z \mapsto \frac{ez + f}{gz + h} \quad (1.2.40)$$

with $a, b, c, d, e, f, g, h \in \mathbb{C}$ such that $ab - cd \neq 0$ and $eh - fg \neq 0$ then the composite performed in the order above is

$$z \mapsto \frac{(ae + bg)z + (af + bh)}{(ce + dg)z + (cf + dh)}. \quad (1.2.41)$$

We see that this is again a Möbius transformation since we can show that the condition on invertibility is $(bc - ad)(fg - eh) \neq 0$, which is true. Note that while these Möbius transformations are individually defined on $\mathbb{C} \setminus \{-d/c\}$ and $\mathbb{C} \setminus \{-h/g\}$ the composite may be taken to be defined on $\mathbb{C} \setminus \{-(cf + dh)/(ce + dg)\}$, despite passing through undefined points in the composite.

Noticing that $ad - bc$ looks like a determinant we recognise that we may package the parameters of the Möbius transformation into a matrix,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (1.2.42)$$

with $\det A \neq 0$. That is, $A \in \text{GL}(2, \mathbb{C})$. Further, a simple calculation shows that if we have two such matrices then their product exactly corresponds to the composition of the corresponding matrices. That is, we have a homomorphism $\text{GL}(2, \mathbb{C}) \rightarrow G$ where G is the group of all Möbius transformations under composition.

This mapping is surjective, but not injective, since, for example, with $A \in \text{GL}(2, \mathbb{C})$ both A and λA (for $\lambda \in \mathbb{C}^\times$) map to the same Möbius transformations since

$$\frac{az + b}{cz + d} = \frac{\lambda az + \lambda b}{\lambda cz + \lambda d}. \quad (1.2.43)$$

We can use this rescaling to always choose our parameters such that $ad - bc = 1$. That is, we can restrict our consideration to matrices of unit determinant. This gives a homomorphism $\text{SL}(2, \mathbb{C}) \rightarrow G$. Note that under this correspondence the following are a translation, $z \mapsto z + b$, dilation, $z \mapsto az$, special conformal transformation, $z \mapsto z/(cz + 1)$, and orientation-preserving inversion, $z \mapsto 1/z$:

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} a^{1/2} & 0 \\ 0 & a^{-1/2} \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \quad (1.2.44)$$

This map is still not injective. If $A \in \text{SL}(2, \mathbb{C})$ then $-A \in \text{SL}(2, \mathbb{C})$ both A and $-A$ map to the same Möbius transformation. The solution is to consider the group¹

$$\text{PSL}(2, \mathbb{C}) := \text{SL}(2, \mathbb{C}) / \{\pm \mathbb{1}\} \quad (1.2.45)$$

called the **projective special linear group**. Then we finally have $\text{PSL}(2, \mathbb{C}) \cong G$. For the most part we will look at matrices in $\text{SL}(2, \mathbb{C})$, since it's easier than working with the cosets of $\text{PSL}(2, \mathbb{C})$, and we'll just remember that we have this overall freedom in the sign of the matrix.

¹ $\mathbb{1}$ is the identity matrix.

We will call translations, rotations/dilations, special conformal transformations, and the inversion $z \mapsto 1/z$ the **elementary conformal transformations**. Any Möbius transformation is generated by these elementary transformations.

An important fact about Möbius transformations is that any Möbius transformation maps a circle on the plane to either another circle, or a line (which we may view as a circle of infinite radius). This is really a fact about inversions and special conformal transformations, since translations, rotations, and dilations trivially map circles to circles.

1.2.3.1 Dealing With Infinity

Here we look at two ways to deal with infinity. The first is simpler, but doesn't really give us anything new. The second is more complicated, but gives us a new way to view the plane.

Extended Complex Numbers. Define the set

$$\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\} \quad (1.2.46)$$

where ∞ is a formal symbol. We can extend the natural structure of \mathbb{C} to \mathbb{C}_∞ by making the following definitions for $z \in \mathbb{C}$:

$$z + \infty = \infty + z = \infty, \quad \infty z = \infty, \quad \frac{1}{\infty} = 0, \quad \text{and} \quad \frac{1}{0} = \infty. \quad (1.2.47)$$

Then if we have a Möbius transformation, φ , defined by

$$\varphi(z) = \frac{az + b}{cz + d} \quad (1.2.48)$$

we can take it as a function $\mathbb{C}_\infty \rightarrow \mathbb{C}_\infty$ by defining

$$\varphi(-d/c) = \infty, \quad \text{and} \quad \varphi(\infty) = \frac{a}{c}. \quad (1.2.49)$$

Further, since this is now defined at $-d/c$ and ∞ we see that this is a bijection of \mathbb{C}_∞ .

This definition works for our set-theoretical concerns, but it doesn't really help with the geometry. Where is this point, ∞ , with respect to the other points of \mathbb{C} ? It should be "infinitely far away", but in what direction? The solution to these problems is to use the next method of dealing with infinity.

Conformal Compactification. The Riemann sphere from complex analysis gives a bijection mapping of \mathbb{C}_∞ to the unit sphere, S^2 . This mapping is by stereographic projection. The construction is as follows:

- Embed the plane in \mathbb{R}^3 with coordinates (u, v, w) as the $w = 0$ plane.
- Embed the sphere in \mathbb{R}^3 as the set $\{(u, v, w) \in \mathbb{R}^3 \mid u^2 + v^2 + w^2 = 1\}$.
- To map a point of the plane, (u, v, w) , to a point on the sphere draw a line from the point to the point $(0, 0, 1)$, which we think of as the north pole of the sphere. This line intersects the sphere in two points, the north pole and one other. Map (u, v, w) to this other point.

- To map the other way draw a line from the north pole to a point on the sphere (not equal to the north pole) and where this line intersects the plane is where the point on the sphere maps. The north pole maps to ∞ .

We call S^2 constructed in this way the conformal compactification of the plane, or the one-point compactification of the sphere. Conformal because, as we'll show shortly, this mapping is conformal, and one-point because we add a point corresponding to the north pole. Compactification because the sphere is compact when the plane wasn't.

After some geometry we end up with a map

$$s: S^2 \setminus \{N\} \rightarrow \mathbb{C} \quad (1.2.50)$$

$$(u, v, w) \mapsto \frac{u + iv}{1 - w} \quad (1.2.51)$$

where we've identified the $w = 0$ plane with \mathbb{C} taking the u and v axes as the real and imaginary axes respectively. The inverse of this mapping is

$$s^{-1}(x + iy) = \left(\frac{2x}{x^2 + y^2 + 1}, \frac{2y}{x^2 + y^2 + 1}, \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1} \right). \quad (1.2.52)$$

This mapping is conformal with respect to the standard metrics on S^2 and \mathbb{C} . The standard metric on S^2 is just the pullback of the Euclidean metric on \mathbb{R}^3 along the embedding map. So, to check that s is conformal it is sufficient to calculate the pullback $(s^{-1})^*$ of the standard metric of \mathbb{R}^3 . The calculation is straightforward, but a little heavy on the algebra. The result is that

$$\begin{aligned} (du)^2 + (dv)^2 + (dw)^2 &= \left(\frac{2dx}{r^2 + 1} - \frac{2x dr^2}{(r^2 + 1)^2} \right)^2 + \left(\frac{2dy}{r^2 + 1} - \frac{2y dr^2}{(r^2 + 1)^2} \right)^2 \\ &\quad + \left(d \left(\frac{r^2 - 1}{r^2 + 1} \right) \right)^2 \end{aligned} \quad (1.2.53)$$

$$= \frac{4}{(r^2 + 1)^2} (dx^2 + dy^2) \quad (1.2.54)$$

$$= \frac{4}{(|z|^2 + 1)^2} dz d\bar{z}. \quad (1.2.55)$$

Here $r^2 = x^2 + y^2$. This shows that s^{-1} is conformal with

$$\Lambda = \frac{4}{(|z|^2 + 1)^2}. \quad (1.2.56)$$

This means that if we have a Möbius transformation defined on $\mathbb{C} \setminus \{p\}$ we can uniquely extend it to a Möbius transformation defined on S^2 by taking $\hat{\phi} = s^{-1} \circ \phi \circ s$ on $S^2 \setminus \{N, s^{-1}(p)\}$, and then defining

$$\hat{\phi}(N) = s^{-1}(a/c), \quad \text{and} \quad \hat{\phi}(s^{-1}(p)) = N. \quad (1.2.57)$$

This allows us to identify N with ∞ from \mathbb{C}_∞ . Further, it can be shown that all conformal transformations of the sphere arise in this way, so we don't lose or gain anything by looking at S^2 in place of \mathbb{C} or \mathbb{R}^2 .

One way to view this process is that (\mathbb{C}, s) is a coordinate chart on S^2 covering all points except N . We can then get another coordinate chart (\mathbb{C}, \bar{s}) by repeating

the stereographic projection process but with the south pole instead. These two charts cover the sphere and the transition maps between them are smooth on their intersection, and as such S^2 is a manifold. On the overlap of these two charts they are related by the inversion $z \mapsto 1/z$. To see this note that if a point is on the top hemisphere of the sphere then stereographic projection from the north pole takes it to the outside of the circle in the plane, and if it is on the bottom hemisphere of the sphere then stereographic projection maps it to a point inside the circle. The equator maps to the circle. Projecting from the south pole instead reverses this.

