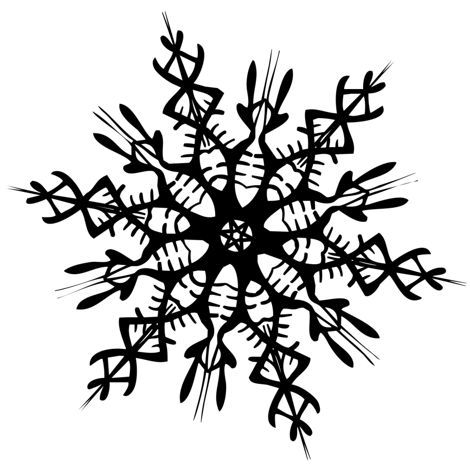
rijksuniversiteit groningen

faculteit wiskunde en natuurwetenschappen

Non-linear realisations of supersymmetry using constrained superfields

July 7, 2016



Bachelor Thesis Mathematics & Physics

Author: W.H. Elbers

Supervisors: D. Roest

A.V. Kiselev

Abstract

A non-linear group realisation is a homomorphism from a group to a group of transformations of a topological space. Non-linear realisations of symmetry groups are important in physics, where they occur in quantum field theories with spontaneous symmetry breaking. For this reason, one studies the non-linear Volkov-Akulov realisation of supersymmetry in theories with supersymmetry breaking. One popular technique for obtaining the Volkov-Akulov realisation makes use of a linear realisation that is made to satisfy a non-linear constraint. We investigate the validity of this approach. In order to tackle this problem, we prove a number of results concerning non-linear realisations in general. In particular, we show that imposing an algebraic constraint on a (non-)linear realisation yields another non-linear realisation, provided that the constraint itself is invariant under the transformation group. Subsequently, we formally derive linear and non-linear realisations of the superPoincaré group using the notion of smooth superfunctions in superspace. Finally, for a general non-linear sigma model with global supersymmetry and n chiral superfields, we derive conditions on the Kähler and superpotential that guarantee that the goldstino superfield Φ satisfies a given nilpotency condition $\Phi^k = 0$ for $k = 2, 3$ in the limit of infinite UV cut-off scale Λ and at energies far below the mass of the sgoldstino.

Contents

1

Introduction

1.1 Introduction

Today, two of the most important theories describing physics beyond the Standard Model are inflation and supersymmetry. The theory of inflation was proposed in an attempt to solve a number of problems caused by the standard Big-Bang model of the Universe. It postulates that during a short period after the Big Bang, spacetime underwent a rapidly accelerating expansion. An important theoretical challenge is to uncover the particle physics mechanism behind inflation. Models of inflation usually employ a scalar field, called the inflaton field, whose potential energy drives the expansion of spacetime. However, no known particle fits the bill, except perhaps the Higgs boson. Meanwhile, the theory of supersymmetry postulates the existence of an additional symmetry of Nature, beyond the Poincaré and gauge symmetries of the Standard Model. Supersymmetry, too, solves a number of problems in contemporary physics, which we consider in detail in chapter 5. One of the predictions of supersymmetric theories is the existence of a large number of undiscovered "superparticles". A tantalising possibility emerges: that one of the superparticles is the inflaton field that drove the expansion of the early Universe. And so theories of supersymmetric inflation were born.

This thesis is mostly concerned with supersymmetry breaking. We know that supersymmetry must be broken, because theories with unbroken supersymmetry predict that superparticles have the same mass as known Standard Model particles, which directly contradicts observations. If supersymmetry is broken, the masses of superparticles are essentially unconstrained. Fortunately, there is reason to believe that the lightest superparticles have masses of around 1 TeV, which is a feasible scale for detection at the Large Hadron Collider in Geneva. See chapter 5 for a detailed discussion. Describing exactly how supersymmetry is broken remains an important challenge in theoretical particle physics. The focus in this thesis is on non-linear realisations of supersymmetry. The reason is that, in quantum field theories, spontaneous symmetry breaking is associated with non-linearly transforming fields. An important goal is to describe in detail how such non-linear realisations arise and to make such notions as "non-linear realisation" more precise.

One technique used to describe theories with non-linear supersymmetry makes use of so-called constrained superfields. We will review this technique in detail and we shall derive general conditions which guarantee that the use of the technique is valid. In the end, it turns out that constrained superfields are incredibly useful in theories of supersymmetric inflation. We see this as additional motivation to study non-linear supersymmetry.

1.2 Looking ahead

Instead of making an artificial distinction between the mathematical and the physical parts of the thesis, I opted for a more logical structure. Chapter 2 serves as an accessible introduction to the topic of spontaneous symmetry breaking, whereas chapter 3 provides the mathematical underpinnings. These first two chapters apply mostly to non-supersymmetric theories, although some of the findings (most notably in section 3.4) carry over directly to supersymmetry breaking. In chapter 4, some additional mathematical concepts are introduced, allowing us to formally describe representations of supersymmetry. Then, chapter 5 actually introduces the physics of supersymmetry. Next, non-linear realisations of supersymmetry are dealt with in chapter 6. Here, we also consider the use of constrained superfields. Finally, in the last chapter, we offer some concluding remarks.

Those readers who are only interested in the physical aspects of non-linear supersymmetry may skip chapters 3 and 4, although it may be worth reading section 3.1, which is intended as an accessible introduction to the rest of the chapter.

1.3 Conventions

Our conventions are mostly the same as [8], but an important difference is the replacement of a non-standard minus sign in the definition of the auxiliary field F of a chiral superfield $\Phi = (\phi, \psi, F)$. We use the Minkowski metric $\eta = \mathrm{diag}(1, -1, -1, -1)$ and always sum over repeated indices. Especially important are the Weyl spinors, which are two-component complex vectors

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix},$$

where ψ_1 and ψ_2 are anticommuting elements of a Grassmann algebra $\mathcal{G}_L$ (see chapter 4). We use both left-handed (ψ_α) and right-handed ($\bar{\psi}_{\dot{\alpha}}$) representations, where left-handed indices $\alpha = 1, 2$ are always undotted and right-handed indices $\dot{\alpha} = 1, 2$ are dotted. The spinors transform under a matrix $\mathcal{M} \in \mathrm{SL}(2, \mathbb{C})$, according to

$$\begin{aligned} \psi_\alpha &\to \psi'_\alpha = \mathcal{M}_\alpha{}^\beta \psi_\beta, \\ \bar{\psi}_{\dot{\alpha}} &\to \bar{\psi}'_{\dot{\alpha}} = \mathcal{M}^{*}_{\dot{\alpha}}{}^{\dot{\beta}} \bar{\psi}_{\dot{\beta}}. \end{aligned}$$

Here, the special linear group $\mathrm{SL}(2, \mathbb{C})$ is the group of 2×2-matrices with determinant 1. Usually, we consider representations of subgroups of $\mathrm{SL}(2, \mathbb{C})$, such as $\mathrm{SU}(2, \mathbb{C})$, but this has no bearing on our conventions here. Spinor indices can be raised and lowered with matrices

$$\epsilon_{\alpha\beta} = \epsilon_{\dot{\alpha}\dot{\beta}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \epsilon^{\alpha\beta} = \epsilon^{\dot{\alpha}\dot{\beta}} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For example, $\psi^\alpha = \epsilon^{\alpha\beta}\psi_\beta$. We define scalar products as

$$\begin{aligned} \psi\chi &\equiv \psi^\alpha \chi_\alpha = -\psi_\alpha \chi^\alpha = \chi^\alpha \psi_\alpha = \chi\psi, \\ \bar{\psi}\bar{\chi} &\equiv \bar{\psi}_{\dot{\alpha}} \bar{\chi}^{\dot{\alpha}} = -\bar{\psi}^{\dot{\alpha}} \bar{\chi}_{\dot{\alpha}} = \bar{\chi}_{\dot{\alpha}} \bar{\psi}^{\dot{\alpha}} = \bar{\chi}\bar{\psi}, \end{aligned}$$

where the order is important, because the spinor components anticommute. Under Hermitian conjugation, one finds

$$(\psi\chi)^\dagger = (\psi^\alpha \chi_\alpha)^\dagger = \chi^\dagger_\alpha \psi^{\alpha\dagger} = \bar{\chi}_{\dot{\alpha}} \bar{\psi}^{\dot{\alpha}} = \bar{\chi}\bar{\psi}.$$

Also important are the σ-matrices $\sigma^\mu = (\sigma^0, \sigma^1, \sigma^2, \sigma^3)$, where $\sigma^0 = I$ is the 2×2 identity matrix and σ^i, for $i = 1, 2, 3$ are the Pauli matrices. We use the convention

$$\psi\sigma^\mu\bar{\chi} \equiv \psi^\alpha \sigma^\mu_{\alpha\dot{\alpha}} \bar{\chi}^{\dot{\alpha}}.$$

Similarly, we use $(\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} = \epsilon^{\dot{\alpha}\dot{\beta}}\epsilon^{\alpha\beta}\sigma^\mu_{\beta\dot{\beta}} = (\sigma_0, \sigma_i)$ and write $\bar{\chi}\bar{\sigma}^\mu\psi \equiv \bar{\chi}_{\dot{\alpha}} \left(\bar{\sigma}^\mu\right)^{\dot{\alpha}\alpha} \psi_\alpha$. Finally, we also make use of the matrices

$$\begin{aligned} \left(\sigma^{\mu\nu}\right)_\alpha{}^\beta &= \tfrac{1}{4} \left(\sigma^\mu_{\alpha\dot{\gamma}} \left(\bar{\sigma}^\nu\right)^{\dot{\gamma}\beta} - \sigma^\nu_{\alpha\dot{\gamma}} \left(\bar{\sigma}^\mu\right)^{\dot{\gamma}\beta} \right), \\ \left(\bar{\sigma}^{\mu\nu}\right)^{\dot{\alpha}}{}_{\dot{\beta}} &= \tfrac{1}{4} \left(\left(\bar{\sigma}^\mu\right)^{\dot{\alpha}\gamma} \sigma^\nu_{\gamma\dot{\beta}} - \left(\bar{\sigma}^\nu\right)^{\dot{\alpha}\gamma} \sigma^\mu_{\gamma\dot{\beta}} \right). \end{aligned}$$

2

Spontaneous Symmetry Breaking

2.1 Introduction

At its heart, this thesis is about spontaneous symmetry breaking. We speak of spontaneous symmetry breaking when a physical system settles in a ground state that breaks the symmetry of the underlying laws. In contrast, explicit symmetry breaking occurs when an external force breaks the underlying symmetry. This means that the laws were not really symmetric to begin with. In quantum field theory, the latter case is modelled by explicitly adding a symmetry breaking term to the Lagrangian.