Using this manifold structure we can look at the neighbourhood of the north pole by working in the $(\mathbb{C}, \mathfrak{s})$ chart. This let's us do things like take the tangent space $T_N S^2$. This is an advantage over \mathbb{C}_∞ where we don't have a sensible notion of a tangent space at infinity.

This relation between the two charts further shows that the standard metric on S^2 is locally conformally flat on S^2 , in particular, it's flat on $S^2 \setminus \{N\}$. By this we simply mean that a conformal transformation takes the sphere metric on $S^2 \setminus \{N\}$ to the standard flat metric on \mathbb{C} . In a sense, we can push all of the curvature to a single point, N . Locality is important here, there is no globally conformally flat metric, as there is a theorem of differential geometry (the Gauss–Bonnet theorem) that says that the curvature of a manifold, M , is $2\pi\chi(M)$ where χ is the Euler characteristic, which is 2 for S^2 , and so there cannot be globally zero curvature. This is fine in our case because the curvature is given by the integral

$$\int_M \sqrt{g} R dA \quad (1.2.58)$$

where R is the Ricci scalar. Taking R to be an appropriately scaled Dirac delta supported at N allows this to evaluate to 4π as needed while retaining zero curvature on $S^2 \setminus \{N\}$.

We will swap back and forth between \mathbb{C} , \mathbb{C}_∞ , and S^2 as needed, and probably refer to all three as “the plane” and by each others names.

1.3 Infinitesimal Conformal Transformations

Given an open subset, $U \subseteq \mathbb{R}^2$ a **vector field** on U is a map

$$\varepsilon : U \rightarrow T\mathbb{R}^2 \quad (1.3.1)$$

where $T\mathbb{R}^2$ is the tangent bundle. That is, ε assigns a tangent vector, ε_p , to each point $p \in U$. Let $x = (x^1, x^2)$ be coordinates covering U . Then we can take the standard basis for $T_p \mathbb{R}^2$, given by $\{\partial/\partial x^i|_p\}$, although we'll drop the $|_p$ from the notation. Then in these coordinates the vector field ε may be expressed as

$$\varepsilon^i(x) \frac{\partial}{\partial x^i} \quad (1.3.2)$$

where $\varepsilon^i : U \rightarrow \mathbb{R}$ are smooth functions (assuming we want a smooth vector field).

This operator (as with all tangent vectors) acts on functions defined on U . This action can be exponentiated to obtain a finite transformation, which is done by solving the differential equation

$$\frac{d}{dt} \varphi^\varepsilon(x, t) = \varepsilon(\varphi^\varepsilon(x, t)) \quad (1.3.3)$$

where $\varphi^\varepsilon : U \times I_0 \rightarrow \mathbb{R}$ is a one-parameter group of finite transformations labelled by a parameter $t \in I_0 = (a, b) \subseteq \mathbb{R}$ and we assume $0 \in I_0$. Moving to coordinates this gives us the differential equation

$$\frac{d}{dt} \varphi^j(x, t) = \varepsilon^i(x) \frac{\partial}{\partial x^i} \varphi^j(x, t). \quad (1.3.4)$$

We impose the initial condition that $\varphi^j(x, 0) = x^j$, so we start at the desired point. This has a solution so long as $|t| < \delta$ for some positive δ . This solution can be expressed formally as

$$\varphi^j(x, t) = e^{t\varepsilon^i(x)\partial_i} x^j \quad (1.3.5)$$

where the exponential is understood via its Taylor series. The finite transformation we end up with is

$$x^j \mapsto \varphi^j(x, t). \quad (1.3.6)$$

The infinitesimal transformation is given by truncating the power series after the first order term:

$$x^j \mapsto x^j + t\varepsilon^j(x). \quad (1.3.7)$$

A vector field, X , is called a **conformal Killing vector field** if the transformations $\varphi^X(x, t)$ are local conformal transformations for sufficiently small intervals of t containing $t = 0$. Using this definition one can show that this is equivalent to the equation

$$\mathcal{L}_\varepsilon g = \lambda g \quad (1.3.8)$$

where \mathcal{L}_ε is the Lie derivative, g is the standard metric on \mathbb{R}^2 , and $\lambda : U \rightarrow \mathbb{R}$ is a (not necessarily positive) function on U . This is the infinitesimal version of the pullback requirement.

In coordinates we can take the transformation $x^j \mapsto x^j + t\varepsilon^j(x)$ and compute the change in coordinates of the metric. Dropping all higher order terms we are left with

$$\partial_i \varepsilon_j + \partial_j \varepsilon_i = \lambda(x) \delta_{ij}. \quad (1.3.9)$$

Note that we are freely lowering all indices here since the standard metric is δ_{ij} . This is the coordinate version of the above statement with the Lie derivative having substituted δ_{ij} for the components of the standard metric. Taking the trace of this equation we are left with

$$\partial_i \varepsilon^i = \lambda(x). \quad (1.3.10)$$

Substituting this back into the previous equation we have

$$\partial_1 \varepsilon^2 = -\partial_2 \varepsilon^1, \quad \text{and} \quad \partial_1 \varepsilon^1 = \partial_2 \varepsilon^2. \quad (1.3.11)$$

If we take our coordinates to be x and y and let $\varepsilon = u(x, y) + iv(x, y)$ then we can see that these are exactly the Cauchy–Riemann equations of complex analysis. This means that the complex coordinate components $\varepsilon^z = \varepsilon(z)$ and $\varepsilon^{\bar{z}} = \overline{\varepsilon(z)}$

must be holomorphic. Note that we don't get antiholomorphic solutions. This is because the orientation reversing conformal transformations do not have infinitesimal versions, which is because they are in a different connected component of the Lie group of conformal transformations.

An infinitesimal local conformal transformation is then written as

$$z \mapsto z + \varepsilon(z), \quad \text{and} \quad \bar{z} \mapsto \bar{z} + \overline{\varepsilon(z)} \quad (1.3.12)$$

on some open set, U . The collection of all conformal Killing vector fields forms a real vector space, V_0 . We know that $\varepsilon(z)$ is holomorphic, and so analytic on U . This means that in a neighbourhood of any point ε has a convergent Taylor series. For the following we'll take U to be the punctured disc at the origin,

$$U = \{z \mid |z| < r \text{ and } z \neq 0\} \quad \text{where} \quad r > 0, \quad (1.3.13)$$

but, this will hold for any sufficiently nice U . We take the puncture disc to allow for ε to have a pole at 0. Then ε has a Laurent series in a neighbourhood of $z = 0$ given by

$$\varepsilon(z) = \sum_{n \in \mathbb{Z}} c_n z^{n+1}. \quad (1.3.14)$$

The fact that we have z^{n+1} rather than z^n is just a convention to make other notation line up later, since the sum is over all of \mathbb{Z} this doesn't change the sum at all. The coefficients c_n with $n \ll 0$ must vanish, since we assume that we only have a simple pole. The corresponding vector field is then given in the basis $\{\partial_z, \partial_{\bar{z}}\}$ for the tangent space at z by

$$\varepsilon^z \partial_z + \varepsilon^{\bar{z}} \partial_{\bar{z}} = - \sum_{n \in \mathbb{Z}} (c_n l_n + \bar{c}_n \bar{l}_n) \quad (1.3.15)$$

where the l_n and \bar{l}_n are the differential operators

$$l_n = -z^{n+1} \partial_z, \quad \text{and} \quad \bar{l}_n = -\bar{z}^{n+1} \partial_{\bar{z}}. \quad (1.3.16)$$

These operators are such that $\{l_n + \bar{l}_n, i(l_n - \bar{l}_n)\}_{n \in \mathbb{Z}}$ forms a (real) basis for V_0 .

A quick calculation shows that V_0 is closed under the commutator. As such V_0 is an (infinite-dimensional) real Lie algebra. The commutation relations of the basis vectors $l_n + \bar{l}_n$ and $i(l_n - \bar{l}_n)$ can be computed in a straightforward way, but they're somewhat complicated. It is nicer to state the commutation relations for l_n and \bar{l}_n separately. We can then reconstruct the commutation relations for $l_n + \bar{l}_n$ using linearity. To reconstruct the commutation relations for $i(l_n - \bar{l}_n)$ we need to allow complex scalars, which results in us looking at the complexification $V_0^{\mathbb{C}} \cong V_0 \oplus iV_0 \cong V_0 \otimes_{\mathbb{R}} \mathbb{C}$. The corresponding Lie algebra then has the (complex) basis $\{l_n, \bar{l}_n\}_{n \in \mathbb{Z}}$. These satisfy the commutation relations

$$[l_n, l_m] = (n-m)l_{n+m}, \quad [\bar{l}_n, \bar{l}_m] = (n-m)\bar{l}_{n+m}, \quad \text{and} \quad [l_n, \bar{l}_m] = 0. \quad (1.3.17)$$

We see from this that what we actually have is that $V_0^{\mathbb{C}}$ consists of two commuting copies of the Lie algebra $\{l_n\}_{n \in \mathbb{Z}}$, where the second copy is marked with a bar to distinguish it. From basic Lie theory we recognise that this means that $V_0^{\mathbb{C}}$ may be written as a direct sum

$$V_0^{\mathbb{C}} = \text{Witt} \oplus \text{Witt}. \quad (1.3.18)$$

Here Witt is the **Witt** algebra, which abstractly is defined as the complex Lie algebra generated by $\{l_n\}_{n \in \mathbb{Z}}$ with the commutation relation $[l_n, l_m] = (n - m)l_{n+m}$. Less abstractly the Witt algebra is the algebra of derivations of the space of Laurent polynomials $\mathbb{C}[z, z^{-1}]$.