The subject of spontaneous symmetry breaking is usually introduced by means of a few tried and tested examples, such as that of a rod under pressure or a ferromagnet. We shall deal with the rod later, but let us begin with the example of freezing water. In its liquid phase, water is a rotationally invariant system. However, as the temperature drops, water molecules settle in a crystalline structure dictated by hydrogen bonds. As each molecule is fixed at its location in the crystal, an arbitrary rotation changes the properties of the system; the continuous rotational symmetry has been broken. Nevertheless, ice crystals usually retain a sixfold rotational symmetry at a microscopic level[1], as depicted in figure 2.1. This is common in spontaneous symmetry breaking. The full group of symmetry transformations is broken down to a smaller subgroup. In this case, the rotation group SO(3) was broken down to the much smaller dihedral group D_6.

As a second example, imagine a rod placed vertically on a surface [65]. See figure 2.2. The rod has a circular cross section and as a result, the system has a rotational symme-

[1] If the conditions are right, the microscopic symmetry may even give rise to the macroscopic sixfold symmetry of snowflakes. This occurs, because it is generally more favourable to attach a molecule to a flat side than at a corner. Hence, sides tend to be filled up resulting in a structure with 120° bends.

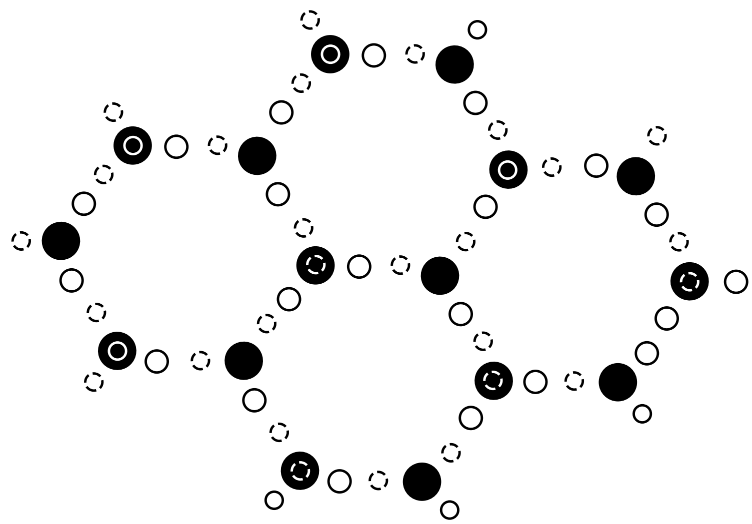


Figure 2.1: Top view of the hexagonal crystal structure of ice. Due to the electronegativity of the oxygen atoms (filled circles), water molecules can form four hydrogen bonds involving its two hydrogen atoms (open circles) and its two lone electron pairs (dashed circles). The angle between two bonds is 104.5°, just right for a hexagonal structure, which generally requires angles of 109.5°, rather than 120°, in order to account for connections between layers [54].

try. We can break the symmetry by exerting a force on top of the rod. As the pressure increases, at some point the rod will start to bend in a particular direction. What direction is inconsequential: all bent states are physically equivalent, but one must be chosen. The act of choosing a direction is what breaks the rotational symmetry. This too is a common feature in spontaneous symmetry breaking. We can also use the rod to demonstrate explicit symmetry breaking. This would amount to exerting a horizontal force on the rod. Now the direction in which the rod bends is not random and the resulting ground state of the system is not degenerate.

From these two examples of spontaneous symmetry breaking, we can distil three common features. First of all, some quantity reaches a critical value at which the nature of the system changes dramatically. Second, the new ground state is not invariant under the original symmetry transformation. Third, the resulting ground state is one of many equivalent possibilities. In the case of the ice crystal, we also saw that the system retained some of its original symmetries. This will become important later on. We shall now study the role of spontaneous symmetry breaking in a quantum field theory.