Restricting to just the Witt algebra briefly we can look at the subalgebra generated by $\{l_{-1}, l_0, l_1\}$. This subalgebra is isomorphic to $\mathfrak{sl}(2, \mathbb{C})$, which can be shown by taking the basis

$$e_{-1} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} = -f, \quad e_0 = \frac{1}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -\frac{1}{2}h, \quad \text{and} \quad e_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = e \quad (1.3.19)$$

for $\mathfrak{sl}(2, \mathbb{C})$, and we can then show, either from the matrices or from the standard commutation relations of $\{h, e, f\}$, that we have $[e_n, e_m] = (n - m)e_{n+m}$.

Considering both Witt algebras now, still restricting to $n = -1, 0, 1$, we can compute the following exponentials for $a, b, c \in \mathbb{C}$:

$$e^{-bl_{-1} - \bar{b}\bar{l}_{-1}}z = e^{b\partial_z}z = z + b \quad (1.3.20)$$

$$e^{-al_0 - \bar{a}\bar{l}_0}z = e^{az\partial_z}z = e^a z \quad (1.3.21)$$

$$e^{cl_1 + \bar{c}\bar{l}_1}z = e^{-cz^2\partial_z}z = \frac{z}{cz + 1} \quad (1.3.22)$$

The first two expressions here are standard using the power series, the third follows by summing the power series using the geometric series when $|zc| < 1$, and then analytically continuing to $\mathbb{C} \setminus \{-1/c\}$. We see that we generate the translations, rotations/dilations (since $a \in \mathbb{C}$ so $e^a \in \mathbb{C}^\times$) and special conformal transformations from this subalgebra of $V_0^\mathbb{C}$. By moving to the universal enveloping algebra, allowing us to consider products such as $l_0 l_1$, we get composites of these conformal transformations upon exponentiation, and thus we generate all Möbius transformations this way.

The fact that the algebra $W^\mathbb{C}$ is infinite-dimensional is special to two dimensions. In d dimensions with $d > 2$ it turns out that the conformal algebra is $\mathfrak{so}(d+2)$, which is finite dimensional (the dimension is actually $(d+2)(d+3)/2$).

Two

Quantum Field Theory

In this chapter we look at quantum field theory in general. We do not impose any conformal symmetry. We will develop QFT axiomatically, mostly focusing on some special functions called correlators. We do not interest ourselves with the physical interpretation of any of the objects discussed. We will introduce functional integration (also known as path integrals), which is just one of two main ways that a QFT can be constructed, the other being second quantisation, which we won't look at. The path integral construction is, at best, ill-defined, but we won't let this bother us and work with the path integrals as formal objects possessing the desired properties.

2.1 Correlation Functions

A quantum field theory is a quantum theory regarding fields, which are generally operator-valued functions of spacetime. We shall restrict our study to Euclidean QFTs in two-dimensions, where spacetime is \mathbb{R}^2 with the standard metric, or equivalently, \mathbb{C} with the standard metric.

We follow the standard physics practice of not really specifying what our fields are, or what they act on, and just say that we have some fields, which we'll denote with a calligraphic font for now, so \mathcal{O}, \mathcal{A} , and so on (later we'll see \mathcal{D} , which is not a field). We call these **local fields**, because they depend on the position of a single point.

While we don't specify what the fields are as functions of spacetime they do have a natural notion of addition and scaling, and so it's natural to assume that we have a vector space, W , of fields. We can take W to be a real or complex vector space. If we take W to be a real vector space then we can always complexify to get $W^{\mathbb{C}}$ and identifying W as a real subspace of this complex vector space. This allows us to always assume that W is a vector space over \mathbb{C} .

Definition 2.1.1 — Correlator Let W be the vector space of local fields, and let $\mathcal{O}_1, \dots, \mathcal{O}_n \in W$ be local fields. Define the set

$$\Delta_n := \{(z_1, \dots, z_n) \in (\mathbb{R}^2)^n \mid z_i = z_j \text{ for some } i \neq j\}, \quad (2.1.2)$$

this consists of all n -tuples of elements of \mathbb{R}^2 where two (or more) elements are equal. Then we define the **n -point correlation function** (or

correlator) as a function

$$\mathbb{R}^{2n} \setminus \Delta_n \rightarrow \mathbb{C} \quad (2.1.3)$$

defined by

$$(z_1, \dots, z_n) \mapsto \langle \mathcal{O}_1(z_1) \cdots \mathcal{O}_n(z_n) \rangle \in \mathbb{C}. \quad (2.1.4)$$

We will identify \mathbb{R}^2 with \mathbb{C} and then we instead have

$$\Delta_n = \{(z_1, \dots, z_n) \in \mathbb{C}^n \mid z_i = z_j \text{ for some } i \neq j\} \quad (2.1.5)$$

and the correlation functions are functions

$$\mathbb{C}^n \setminus \Delta_n \rightarrow \mathbb{C}. \quad (2.1.6)$$

Note that while we write $\mathcal{O}_i(z_i)$ we do not require that the \mathcal{O}_i are complex-analytic, and could just as well write $\mathcal{O}_i(z_i, \bar{z}_i)$, but we'll soon see that expressions in QFT already get quite large, so we don't want to add even more symbols.

We do not specify how the correlation function is evaluated, for now it's just notation. We'll look at one way to evaluate it later. For now we just demand that these functions have some desirable properties.

The first property is that the correlators are independent¹ of the order in which we write the $\mathcal{O}_i(z_i)$. That is, for $\pi \in S_n$ we should have

$$\langle \mathcal{O}_1(z_1) \cdots \mathcal{O}_n(z_n) \rangle = \langle \mathcal{O}_{\pi(1)}(z_{\pi(1)}) \cdots \mathcal{O}_{\pi(n)}(z_{\pi(n)}) \rangle. \quad (2.1.7)$$

Note that we have to permute both the fields and the evaluation points to get the equality².

The second property that we require is linearity in the fields. That is, if $\mathcal{O}_1 = \mathcal{A}_1 + \lambda \mathcal{B}_1$ for $\lambda \in \mathbb{C}$ and $\mathcal{A}_1, \mathcal{B}_1 \in W$ then we have

$$\begin{aligned} \langle \mathcal{O}_1(z_1) \mathcal{O}_2(z_2) \cdots \mathcal{O}_n(z_n) \rangle \\ = \langle \mathcal{A}_1(z_1) \mathcal{O}_2(z_2) \cdots \mathcal{O}_n(z_n) \rangle + \lambda \langle \mathcal{B}_1(z_1) \mathcal{O}_2(z_2) \cdots \mathcal{O}_n(z_n) \rangle. \end{aligned} \quad (2.1.8)$$

By requiring that the correlators are functions $\mathbb{C}^n \setminus \Delta_n \rightarrow \mathbb{C}$ we are requiring that they achieve finite values everywhere on $\mathbb{C}^n \setminus \Delta_n$. However, we can naturally extend the correlators to all of \mathbb{C}^n by considering them as functions $\mathbb{C}^n \rightarrow \mathbb{C}_\infty$ which are allowed to have singularities only at elements of Δ_n . The assumption that we have poles only at coincident points is called the **locality assumption**. Physically the idea is that the fields evaluated at different points can't interact with each other, but by evaluating the fields at the same point they can interact, and often do to produce divergent results. In a Euclidean QFT the result of this is that the local fields must be real analytic in a neighbourhood free of points in Δ_n . This means that they have a convergent power series on such neighbourhoods. It does not mean that they are *complex* analytic, and so there is no general requirement that the local fields are in any way holomorphic or antiholomorphic.

It is natural to require that as well as the fields \mathcal{O}_i being in W there are also fields $\partial^m \bar{\partial}^n \mathcal{O}_i$ in W which can reasonably be interpreted as the derivatives of the fields \mathcal{O}_i . We don't define the derivative of a field, we simply require that the following

¹Here we assume that the fields represent bosons (particles with integer spin). If instead they represent fermions (particles with half-integer spin) then exchanging two fermionic fields should pick up a minus sign. The result is that in the following formula we need a $\text{sgn } \pi$ factor for fermionic fields.

²The correlator may be thought of as a function of the product of the fields. When we think of the fields as operators as in second quantisation we do not assume that they commute. However, when in a correlator we are either doing path integrals, in which case the fields are taken to commute, or if we are doing second quantisation then for these expressions to make sense we must normal order the terms in the product, and this fixes an ordering independently of the order we write.

holds:

$$\langle \partial^m \bar{\partial}^n \mathcal{O}_1(z_1) \cdots \mathcal{O}_k(z_k) \rangle = \langle \partial_{z_1}^m \partial_{\bar{z}_1}^n \mathcal{O}_1(z_1) \cdots \mathcal{O}_k(z_k) \rangle \quad (2.1.9)$$

$$= \partial_{z_1}^m \partial_{\bar{z}_1}^n \langle \mathcal{O}_1(z_1) \cdots \mathcal{O}_k(z_k) \rangle. \quad (2.1.10)$$

Here the derivatives on the right are just normal derivatives of a complex function in z_1 and \bar{z}_1 . Note that this defines the derivative of any field, since by the symmetry of fields within the correlator we are always free to reorder the fields to get the one we're taking the derivative of at the front.

Notation 2.1.11 Derivatives are assumed to have higher precedence than multiplication. Thus, $\partial_x f g$ is always interpreted as $(\partial_x f)g$, rather than $\partial_x(fg)$.

2.2 Distributions

Now that we've got derivatives it's natural to consider integrals. Unfortunately, the singularities cause problems when it comes to defining integrals. The solution is to replace functions with distributions.

Definition 2.2.1 — Distribution Let X be a topological space, and $U \subseteq X$ open. Let $K(U)$ denote the set of all smooth functions $f : U \rightarrow \mathbb{C}$ with compact support (that is, the subset of U on which $f(x) \neq 0$ is compact). We call $K(U)$ the space of test functions, and $f \in K(U)$ a **test function**. There is a natural topology on $K(U)$ called the canonical LF-topology. A **distribution** is a continuous (with respect to the canonical LF-topology) linear functional $F \in K(U)^*$, that is $F : K(U) \rightarrow \mathbb{C}$.

We are interested in the case $X = \mathbb{C}^n$, and we assume that test functions are defined on all of \mathbb{C}^n , but we'll only ask that they have properties locally. Given a test function, $f \in K(\mathbb{C}^n)$, and a distribution, $\phi \in K(U)^*$, if f is locally integrable, meaning that

$$\int_U f(z) d^n z \quad (2.2.2)$$

exists for open $U \subseteq \mathbb{C}^n$, then there is a canonical way to evaluate the action of the distribution on f , namely

$$(\phi, f) = \int_U \phi(z) f(z) d^n z. \quad (2.2.3)$$

We assume that U is a region without boundary, or that at least f vanishes on the boundary of U , which allows us to use integration by parts to show that

$$(\phi, \partial_\mu f) = \int_U \phi(z) \partial_\mu f(z) d^n z = - \int_U (\partial_\mu \phi(z)) f(z) d^n z \quad (2.2.4)$$

where the boundary term,

$$\int_{\partial U} \phi(z) f(z) dA, \quad (2.2.5)$$

vanishes for one of the above reasons.

An alternative requirement to compact support for test functions is that test functions must decay to zero at infinity faster than any polynomial. This allows, for example, $z \mapsto e^{-|z|^2}$, to be a test function, despite being everywhere non-zero.

This is quite a formal way to do distributions, which we'll quickly abandon and start considering them to be functions-but-weird. The general rule is that whenever there's a distribution (usually a Dirac delta) it only really makes sense under an integral, even if we don't write the integral. So everywhere I write a Dirac delta just assume that it's under an appropriate integral.

2.3 Functional Integration

One way to produce correlation functions is through **functional integration**, the integration of distributions. Physicists call these **path integrals**, not to be confused with a contour integral, because when these integrals first arose in physics the domain of the functions was all possible paths between two points in space.

We won't define path integrals properly, mostly because there isn't a consistent definition that works in all scenarios. Instead, we'll motivate them by analogy with normal integrals, then we'll treat them as formal objects that we can learn to manipulate in ways that mirror the properties of normal integrals (maybe you're already seeing the similarity to how we treat correlators).

An ordinary integral looks like

$$\int_D f(x) d^n x \quad (2.3.1)$$

where \mathcal{M} is some manifold, $D \subseteq \mathcal{M}$ is a region of that manifold, $f : D \rightarrow \mathbb{C}$ is an integrable function, and $d^n x$ is an appropriate measure, where we've already assumed that D can be covered by some coordinates³ x .

A functional integral looks like

$$\int_C F[\phi] \mathcal{D}\phi \quad (2.3.2)$$

where C is some set of functions $\phi : \mathcal{M} \rightarrow \mathbb{C}$ (with \mathcal{M} a manifold), $F : C \rightarrow \mathbb{C}$ is some functional, and $\mathcal{D}\phi$ is an appropriate (usually ill-defined) measure.

³If this is not the case then we need to split D up into regions and be careful about where they overlap using partitions of unity.

Example 2.3.3 Consider the interval, $I = [0, 1] \subseteq \mathbb{R}$, and define

$$C(I) = \{\phi \mid I \rightarrow \mathbb{R} \mid \phi \text{ is continuous}\}. \quad (2.3.4)$$

We can approximate any such function by a piecewise linear function, $\phi_N \in P_N(I) \subset C(I)$, where $P_N(I)$ is the set of all continuous functions which are linear on the intervals $[i/N, (i+1)/N]$ for $i = 0, 1, \dots, N-1$. The space $C(I)$ is infinite dimensional, whereas the space $P_N(I)$ is finite dimensional. To specify an element of $P_N(I)$ we need only specify $N+1$ real numbers, typically the endpoints of each line segment, which are $\phi(0), \phi(1/N), \phi(2/N), \dots, \phi(1)$. Given a functional $F : C(I) \rightarrow \mathbb{R}$ we can define the functional integral as the "limit" of the integral over the $N+1$

dimensional space as we take $N \rightarrow \infty$:

$$\int_{C(I)} F[\phi] \mathcal{D}\phi = \lim_{N \rightarrow \infty} \int \cdots \int F[\phi_N] d\phi(0) \cdots d\phi(1). \quad (2.3.5)$$

Proving that this limit exists and is independent of the way we approximate the space $C(I)$ by finite-dimensional subspaces is hard, and has only been done rigorously in a few cases.

So, abandoning all hope of a rigorous definition of functional integrals we proceed to list some properties that we would like functional integrals to possess. The first is that, like normal integrals, functional integrals should be linear in the integrand. So, if F_1 and F_2 are functionals and $\lambda \in \mathbb{C}$ is a scalar then we require the following:

$$\int_C (F_1[\phi] + \lambda F_2[\phi]) \mathcal{D}\phi = \int_C F_1[\phi] \mathcal{D}\phi + \lambda \int_C F_2[\phi] \mathcal{D}\phi. \quad (2.3.6)$$

Now consider a normal integral over an unbounded region, or where the integrand vanishes on the boundary. Then by integration by parts we have

$$\int_D \frac{\partial f}{\partial x} d^n x = \int_D 1 \cdot \frac{\partial f}{\partial x} d^n x = - \int_D \frac{\partial 1}{\partial x} f(x) d^n x = 0. \quad (2.3.7)$$

We therefore demand the equivalent result for functional integrals, namely

$$\int_C \frac{\delta F}{\delta \phi(x)} \mathcal{D}\phi = 0. \quad (2.3.8)$$

Of course, we need to define the **functional derivative**, $\delta F / \delta \phi(X)$. This can be done formally as a limit, but again, we just look at the properties it has, including linearity,

$$\frac{\delta(F_1 + \lambda F_2)}{\delta \phi(x)} = \frac{\delta F_1}{\delta \phi(x)} + \lambda \frac{\delta F_2}{\delta \phi(x)}, \quad (2.3.9)$$

and the product rule,

$$\frac{\delta F_1 F_2}{\delta \phi(x)} = \frac{\delta F_1}{\delta \phi(x)} F_2[\phi] + F_1[\phi] \frac{\delta F_2}{\delta \phi(x)}. \quad (2.3.10)$$

Another important property is the functional version of

$$\frac{\partial x^i}{\partial x^j} = \delta^i_j, \quad (2.3.11)$$

which is

$$\frac{\delta \varphi(x)}{\delta \varphi(y)} = \delta(x - y). \quad (2.3.12)$$

The final property that we assume for functional integrals is invariance under a suitable change of variables. We'll go into more detail on how this is done later.

2.4 Path Integrals

In this section we will look at how QFT uses functional integrals (which I'll call path integrals in this context). We assume the simplest case, which is that there's only one field, ϕ , and it is a scalar field. This means that the corresponding particle will have zero spin. We won't worry about what this means, and in later sections we'll introduce other types of field which do have spin.

Looking at the properties that correlators and functional integrals possess we see that they match, and so it makes sense to define the correlators to be the result of evaluating specific path integrals. Specifically, we can define the n -point correlator by⁴

$$\langle \phi(z_1) \cdots \phi(z_n) \rangle = \frac{1}{Z} \int_C e^{-S[\phi]} \phi(z_1) \cdots \phi(z_n) \mathcal{D}\phi. \quad (2.4.1)$$

Here $\phi \in W$ is just a particular field, called the **fundamental field**. We will assume that most of our fields are built out of ϕ and its derivatives. In this Z is the **partition functional**, defined as

$$Z = \int_C e^{-S[\phi]} \mathcal{D}\phi. \quad (2.4.2)$$

It is just a normalisation factor. In these we have the **action functional**, S , which is usually defined via an integral

$$S[\phi] = \int_D \mathcal{L}(\phi, \partial_\mu \phi, t) dt d^n x \quad (2.4.3)$$

where we're working in $(n+1)$ -dimensional spacetime. The function \mathcal{L} is called the **Lagrangian** (or more properly, the Lagrangian density as we integrate it over time and space). Most QFT theories are specified by defining the Lagrangian. The simplest case is a the Lagrangian corresponding to a free scalar massless boson. Breaking this down:

- free means that there are no interactions between particles, which in QFT means that the Lagrangian doesn't have any terms higher than second order in ϕ ;
- scalar boson means that the particle has spin zero, and the corresponding field is just scalar valued (as opposed to a spinor or vector valued field, which we'll see later);
- massless means that the particle has no mass, which in QFT means that the Lagrangian has no ϕ^2 term.