The simplest field theory with spontaneous symmetry breaking involves a complex scalar field $\phi(x)$ with a quartic potential. The following discussion is based on [70]. We let our

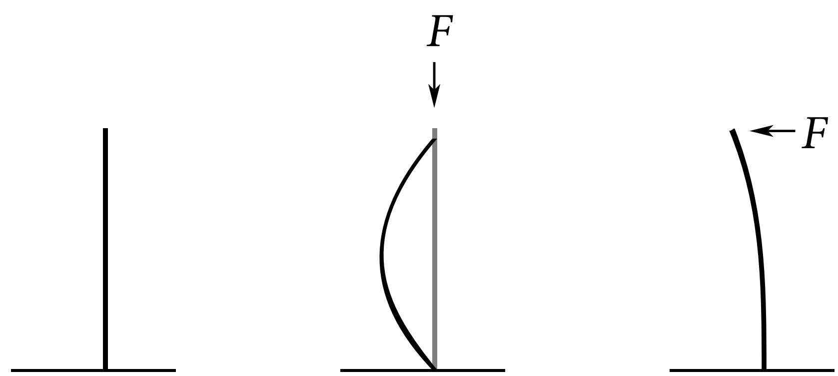


Figure 2.2: A rod placed vertically on a surface. Shown are the unbroken rotational symmetry (left), spontaneous symmetry breaking (middle), and explicit symmetry breaking (right).

Lagrangian be

$$\mathcal{L} = \partial_\mu \phi \, \partial^\mu \phi^* - m^2 \phi\phi^* - \frac{\lambda}{4} \left(\phi\phi^*\right)^2 . \tag{2.1}$$

This Lagrangian is invariant under U(1) transformations which send

$$\phi \to \exp(i\theta)\phi, \qquad \phi^* \to \exp(-i\theta)\phi^*. \tag{2.2}$$

Suppose that this symmetry is broken due to some phase transition. We can model this behaviour and derive predictions without precise knowledge of the underlying mechanisms. Our theory (2.1) becomes an *effective theory* describing the broken phase of some complicated microscopic theory. To do this, we assume that the parameters $m^2 = m^2(T)$ and $\lambda = \lambda(T)$ are continuous functions of some external quantity, say temperature T. Moreover, we assume that the sign of m^2 flips when the temperature reaches some critical point, such as a freezing point. This is perfectly reasonable, because m^2 is just a parameter at this point. Still, in order to prevent problems with interpreting m as mass for $m^2 < 0$, we substitute $\mu^2 = -m^2$ into our Lagrangian:

$$\mathcal{L} = \partial_\mu \phi \, \partial^\mu \phi^* + \mu^2 \phi\phi^* - \frac{\lambda}{4} \left(\phi\phi^*\right)^2 = \partial_\mu \phi \partial^\mu \phi^* - V(\phi).$$

This time, the potential $V(\phi)$ is not minimised at $\phi = 0$, but rather at

$$\phi = \sqrt{\frac{2\mu^2}{\lambda}} e^{i\theta} \equiv v e^{i\theta},$$

up to an arbitrary phase factor. In fact, the point $\phi = 0$ is a local maximum. This tells us that expanding around $\langle\phi\rangle = 0$ would result in an unstable theory. Instead, we should expand around the true vacuum. There are infinitely many possible vacua $\langle\phi\rangle = v e^{i\theta}$, of which the choice $\theta = 0$ is particularly convenient. One possible parametrisation of fluctuations around this vacuum is in terms of two real scalar fiels $\sigma(x)$ and $\pi(x)$. See figure 2.3. We write

$$\phi(x) = \left(v + \frac{\sigma(x)}{\sqrt{2}}\right) \exp\left(\frac{i\pi(x)}{F_\pi}\right), \tag{2.3}$$

for some real constant $F_\pi = v$, which is chosen so that the π-field obtains a canonically normalised kinetic term in the Lagrangian. We immediately see that $V(\phi) = -\mu^2\phi\phi^* + (\lambda/4)(\phi\phi^*)^2$ is independent of $\pi(x)$, which means that $\pi(x)$ must be massless. The presence of massless fields is another common feature of spontaneous symmetry breaking, as we will prove in the next section. Substituting (2.3) into the Lagrangian, we find

$$\mathcal{L} = \tfrac{1}{2}\left(\partial_\mu\sigma\right)^2 + \left(v + \frac{\sigma(x)}{\sqrt{2}}\right)^2 \frac{1}{F_\pi^2}\left(\partial_\mu\pi\right)^2 - V(\sigma).$$

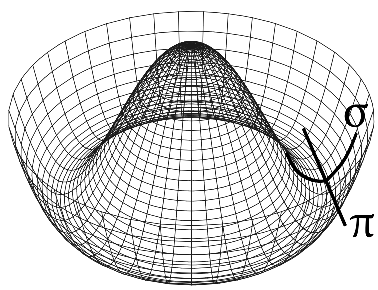


Figure 2.3: Potential energy $V(\phi)$ of the complex scalar field model. After a particular vacuum state has been chosen, the complex scalar field $\phi(x)$ is expanded in terms of two real scalar fields: a radial component $\sigma(x)$ and an angular component $\pi(x)$.