The Lagrangian is then

$$\mathcal{L} = \frac{1}{8\pi} \partial_\mu \phi \partial^\mu \phi. \quad (2.4.4)$$

The single term here is called the kinetic term, and it controls how the particle moves through spacetime. The factor of $1/(8\pi)$ is included here because it will be

⁴We work in Euclidean spacetime. In Minkowski spacetime the exponent is $iS[\phi]$. Also, we're working in units where $\hbar = 1$, else we'd need a $1/\hbar$ in the exponent to give us a dimensionless quantity, with $\hbar = 1$ the action is dimensionless.

nice later. A more standard choice is just a factor of $1/2$. The action functional is then

$$S_{\text{fb}}[\phi] = \frac{1}{8\pi} \int_C \partial_\mu \phi \partial^\mu \phi \, d^2x. \quad (2.4.5)$$

Here we assume that we're working in two dimensional spacetime.

At this point we note that taking a product of composite fields, such as $\phi(z)\phi(z)$, is often going to lead to divergences in our correlators, because here we have two fields evaluated at the same point. The usual way we get around this is to consider such a product as short-hand for

$$\lim_{\varepsilon \rightarrow 0} \phi(z)\phi(z + \varepsilon), \quad (2.4.6)$$

or something similar which prevents these coincident evaluation points, and then only take the limit when it is safe to do so. Sometimes we still get infinities this way, but physics has a whole system for subtracting these off in a consistent way called **renormalisation**. In physics it's common to take $\phi(x)\phi(x + i\varepsilon)$, with an extra factor of i , because the poles we're usually wanting to avoid are on the real axis. and so this is often called the **$i\varepsilon$ prescription**.

It should also be noted that we can take many fundamental fields, for example, if we have n fundamental fields, ϕ_1, \dots, ϕ_n , with $\phi_i \in C_i$ then we may define the correlator of n fields as

$$\langle \phi_1(z_1) \cdots \phi_n(z_n) \rangle = \frac{1}{Z} \int_{C_1 \otimes \cdots \otimes C_n} e^{-S[\phi_1, \dots, \phi_n]} \phi_1(z_1) \cdots \phi_n(z_n) \mathcal{D}\phi_1 \cdots \mathcal{D}\phi_n$$

where

$$Z = \langle \phi_1(z_1) \cdots \phi_n(z_n) \rangle = \frac{1}{Z} \int_{C_1 \otimes \cdots \otimes C_n} e^{-S[\phi_1, \dots, \phi_n]} \mathcal{D}\phi_1 \cdots \mathcal{D}\phi_n. \quad (2.4.7)$$

2.4.1 Computing Correlators for the Free Boson

In this section we'll look at how, despite the fact that the path integrals aren't really defined, we can still use them to evaluate the correlators. We'll do this for the free massless scalar boson, which has the action functional⁵

$$S_{\text{fb}}[\phi] = \frac{1}{8\pi} \int_C d^2x \partial_\mu \phi \partial^\mu \phi. \quad (2.4.8)$$

2.4.1.1 Odd Correlators

One important property that this has is that the integrand (which is the Lagrangian) is invariant under the map

$$\phi(x) \mapsto -\phi(x). \quad (2.4.9)$$

In fact, we actually have a full-blown $U(1)$ symmetry

$$\phi(x) \mapsto e^{i\alpha} \phi(x), \quad \alpha \in [0, 2\pi), \quad (2.4.10)$$

but this is more than we need right now. This symmetry is important because it means that if the measure, $\mathcal{D}\phi$, is also invariant under this mapping then we have

$$\langle \phi(z_1) \cdots \phi(z_n) \rangle \mapsto (-1)^n \langle \phi(z_1) \cdots \phi(z_n) \rangle \quad (2.4.11)$$

⁵We start to use the physics convention where the measure is put before the integrand. This assumes that integration binds less tightly than multiplication but more tightly than addition, so $\int dx f g = \int dx (fg) = \int fg dx$ and $\int dx f + g = (\int dx f) + g = \int f dx + g = g + \int f dx$.

which tells us that if n is odd then the n -point correlator must vanish.

Physically, what this tells us is that there are no n -point interactions for n odd. An n -point interaction (in a general QFT) being some particles coming in, interacting, and some (possibly different) particles going out, such that the total number of particles going in plus the total number of particles going out is n . The reason that there cannot be n -point interactions (n odd) for the free boson is that there are no interactions, and so it is not possible to create or destroy particles, and as such the particles going into an interaction (in which nothing can happen, because we've got free particles) must be the same particles that go out, so if we have k particles coming in we have k going out, so we have a $2k$ -point interaction.

2.4.1.2 2-Point Correlators

Now we can use another property of functional integrals, the integral of a derivative vanishes, which tells us that

$$\int \frac{\delta}{\delta\phi(x)} (e^{-S_{\text{fb}}[\phi]} \phi(y)) \mathcal{D}\phi = 0. \quad (2.4.12)$$

Now, we can use the product rule, chain rule, and Dirac delta properties to take the derivative and we find that

$$\frac{\delta}{\delta\phi(x)} (e^{-S_{\text{fb}}[\phi]} \phi(y)) = \left(\frac{\delta}{\delta\phi(x)} e^{-S_{\text{fb}}[\phi]} \right) \phi(y) + e^{-S_{\text{fb}}[\phi]} \frac{\delta\phi(y)}{\delta\phi(x)} \quad (2.4.13)$$

$$= -\frac{\delta S[\phi]}{\delta\phi(x)} e^{-S_{\text{fb}}[\phi]} \phi(y) + e^{-S_{\text{fb}}[\phi]} \delta(x-y). \quad (2.4.14)$$

Of course, we're assuming that functional derivatives of exponentials work just as they do for normal derivatives, and this is indeed the case. Now all we have to do is compute the derivative of $S_{\text{fb}}[\phi]$:

$$\frac{\delta}{\delta\phi(x)} S_{\text{fb}}[\phi] = \frac{\delta}{\delta\phi(x)} \frac{1}{8\pi} \int d^2\tilde{x} \partial_{\mu}^{\tilde{x}} \phi(\tilde{x}) \partial_{\tilde{x}}^{\mu} \phi(\tilde{x}). \quad (2.4.15)$$

Here we're introducing \tilde{x} as a dummy variable distinct from x . The first step is to assume that the integral and the functional derivative commute, giving us

$$\frac{\delta}{\delta\phi(x)} S_{\text{fb}}[\phi] = \frac{1}{8\pi} \int d^2\tilde{x} \frac{\delta}{\delta\phi(x)} [\partial_{\mu}^{\tilde{x}} \phi(\tilde{x}) \partial_{\tilde{x}}^{\mu} \phi(\tilde{x})]. \quad (2.4.16)$$

Next we apply the chain rule to get

$$\frac{\delta}{\delta\phi(x)} S_{\text{fb}}[\phi] = \frac{1}{8\pi} \int d^2\tilde{x} \left[\frac{\delta\partial_{\mu}^{\tilde{x}} \phi(\tilde{x})}{\delta\phi(x)} \partial_{\tilde{x}}^{\mu} \phi(\tilde{x}) + \partial_{\mu}^{\tilde{x}} \phi(\tilde{x}) \frac{\delta\partial_{\tilde{x}}^{\mu} \phi(\tilde{x})}{\delta\phi(x)} \right]. \quad (2.4.17)$$

Now we assume that the partial derivatives commute with the functional derivative, so we have

$$\frac{\delta}{\delta\phi(x)} S_{\text{fb}}[\phi] = \frac{1}{8\pi} \int d^2\tilde{x} \left[\left(\partial_{\mu}^{\tilde{x}} \frac{\delta\phi(\tilde{x})}{\delta\phi(x)} \right) \partial_{\tilde{x}}^{\mu} \phi(\tilde{x}) + \partial_{\mu}^{\tilde{x}} \phi(\tilde{x}) \left(\partial_{\tilde{x}}^{\mu} \frac{\delta\phi(\tilde{x})}{\delta\phi(x)} \right) \right]. \quad (2.4.18)$$

Now, we the derivatives $\delta\phi(\tilde{x})/\delta\phi(x)$ give Dirac deltas, $\delta(\tilde{x}-x)$, and we can't really take the derivative of these. Instead, we use integration by parts to move the

derivative onto the other term at the cost of a minus sign, giving

$$\frac{\delta}{\delta\phi(x)} S_{\text{fb}}[\phi] = -\frac{1}{8\pi} \int d^2\tilde{x} \left[\frac{\delta\phi(\tilde{x})}{\delta\phi(x)} \partial_\mu^\tilde{x} \partial_x^\mu \phi(\tilde{x}) + \partial_\mu^\tilde{x} \partial_x^\mu \phi(\tilde{x}) \frac{\delta\phi(\tilde{x})}{\delta\phi(x)} \right] \quad (2.4.19)$$

$$= -\frac{1}{8\pi} \int d^2\tilde{x} \left[\delta(\tilde{x} - x) \partial_\mu^\tilde{x} \partial_x^\mu \phi(\tilde{x}) + \partial_\mu^\tilde{x} \partial_x^\mu \phi(\tilde{x}) \delta(\tilde{x} - x) \right] \\ = -\frac{1}{8\pi} (\partial_\mu^x \partial_x^\mu \phi(x) + \partial_x^\mu \partial_\mu^x \phi(x)) \quad (2.4.20)$$

$$= -\frac{1}{8\pi} (\partial_\mu^x \partial_x^\mu \phi(x) + \partial_x^\mu \partial_\mu^x \phi(x)) \quad (2.4.21)$$

$$= -\frac{1}{4\pi} \partial_\mu^x \partial_x^\mu \phi(x). \quad (2.4.22)$$

Using this work to compute the derivative in [Equation \(2.4.12\)](#) we get

$$0 = \int \mathcal{D}\phi \left(\frac{1}{4\pi} \phi(y) \partial_\mu^x \partial_x^\mu \phi(x) + \delta(x - y) \right) e^{-S_{\text{fb}}[\phi]}. \quad (2.4.23)$$

This tells us that

$$\int \mathcal{D}\phi \phi(y) \partial_\mu^x \partial_x^\mu \phi(x) e^{-S_{\text{fb}}[\phi]} = -4\pi \int \mathcal{D}\phi \delta(x - y) e^{-S_{\text{fb}}[\phi]}. \quad (2.4.24)$$

We can write this a bit more compactly by pulling the $\delta(x - y)$ out of the integral, which is valid as $\delta(x - y)$ doesn't depend on ϕ . Of course, a Dirac delta on its own doesn't make sense, so the following result only holds when we're within the integrand of some integral over x or y :

$$\int \mathcal{D}\phi \phi(y) \partial_\mu^x \partial_x^\mu \phi(x) e^{-S_{\text{fb}}[\phi]} = -4\pi \delta(x - y) \int \mathcal{D}\phi e^{-S_{\text{fb}}[\phi]} = -4\pi \delta(x - y) Z \quad (2.4.25)$$

where Z is the partition functional. If we once again assume that the integral commutes with the derivatives then we can write the left-hand side as

$$\int \mathcal{D}\phi \phi(y) \partial_\mu^x \partial_x^\mu \phi(x) e^{-S_{\text{fb}}[\phi]} \\ = \partial_\mu^x \partial_x^\mu \int \mathcal{D}\phi \phi(y) \phi(x) e^{-S_{\text{fb}}[\phi]} = \partial_\mu^x \partial_x^\mu \langle \phi(x) \phi(y) \rangle. \quad (2.4.26)$$

Thus, we're left with the equality (under an integral)

$$\partial_\mu^x \partial_x^\mu \langle \phi(x) \phi(y) \rangle = -4\pi \delta(x - y). \quad (2.4.27)$$

Writing $G(x, y) = \langle \phi(x) \phi(y) \rangle$ and writing out the derivatives in more detail this becomes

$$\left(\frac{\partial^2}{\partial(x^1)^2} + \frac{\partial^2}{\partial(x^2)^2} \right) G(x, y) = -4\pi \delta(x - y). \quad (2.4.28)$$

We therefore see that G is (up to some arbitrary choice of an overall constant factor) the Green's function for Laplace's equation

$$\nabla^2 f = 0 \quad (2.4.29)$$

in two dimensional Euclidean spacetime.

Remark 2.4.30 It is a general fact in QFT that the two-point correlators are Green's functions for the differential operator, D , appearing in the kinetic and mass terms when we write these two terms as $\phi D\phi$. One of the first examples of this in most QFT courses is given (in Minkowski space) by adding a mass term, $-m^2\phi^2/2$, to the Lagrangian. The resulting two-point correlator of the massive free scalar boson is the Green's function of the Klein–Gordon operator $\partial^2 + m^2$, where^a $\partial^2 = \partial^\mu \partial_\mu$ is the Laplacian operator in Minkowski space. The example we have here is the two-dimensional Euclidean version of this in the $m = 0$ case.

^aThis operator, called the d'Alembert operator or d'Alembertian, is also written as \square , and, confusingly, \square^2 (in analogy with ∇^2).

⁶well, someone's worked it out

The reason that this is useful is that we⁶ know what the Green's function for Laplace's equation in two-dimensions is, and so we know what the 2-point correlator is, it's

$$\langle \phi(x)\phi(y) \rangle = -\ln((x^1 - y^1)^2 + (x^2 - y^2)^2) = -2\ln(|x - y|)^2, \quad (2.4.31)$$

or, in complex coordinates $z_1 = x^1 + ix^2$ and $z_2 = y^1 + iy^2$,

$$\langle \phi(z_1)\phi(z_2) \rangle = -\ln(|z_1 - z_2|^2). \quad (2.4.32)$$

We can check, with a bit of work, that this is indeed a solution to Equation (2.4.27) by integrating Equation (2.4.27) using Stoke's theorem. This is easiest to do in complex coordinates. To do this we'll need to use

$$\partial_\mu \partial^\mu = 4\partial_z \partial_{\bar{z}} \quad (2.4.33)$$

which follows from the fact that in complex coordinates the inverse metric is

$$g^{\mu\nu} = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \quad (2.4.34)$$

so

$$\partial_\mu \partial^\mu = g^{\mu\nu} \partial_\mu \partial_\nu \quad (2.4.35)$$

$$= g^{zz} \partial_z \partial_z + g^{z\bar{z}} \partial_z \partial_{\bar{z}} + g^{\bar{z}z} \partial_{\bar{z}} \partial_z + g^{\bar{z}\bar{z}} \partial_{\bar{z}} \partial_{\bar{z}} \quad (2.4.36)$$

$$= 2\partial_z \partial_{\bar{z}} + 2\partial_{\bar{z}} \partial_z \quad (2.4.37)$$

$$= 4\partial_z \partial_{\bar{z}}. \quad (2.4.38)$$

Using this we can write Equation (2.4.27) as

$$\partial_{z_1} \partial_{\bar{z}_1} \langle \phi(z_1)\phi(z_2) \rangle = -\pi \delta(z_1 - z_2). \quad (2.4.39)$$

Stoke's theorem in complex coordinates is

$$\iint_D \partial_{\bar{z}} F(z, \bar{z}) d^2z = -\frac{i}{2} \oint_{\partial D} F(z, \bar{z}) dz. \quad (2.4.40)$$

The contour integral is along the boundary of D and oriented anticlockwise. Taking a region D containing z_2 we can integrate the left hand side of Equation (2.4.39) to get

$$\iint_D \partial_{\bar{z}_1} \partial_{z_1} \langle \phi(z_1)\phi(z_2) \rangle dz_1 = -\frac{i}{2} \oint_{\partial D} \partial_{z_1} \langle \phi(z_1)\phi(z_2) \rangle dz_1 \quad (2.4.41)$$

and on the right we can use the Dirac delta to simply get

$$-\pi \iint_D \delta(z_1 - z_2) d^2 z_1 = -\pi. \quad (2.4.42)$$

Thus, we have

$$\frac{i}{2} \oint_{\partial D} \partial_{z_1} \langle \phi(z_1) \phi(z_2) \rangle dz_1 = \pi. \quad (2.4.43)$$

Now we can substitute in Equation (2.4.32) and we have

$$\partial_{z_1} \langle \phi(z_1) \phi(z_2) \rangle = -\partial_{z_1} \ln(|z_1 - z_2|^2) = -\frac{2}{z_1 - z_2} \quad (2.4.44)$$

and so we have

$$-i \oint_{\partial D} \frac{1}{z_1 - z_2} dz_1 = \pi. \quad (2.4.45)$$

This is true by Cauchy's integral formula,

$$f(a) = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(z_1)}{z_1 - z_2} dz_1 \quad (2.4.46)$$

for $f(z) = 1$.

2.4.1.3 4-Point Correlators

We can compute 4-point correlation functions now, but we won't because the 2-point correlator took long enough. We start again with the integral of a derivative vanishing:

$$\int \frac{\delta}{\delta \phi(z_1)} (e^{-S_{\text{fb}}[\phi]} \phi(z_2) \phi(z_3) \phi(z_4)) \mathcal{D}\phi = 0. \quad (2.4.47)$$

We then compute the derivative and go through the same steps as for the 2-point correlator and we find that

$$\partial_\mu^{z_1} \partial_{z_1}^\mu \langle \phi(z_1) \phi(z_2) \phi(z_3) \phi(z_4) \rangle = -\pi \delta(z_1 - z_2) \langle \phi(z_3) \phi(z_4) \rangle \quad (2.4.48)$$

$$- \pi \delta(z_1 - z_3) \langle \phi(z_2) \phi(z_4) \rangle \quad (2.4.49)$$

$$- \pi \delta(z_1 - z_4) \langle \phi(z_2) \phi(z_3) \rangle. \quad (2.4.50)$$

This, and equations like it, are called either

- **Schwinger–Dyson equations;** or
- **quantum equations of motion.**

The second comes from identifying that the classical equations of motion, as one would find from the Euler–Lagrange equations, which come from requiring that $\delta S[\phi] = 0$, are in this case just the usual Laplace equation, $\partial_z \partial_{\bar{z}} \phi = 0$. The quantum equations of motion then agree with this (replacing the field with the correlator) at all points, except where we have two coincident field evaluations in the correlator, where we instead get Dirac deltas. These are called **contact terms**, and

the physical intuition is that these infinities arise because we're trying to put two particles in the same place, allowing them to interfere with each other.