Let's take a step back and consider the symmetry that was broken. Of course, the vacuum $\langle\phi\rangle = v$ is not invariant under the transformations of (2.2). In terms of equation (2.3), the transformations are now realised as

$$\sigma(x) \to \sigma(x), \qquad \text{and} \qquad \pi(x) \to \pi(x) + F_\pi\theta. \tag{2.4}$$

In this case, the original infinitesimal transformation $\delta\phi = i\theta\phi$ was linear in ϕ, whereas the new transformation $\delta\pi = F_\pi\theta$ is not linear in π. In general, spontaneous symmetry breaking induces non-linear realisations. This explains our interest in non-linear realisations of supersymmetry, as we will see in chapter 6. We also observe that the field $\sigma(x)$ is not involved in the transformations at all. It is convenient to decouple $\sigma(x)$ by sending $\mu, \lambda \to \infty$, whilst keeping the ratio $\mu/\sqrt{\lambda}$ fixed. This is equivalent to imposing the constraint

$$\phi\phi^* = v^2. \tag{2.5}$$

Either way, we obtain a theory of a free scalar field:

$$\mathcal{L} = \tfrac{1}{2} \left(\partial_\mu \pi\right)^2 .$$

Here the kinetic term is canonical, because our choice of $F_\pi = v$ was held constant in the limit $m, \lambda \to \infty$. This free scalar field is a Goldstone boson, which we dub the *pion field*, due to its similarity to the real world particles π^0 and $\pi^\pm$. We will meet the pions in section 2.3, but first, in the next section, we shall see that every spontaneously broken continuous symmetry results in such a massless scalar field.

2.2 Goldstone's theorem

Goldstone's theorem [35] states that every spontaneously broken global continuous symmetry gives rise to a massless scalar field. Goldstone's theorem is closely connected to Noether's theorem [55], which states that every continuous global symmetry implies the existence of a conserved current. We will see that the Noether current is related to the Goldstone boson. For completeness, we start with a demonstration of Noether's theorem [73, 65]. We consider a group of global continuous transformations on a set of fields $\phi_\alpha(x)$, such that the action $S = S\left(\phi_\alpha, \partial_\mu \phi_\alpha, x^\mu\right)$ remains invariant. This is an idea that we will consider in much greater detail in chapter 3. For now, we shall content ourselves with writing an arbitrary transformation as

$$\phi_\alpha(x) \to \phi_\alpha(x) + \delta\phi_\alpha(x) = \phi'_\alpha(x).$$

Note that this equation is evaluated at the same spacetime point x^μ. If the group of transformations also causes a change $x^\mu \to x^\mu + \delta x^\mu = x'^\mu$, then we might consider the total variation in ϕ_α as consisting of two parts:

$$\begin{aligned} \Delta\phi_\alpha(x) &= \phi'_\alpha(x') - \phi_\alpha(x') + \phi_\alpha(x') - \phi_\alpha(x) \\ &\cong \delta\phi_\alpha(x) + \left(\partial_\mu \phi_\alpha\right) \delta x^\mu. \end{aligned} \tag{2.6}$$

The transformation group is considered to be a symmetry if the action remains unchanged:

$$0 = \delta S = \int_\Omega \mathrm{d}^4x' \, \mathcal{L}\left(\phi'_\alpha, \partial_\mu \phi'_\alpha, x'^\mu\right) - \int_\Omega \mathrm{d}^4x \, \mathcal{L}\left(\phi_\alpha, \partial_\mu \phi_\alpha, x^\mu\right).$$

Here, the integral is taken over some bounded region Ω in spacetime, where the equations of motion of the fields hold at each point. We can work out the variation in S, by noting that

$$\int \mathrm{d}^4x' \, f(x') = \int \mathrm{d}^4x \, f(x') \left(1 + \partial_\mu \delta x^\mu\right).$$

Hence, we find to first order that

$$\begin{aligned}\delta S &= \int_\Omega \mathrm{d}^4x \ (\mathcal{L} + \delta\mathcal{L})(1 + \partial_\mu \delta x^\mu) - \int_\Omega \mathrm{d}^4x \ \mathcal{L} \\ &\cong \int_\Omega \mathrm{d}^4x \ (\mathcal{L}(\partial_\mu \delta x^\mu) + \delta\mathcal{L}),\end{aligned}$$

where the variation in the Lagrangian density $\mathcal{L}$ is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_\alpha}\delta\phi_\alpha + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\delta(\partial_\mu\phi_\alpha) + (\partial_\mu\mathcal{L})\,\delta x^\mu.$$

If we also use the fact that $\partial_\mu(\mathcal{L}\delta x^\mu) = (\partial_\mu\mathcal{L})\delta x^\mu + \mathcal{L}(\partial_\mu\delta x^\mu)$, we derive

$$\delta S = \int_\Omega \mathrm{d}^4x \ \left(\frac{\partial\mathcal{L}}{\partial\phi_\alpha}\delta\phi_\alpha + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\delta(\partial_\mu\phi_\alpha) + \partial_\mu(\mathcal{L}\delta x^\mu)\right).$$

Subtracting and adding a term, we then find

$$\delta S = \int_\Omega \mathrm{d}^4x \ \left(\frac{\partial\mathcal{L}}{\partial\phi_\alpha} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\right)\delta\phi_\alpha + \int_\Omega \mathrm{d}^4x \ \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\delta\phi_\alpha + \mathcal{L}\delta x^\mu\right). \tag{2.7}$$

Note that the first pair of brackets surround the Euler-Lagrange equations. Since we assumed that the equations of motion hold everywhere in Ω, it follows that the first integral must vanish. Furthermore, adding and subtracting a term, we can write the second integral as:

$$\delta S = \int_\Omega \mathrm{d}^4x \ \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}[\delta\phi_\alpha + (\partial_\nu\phi_\alpha)\delta x^\nu] - \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\partial_\nu\phi_\alpha - \delta^\mu_\nu\mathcal{L}\right]\delta x^\nu\right).$$

Now, looking back at (2.6), we recognise the first expression in square brackets as the total variation $\Delta\phi_\alpha(x)$. It can also be shown that the second expression in square brackets, which we denote as $T^\mu{}_\nu$, is the energy-momentum tensor. This is the conserved current associated with a global translation in spacetime[2]. Now, as the final step, we let ω^a be some infinitesimal parameters, so that the transformations can be written as

$$\Delta\phi_\alpha(x) = \Phi_{\alpha a}(x)\omega^a, \qquad \delta x^\mu(x) = X^\mu{}_a(x)\omega^a, \tag{2.8}$$

where $\Phi_{\alpha a}(x)$ and $X^\mu{}_a(x)$ are two sets of functions that are characteristic of the transformation group. We then demand that the integral

$$\delta S = \int_\Omega \mathrm{d}^4x \ \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\Phi_{\alpha a} - T^\mu{}_\nu X^\nu{}_a\right)\omega^a$$

is zero. Since Ω and ω^a are arbitrary, this implies that

$$\partial_\mu J^\mu{}_a = 0, \qquad \text{with} \qquad J^\mu{}_a \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_\alpha)}\Phi_{\alpha a} - T^\mu{}_\nu X^\nu{}_a. \tag{2.9}$$

[2]For one field $\phi(x)$ and a transformation $x^\mu \to x^\mu + \epsilon^\mu$, simply substitute $X^\mu{}_\nu = \delta^\mu_\nu$ and $\Phi_\mu = 0$ into (2.9).

This is the conserved Noether current. Now, if V is a 3-dimensional volume in space, then

$$Q_a = \int_V \mathrm{d}^3x\; J^0{}_a, \tag{2.10}$$

is a conserved charge. To see this, observe that

$$\dot{Q}_a = \int_V \mathrm{d}^3x\; \partial_0 J^0{}_a = -\int_V \mathrm{d}^3x\; \partial_i J^i{}_a = -\int_{\partial V} \mathrm{d}\sigma_i\; J^i{}_a, \tag{2.11}$$

where we applied the divergence theorem in the final step. This equation demonstrates that any change in the charge contained in the volume V must flow through the surface ∂V. If V is taken to be large enough, then the surface integral may be assumed to vanish, which implies that total charge is conserved.

How do we know that the Noether current says anything meaningful? In general, there are infinitely many trivially conserved currents. This is easiest to see in 2 dimensions, where for every arbitrary scalar-valued function $f(x)$, we have a conserved current of the form [44]

$$J^\mu = \epsilon^{\mu\nu}\partial_\nu f(x),$$

where $\epsilon^{\mu\nu}$ is the Levi-Civita symbol. It is easy to see that J^μ is conserved:

$$\partial_\mu J^\mu = \partial_0 J^0 + \partial_1 J^1 = \partial_0\partial_1 f(x) - \partial_1\partial_0 f(x) = 0.$$

The conservation of such currents is completely independent of the properties of the Lagrangian[3]. Furthermore, the conserved charge associated with this current is rather trivial:

$$Q = \int_a^b \mathrm{d}x\; J^0 = \int_a^b \mathrm{d}x\; \partial_1 f(x) = f(b) - f(a).$$

It is reasonable to assume that $f(b) = f(a) = 0$ for large enough a and b, in which case the total conserved charge Q is identically zero. In contrast with these trivial currents, the Noether current (2.9) is only conserved on-shell, when the equations of motion are satisfied. If this is not the case, then the first integral in (2.7) is nonzero and must be accounted for. Moreover, there is no reason to assume that the conserved charge associated with the Noether current is identically zero, even for very large volumes V, because $J^\mu{}_a$ cannot generally be written as a total divergence.

Let us now turn to a demonstration of Goldstone's theorem. In general, Goldstone's theorem applies to all spontaneously broken continuous symmetries. Yet, the proof most commonly found in textbooks (e.g. [65, p. 291]) applies only to internal symmetries. Including spacetime symmetries is somewhat involved [51] and the examples dealt with in this

[3]These currents are not necessarily trivial when spacetime has nontrivial topological properties. We do not deal with such "topological currents" here.