We can substitute in the known 2-point correlators and we get

$$\partial_\mu^{z_1} \partial_{z_1}^\mu \langle \phi(z_1) \phi(z_2) \phi(z_3) \phi(z_4) \rangle = \ln(|z_1 - z_2|^2) \ln(|z_3 - z_4|^2) \quad (2.4.51)$$

$$+ \ln(|z_1 - z_3|^2) \ln(|z_2 - z_4|^2) \quad (2.4.52)$$

$$+ \ln(|z_1 - z_4|^2) \ln(|z_2 - z_3|^2). \quad (2.4.53)$$

In general if we choose a more complicated theory, in particular one with interactions, the Schwinger–Dyson equations will not be solvable.

2.4.1.4 Wick's Theorem

This process generalises to n -point correlators for any even n , and we find that we have the following general result.

Theorem 2.4.54 — Wick's Theorem For A Free Massless Scalar Boson. The n -point correlator of a free massless scalar boson field is either 0, if n is odd, or

$$\langle \phi(z_1) \cdots \phi(z_{2k}) \rangle = \sum_{\text{pairs } (i_\ell, j_\ell)} (-1)^k \ln(|z_{i_1} - z_{j_1}|) \ln(|z_{i_2} - z_{j_2}|) \cdots \ln(|z_{i_k} - z_{j_k}|)$$

if $n = 2k$ is even. The sum here is taken over ways of partitioning the indices $\{1, \dots, 2k\}$ into pairs of indices with no repeated indices. We consider two partitions into pairs to be the same if they differ only by the order of the pairs or the order of the two indices within a pair.

To demonstrate what we mean by the pairs thing in the $n = 2$ case we simply have a single pair, $(1, 2)$, and in the $n = 4$ case the sum has three terms, corresponding to the pairs $(1, 2)(3, 4)$, $(1, 3)(2, 4)$, and $(1, 4)(2, 3)$. We do not consider $(1, 2)(4, 3)$, because this is the same as $(1, 2)(3, 4)$, and we do not consider $(3, 4)(1, 2)$, because this is the same as $(1, 2)(3, 4)$.

Remark 2.4.55 — Wick Contractions and Generalisations Noting that the logarithms here are just 2-point correlators and writing

$$\overline{\phi(z_1) \phi(z_2)} = \langle \phi(z_1) \phi(z_2) \rangle, \quad (2.4.56)$$

which we call a **Wick contraction**, it is common to write these results using this notation, for example if $n = 4$ we write

$$\langle \phi(z_1) \cdots \phi(z_4) \rangle = \overbrace{\phi(z_1) \phi(z_2)} \overbrace{\phi(z_3) \phi(z_4)} \quad (2.4.57)$$

$$+ \overbrace{\phi(z_1) \phi(z_2) \phi(z_3)} \phi(z_4) \quad (2.4.58)$$

$$+ \overbrace{\phi(z_1) \phi(z_2) \phi(z_4)} \phi(z_3) \quad (2.4.59)$$

It should be noted that in the full generalisation of Wick's theorem for a non-free field there is no reason that the odd correlators should vanish. Instead, the odd correlators typically have some product of contracted fields,

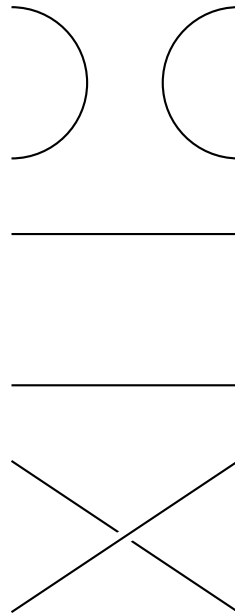


Figure 2.1: The three ways to have two particles enter on the left and two particles leave on the right (to first order, we can have weird loops at higher order, but higher orders are hard and (usually) less relevant than first order).

This corresponds to the three different ways to contract four fields.

and some non-contracted fields. For a fully general result of $n = 2k$ or $n = 2k + 1$ interacting fields we have to sum over all terms with 2ℓ contracted fields where $\ell = 1, \dots, k$. Because this leaves us with other fields we also have to normal order the non-contracted fields. This also assumes that the original fields are time ordered.

In Feynman diagrams the contracted fields correspond to propagators between the points of evaluation. Uncontracted fields correspond to vertices. So in the free theory since all fields are contracted there are no vertices, which makes sense because the vertices are interactions, and all of the particles just propagate from start to end. The reason that we end up with multiple terms is that the particles are indistinguishable, so in the $n = 4$ case if we assume that 1 and 2 refer to the incoming particles and 3 and 4 the outgoing particles we can't tell which of the three following diagrams in Figure 2.1 actually corresponds to what is actually happening.

(Obligatory “in QFT Feynman diagrams aren't real, the particles aren't real, and they aren't really following nice paths on the page, instead the Feynman diagrams correspond to integrals in a big series expansion for the scattering matrix and shouldn't be interpreted physically”. Despite this everyone does think of the Feynman diagrams as physical, after all, the fields aren't really any more “real” than the particles are.)

Despite having spent all of this time on functional integrals they turn out not to be that useful in CFT. Instead we will work with the correlation functions di-

rectly. We will impose upon them conditions that arise from demanding a conformal symmetry. This, plus a few more assumptions on which local fields appear, allows us to solve these the Schwinger–Dyson equations to obtain correlation functions.

2.5 Continuous Symmetries

A continuous symmetry in QFT is given by taking W , the space of fields, to be a representation of some Lie group such that the action functional, S , is invariant under this action. Since we have a Lie group the natural thing to do is consider the Lie algebra, and look at infinitesimal transformations. Suppose that we have m fundamental fields, ϕ_1, \dots, ϕ_m , upon which our action functional depends. Then an infinitesimal transformation of these fields is a map

$$\phi_i(x) \mapsto \phi_i(x) + \omega^a \delta_a \phi_i(x). \quad (2.5.1)$$

Here $\omega^a \in \mathbb{R}$ are some parameters (assuming a real Lie group, \mathbb{C} for a complex Lie group), and we shall typically assume these parameters to be “small”, in the way that physicists do, meaning that anything of second order in ω^a is negligible. The $\delta_a \phi_i(x)$ are some other fields which may depend on i or x . If $\delta_a \phi_i(x)$ is independent of x and i , so it’s just a constant, then we call this a **global symmetry**. Note that this is a different use of “global” to that in global conformal symmetry, which simply means that it’s defined everywhere. If $\delta_a \phi_i(x)$ depends on x then it is called a **local symmetry**. Finally, if $\delta_a \phi_i(x)$ depends on i then it is called an internal symmetry.

The change in the action functional under an arbitrary infinitesimal transformation is

$$\delta S = \int d^2x \omega^a \delta_a \phi_i(x) \frac{\delta S}{\delta \phi_i(x)}. \quad (2.5.2)$$

The infinitesimal transformation is a symmetry (by which we mean it exponentiates to a symmetry) if and only if $\delta S = 0$.

Example 2.5.3 Consider the free massless scalar boson with action functional

$$S_{\text{fb}}[\phi] = \frac{1}{8\pi} \int d^2x \partial_\mu \phi \partial^\mu \phi. \quad (2.5.4)$$

This is invariant under a translation

$$\phi(x) \mapsto \phi(x + \omega) \quad (2.5.5)$$

for some constant $\omega \in \mathbb{R}^2$ (working in 2-dimensional Euclidean space). The infinitesimal version of this is

$$\phi(x) \mapsto \phi(x) + \omega^\mu \partial_\mu \phi(x). \quad (2.5.6)$$

This follows by Taylor expanding,

$$\phi(x + \omega) = \phi(x) + \omega^\mu \partial_\mu \phi(x) + \mathcal{O}(\omega^2), \quad (2.5.7)$$

and truncating after the first order term. To see that this is a symmetry we can compute the variation in the action:

$$\delta S_{\text{fb}} = \frac{1}{8\pi} \int d^2x \omega^\mu \partial_\mu \phi(x) \frac{\delta S}{\delta \phi(x)} \quad (2.5.8)$$

$$= -\frac{1}{4\pi} \int d^2x \omega^\mu \partial_\mu \phi(x) \partial_\nu \partial^\nu \phi(x) \quad (2.5.9)$$

$$= \frac{1}{4\pi} \int d^2x \omega^\mu \partial_\mu \partial_\nu \phi(x) \partial^\nu \phi(x) \quad (2.5.10)$$

$$= \frac{1}{8\pi} \omega^\mu \int d^2x \partial_\mu (\partial_\nu \phi(x) \partial^\nu \phi(x)) \quad (2.5.11)$$

$$= 0. \quad (2.5.12)$$

The second line follows by similar calculations to the 2-point correlator after Equation (2.4.12). The next line is integration by parts to move one of the derivatives across. The line after that is just the product rule,

$$\partial_\mu (\partial_\nu \phi \partial^\nu \phi) = \partial_\mu \partial_\nu \phi \partial^\nu \phi + \partial_\nu \phi \partial_\mu \partial^\nu \phi \quad (2.5.13)$$

$$= \partial_\mu \partial_\nu \phi \partial^\nu \phi + \partial^\nu \phi \partial_\mu \partial_\nu \phi \quad (2.5.14)$$

$$= 2\partial_\mu \partial_\nu \phi \partial^\nu \phi, \quad (2.5.15)$$

and finally the whole expression is zero because it's the integral of a total derivative, and this vanishes by integration by parts with a constant function and vanishing on the boundary.