chapter are all internal symmetries. For these reasons, we restrict ourselves to the usual proof. We thus assume that ${X^\nu}_a = 0$, so that the conserved charge can be written as

$$\begin{aligned} Q_a &= \int_V \mathrm{d}^3x \; {J^0}_a(x) \\ &= \int_V \mathrm{d}^3x \; \frac{\partial \mathcal{L}}{\partial\left(\partial_0 \phi_\alpha\right)} \Phi_{\alpha a}(x) \\ &= \int_V \mathrm{d}^3x \; \pi_\alpha(x) \Phi_{\alpha a}(x), \end{aligned}$$

where π_α is the conjugate momentum of ϕ_α. We can now use the canonical commutation relation $[\phi_\alpha(\vec{x}), \pi_\beta(\vec{y})] = i\delta^3(\vec{x}-\vec{y})\delta^\alpha_\beta$ to derive the commutator between Q_a and ϕ_α. We find

$$[Q_a, \phi_\alpha(\vec{y})] = \int_V \mathrm{d}^3x \; \left[\pi_\alpha(\vec{x}), \phi_\alpha(\vec{y})\right] \Phi_{\alpha a}(\vec{x}) = i\Phi_{\alpha a}(\vec{y}).$$

Now, let's assume that the charges Q_a are generators of the Lie algebra corresponding to the transformation group. We can then construct unitary operators $U = \exp(i\omega^a Q_a)$ that operate on the fields through

$$\begin{aligned} U^\dagger \phi_\alpha U &= \left(1 - i\omega^a Q_a + \ldots\right) \phi_\alpha \left(1 + i\omega^a Q_a + \ldots\right) \\ &\cong \phi_\alpha - i\omega^a [Q_a, \phi_\alpha] \\ &= \phi_\alpha + \omega^a \Phi_{\alpha a}. \end{aligned}$$

Looking back at (2.6) and (2.8), we see that this group action indeed corresponds to the transformation $\phi_\alpha(x) \to \phi'_\alpha(x)$. At this point, we could introduce spontaneous symmetry breaking by demanding that the vacuum state $|\Omega\rangle$ of the theory is not invariant under the transformation group. Note that a state $|\cdot\rangle$ is invariant under the transformation group if and only if $Q_a|\cdot\rangle = 0$ for all a. Hence, we require that[4]

$$Q_a|\Omega\rangle \neq 0.$$

There exists a nice demonstration of Goldstone's theorem [70, p. 564], which uses this relation directly, together with the observation that the charges are conserved, such that $[Q_a, H] = 0$. Nevertheless, we will proceed with a more formal proof. We assume that, for some α, the field $\phi_\alpha(x)$ has a nonzero vacuum expectation value $\langle\Omega|\phi_\alpha(x)|\Omega\rangle \neq 0$. Recall from section 2.1 that this was one of the characteristic features of spontaneous symmetry breaking. Additionally, we assume that $\phi_\alpha(x)$ does not transform as a singlet under Q_a, for some a, so that there exists an operator $\psi(x)$ satisfying

$$[Q_a, \psi(x)] = \phi_\alpha(x). \tag{2.12}$$

[4]Actually, $\|Q_a|\Omega\rangle\| = \infty$ whenever $Q_a|\Omega\rangle \neq 0$, so $Q_a|\Omega\rangle$ does not exist in Hilbert space and is not a proper quantum state. This is the Fabri-Picasso theorem [25]. Ultimately, this has no effect on the results.

Finally, we make the common [78] assumption that the translational invariance of the vacuum is unbroken. Using these assumptions, we will be able to show that there exist massless Goldstone fields. To do this, we use (2.10) to write

$$\langle\Omega|\phi_\alpha(x)|\Omega\rangle = \langle\Omega|[Q_a, \psi(x)]|\Omega\rangle = \int_V \mathrm{d}^3y\ \langle\Omega|[J^0{}_a(y), \psi(x)]|\Omega\rangle \neq 0.$$

Here, the integral is over some 3-dimensional volume V in space, where we must take care to evaluate the integral at equal times $x^0 = y^0$, so that the commutation relation (2.12) is valid. Observe, using (2.11), that the time derivative of this expression is

$$\begin{aligned}\frac{\partial}{\partial y^0}\langle\Omega|\phi_\alpha(x)|\Omega\rangle &= \frac{\partial}{\partial y^0}\int_V \mathrm{d}^3y\ \langle\Omega|[J^0{}_a(y), \psi(x)]|\Omega\rangle \\ &= -\int_{\partial V} \mathrm{d}\sigma_i\ \langle\Omega|[J^i{}_a(y), \psi(x)]|\Omega\rangle,\end{aligned} \tag{2.13}$$

which may be assumed to vanish if the volume V is large enough [36]. Now, introducing a set of intermediate states $|N\rangle$, we obtain [80, 65]

$$\begin{aligned}\langle\Omega|\phi_\alpha(x)|\Omega\rangle &= \int_V \mathrm{d}^3y\ \left\{\langle\Omega|J^0{}_a(x)\psi(x)|\Omega\rangle - \langle\Omega|\psi(x)J^0{}_a(x)|\Omega\rangle\right\} \\ &= \int_V \mathrm{d}^3y\ \left\{\langle\Omega|J^0{}_a(x)|N\rangle\langle N|\psi(x)|\Omega\rangle - \langle\Omega|\psi(x)|N\rangle\langle N|J^0{}_a(x)|\Omega\rangle\right\}.\end{aligned}$$