Example 2.5.16 Consider the action functional for three free massless scalar boson fields

$$S_{\text{fb}}^{(3)}[\phi_1, \phi_2, \phi_3] = \frac{1}{8\pi} \int d^2x \partial_\mu \phi_a \partial^\mu \phi_a. \quad (2.5.17)$$

We sum over both $\mu = 1, 2$ and $a = 1, 2, 3$ here. This is invariant under the mapping

$$\phi_a(x) \mapsto M_{ab} \phi_b(x) \quad (2.5.18)$$

for $M_{ab} \in O(3)$. To see this write the three fields as a vector

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} \quad (2.5.19)$$

and then the action functional becomes

$$S_{\text{fb}}^{(3)}[\phi_1, \phi_2, \phi_3] = \frac{1}{8\pi} \int d^2x (\partial_\mu \phi) \cdot (\partial^\mu \phi) \quad (2.5.20)$$

where the dot product is exactly what you'd expect:

$$\Phi \cdot \Phi = \Phi_1^2 + \Phi_2^2 + \Phi_3^2 \quad \text{for} \quad \Phi = (\Phi_1, \Phi_2, \Phi_3). \quad (2.5.21)$$

The action functional is clearly invariant under $\phi \mapsto M\phi$ since by definition $M\Phi \cdot M\Phi = \Phi \cdot \Phi$ for $M \in O(3)$, and this holds for the derivative terms also. It is then natural to think of M as rotating the fields into each other, since after acting with M the first component of ϕ will generally depend on ϕ_1 , ϕ_2 , and ϕ_3 . In fact, it's common to think of all symmetries of the form $\phi \mapsto U\phi$ for U in some matrix Lie group as “rotating” the fields into each other.

Note that this only works if M is constant (i.e., a global symmetry), if M depends on x then the derivative terms get messed up. There is a canonical way to fix this, we replace the derivatives ∂_μ with “gauge covariant derivatives” $D_\mu = \partial_\mu - \alpha A_\mu$ for some constant^a α and some vector field A_μ , which is chosen to also transform in such a way that the extra terms arising from the derivatives of $M(x)$ perfectly cancel with the extra terms arising from A_μ 's transformation. In this case this requires that A_μ transforms as $A_\mu(x) \mapsto A_\mu(x) + \alpha^{-1} \partial_\mu M(x)$.

The infinitesimal version of this transformation is

$$\phi_a(x) \mapsto \phi_a(x) + \omega^b \varepsilon_{abc} \phi_c(x) \quad (2.5.22)$$

where ε_{abc} is the Levi-Civita totally antisymmetric 3-nr defined by $\varepsilon_{123} = 1$. This comes from the fact that ε_{abc} are the structure constants of the $\mathfrak{o}(3)$ Lie algebra, which consists of antisymmetric 3×3 real matrices. Fixing b we can think of ε_{abc} as a 3×3 matrix with components $A_{ac} = \varepsilon_{abc}$, and these matrices form a basis for $\mathfrak{o}(3)$, so $\omega^b \varepsilon_{abc}$ is just an arbitrary element of $\mathfrak{o}(3)$ expressed in this basis.

We can see that this is a symmetry by computing $\delta S_{\text{fb}}^{(3)}$:

$$\delta S_{\text{fb}}^{(3)} = -\frac{1}{4\pi} \int d^2x \varepsilon_{abc} \omega^b \phi_c(x) \partial_\nu \partial^\nu \phi_a(x) \quad (2.5.23)$$

$$= \frac{1}{4\pi} \int d^2x \omega_{abc} \omega^b \partial_\nu \phi_c(x) \partial^\nu \phi_a(x) \quad (2.5.24)$$

$$= 0. \quad (2.5.25)$$

Here the first equality follows from the same calculations that we used to compute the 2-point correlators following Equation (2.4.12). The second equality is integration by parts to move the derivative across at the cost of a sign. The third equality is because ε_{abc} is antisymmetric in a and c , and $\partial_\nu \phi_c(x) \partial^\nu \phi_a(x)$ is symmetric in a and c .

^aconstant with respect to spacetime, it turns out that α is generally not constant with respect to the energy of the interaction, this is known as running of the gauge coupling, and is one of the weird things about QFT. It leads to cool sounding things like “asymptotic freedom” in QCD.

Appendices

A

Differential Geometry

A.1 Tangent Space

Let \mathcal{M} be a d -dimensional manifold. The tangent space at $p \in \mathcal{M}$ is a d -dimensional vector space $T_p\mathcal{M}$. One definition of this is the vector space of derivations at p , where a derivation is a linear map $D: C^\infty(\mathcal{M}) \rightarrow \mathbb{R}$ satisfying

$$D(fg) = D(f)g(x) + f(x)D(g). \quad (\text{A.1.1})$$

Clearly derivatives are derivations, this is just the product rule, and in fact given a coordinate chart (U, x) with $p \in U$ and $x = (x^1, \dots, x^d)$ we have a basis for $T_p\mathcal{M}$ given by

$$\left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^d} \Big|_p \right\}. \quad (\text{A.1.2})$$

Once we have tangent spaces it makes sense to consider the collection of all tangent vectors at any point $p \in \mathcal{M}$. This gives us the tangent bundle

$$TM = \bigsqcup_{p \in \mathcal{M}} T_p\mathcal{M}. \quad (\text{A.1.3})$$

This is a bundle since we have the natural projection $\pi: TM \rightarrow \mathcal{M}$ sending a tangent vector $v \in T_p\mathcal{M}$ to the point $p \in \mathcal{M}$.

A.2 Pushforward and Pullback

Let $\varphi: \mathcal{M} \rightarrow \mathcal{N}$ be a smooth map between manifolds. The **pushforward**, also called the **differential**, of φ at $p \in \mathcal{M}$ is the linear map

$$d\varphi_p: T_p\mathcal{M} \rightarrow T_{\varphi(p)}\mathcal{N} \quad (\text{A.2.1})$$

defined to act on a derivation, $X: C^\infty(\mathcal{M}) \rightarrow \mathbb{R}$, by sending it to the derivation $d\varphi_p(X): C^\infty(\mathcal{N}) \rightarrow \mathbb{R}$ defined to act on $f \in C^\infty(\mathcal{N})$ by

$$d\varphi_p(X)(f) = X(f \circ \varphi). \quad (\text{A.2.2})$$

That is, $d\varphi_p$ is nothing but precomposition with φ followed by evaluation.

Fix charts (U, x) and (V, y) for neighbourhoods of $p \in \mathcal{M}$ and $\varphi(p) \in \mathcal{N}$. Then $T_p\mathcal{M}$ and $T_{\varphi(p)}\mathcal{N}$ have bases $\{\partial/\partial x^i|_p\}$ and $\{\partial/\partial y^i|_{\varphi(p)}\}$. In these bases $d\varphi_p$ may be expressed as a matrix

$$(d\varphi_p)^i_j = \frac{\partial \varphi^i}{\partial x^j} \quad (\text{A.2.3})$$

where φ^j is such that $y^j = \varphi(x^j)$.

The **pullback** of φ is the map $\varphi^* : C^\infty(\mathcal{N}) \rightarrow C^\infty(\mathcal{M})$ defined by $(\varphi^* f)(x) = f(\varphi(x))$. That is, φ^* is precomposition with φ . We can also define the pullback of a k -form, ω , as

$$(\varphi^* \omega)_p(X_1, \dots, X_k) = \omega_{\varphi(p)}(d\varphi_p(X_1), \dots, d\varphi_p(X_k)). \quad (\text{A.2.4})$$

This will be particularly important for a 2-form, g , where we have

$$(\varphi^* g)_p(X_1, X_2) = g_{\varphi(p)}(d\varphi_p(X_1), d\varphi_p(X_2)). \quad (\text{A.2.5})$$

A.3 Riemannian Manifolds

A **Riemannian manifold**, (\mathcal{M}, g) , is a manifold, \mathcal{M} , equipped with a **Riemannian metric**, g , which assigns to each tangent space, $T_p\mathcal{M}$, a positive-definite inner product

$$g_p : T_p\mathcal{M} \times T_p\mathcal{M} \rightarrow \mathbb{R}, \quad (\text{A.3.1})$$

such that the component functions, $g_{ij} : U \rightarrow \mathbb{R}$, are smooth on any chart (U, x) . These components are defined for a basis $\{e_i\}$ of $T_p\mathcal{M}$ by

$$g_{ij} = g_p(e_i, e_j). \quad (\text{A.3.2})$$

These are such that

$$g = \sum_{i,j} g_{ij} dx^i dx^j \quad (\text{A.3.3})$$

where dx^i is the dual basis to $\{e_i\}$, defined by $dx^i(e_j) = \delta^i_j$.