Next, we assume that the usual relation [36]

$$J^0{}_a(y) = e^{-iPy}J^0{}_a(0)e^{iPy}$$

holds, so that we can use the translational invariance of the vacuum, $e^{iPy}|\Omega\rangle = |\Omega\rangle$, to find

$$\begin{aligned}0 \neq \langle\Omega|\phi_\alpha(x)|\Omega\rangle &= \int_V \mathrm{d}^3y\ \left\{\langle\Omega|J^0{}_a(0)|N\rangle\langle N|\psi(x)|\Omega\rangle e^{ip_N\cdot y} - \langle\Omega|\psi(x)|N\rangle\langle N|J^0{}_a(0)|\Omega\rangle e^{-ip_N\cdot y}\right\} \\ &= (2\pi)^3\delta^3(\vec{p}_N)\left\{\langle\Omega|J^0{}_a(0)|N\rangle\langle N|\psi(x)|\Omega\rangle e^{ip_N^0y^0} - \langle\Omega|\psi(x)|N\rangle\langle N|J^0{}_a(0)|\Omega\rangle e^{-ip_N^0y^0}\right\}.\end{aligned}$$

Finally, because (2.13) is assumed to vanish, it follows that the integral above is independent of y^0. This implies that the masses p_N^0 of the intermediate states must all vanish. Furthermore, the intermediate states must exist for the integral to be nonzero. This concludes the proof of Goldstone's theorem: we have found our massless Goldstone fields.

2.3 Pion physics

In section 2.1, we already alluded to the physical analogue of the Goldstone field π. In this section, we shall see that the machinery of the Goldstone theorem can be used to give a nice description of the real world pions. The pions constitute a class of mesons with masses of around 140 MeV for the charged pions π^+ and π^- and 135 MeV for the neutral pion π^0. Their existence was predicted in 1935 by Yukawa [83] in order to explain the mechanism behind the strong nuclear force, for which the pions can be seen as carrier particles. The mass of such a particle can be predicted using simple physical considerations. Using the diameter of the atomic nucleus $d_{\text{nucl}} \cong 1 \times 10^{-15}$ m and the uncertainty principle, Yukawa predicted that

$$m_{\text{pion}} = \Delta E \cong \frac{1}{\Delta t}\frac{\hbar}{2} \cong \frac{c}{d_{\text{nucl}}}\frac{\hbar}{2} \cong 1.6 \times 10^{-11} \text{ J} \cong 100 \text{ MeV}.$$

The subsequent discovery of the charged pions by Powell, Lattes, and Occhialini, who used photographic emulsions to capture cosmic rays at high altitudes, resulted in Nobel prizes for both Yukawa and Powell. In our modern understanding of QCD, the pions are seen as quark condensates composed of one quark and one antiquark. Furthermore, it turns out that pions can be fruitfully described as pseudo-Goldstone bosons arising from the spontaneous breaking of chiral symmetry. Of course, the pions have (relatively small) masses, which means that they are not proper Goldstone bosons. The reason is that the chiral symmetry is only an approximate symmetry. We will now see how this all fits together. First, we start with the QED and QCD Lagrangian in order to explain chiral symmetry. Then, we will introduce a simpler effective theory, which features the same spontaneous breaking of chiral symmetry. This discussion is largely based on [69, 70, 24]. In our treatment of pion physics, we only require electromagnetism and the strong interaction. In this régime, the gauge symmetries are $\mathrm{SU}(3) \times \mathrm{U}(1)$ and the kinetic contribution to the Lagrangian of the gauge fields is

$$\mathcal{L}_F = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - \tfrac{1}{4}F^I_{\mu\nu}F^{\mu\nu,I},$$

where $I = 1, \ldots, 8$ label the gluons. The contribution of the quarks is

$$\mathcal{L}_q = i\bar{\psi}^i\gamma^\mu \left(\partial_\mu - igA^I_\mu T^I\right)\psi^i,$$

where $i = 1, \ldots, 6$ label the quarks. The matrices T^I are the generators of the SU(3) group and g is the QCD coupling constant. The left- and right-handed components of the quarks are $\psi_L = \frac{1}{2}(1-\gamma^5)\psi$ and $\psi_R = \frac{1}{2}(1+\gamma^5)\psi$. We thus obtain

$$\mathcal{L}_q = i\bar{\psi}^i_L\gamma^\mu \left(\partial_\mu - igA^I_\mu T^I\right)\psi^i_L + i\bar{\psi}^i_R\gamma^\mu \left(\partial_\mu - igA^I_\mu T^I\right)\psi^i_R.$$

Acting independently on the left- and right-handed components, we observe that the Lagrangian $\mathcal{L}_F + \mathcal{L}_q$ is invariant under $\mathrm{U}(6)_L \times \mathrm{U}(6)_R$ transformations. These are global flavour symmetries, which are separate from the local gauge symmetries SU(3) × U(1). The flavour symmetries are broken for various reasons. The first is the fact that quarks have masses:

$$\mathcal{L}_m = -\sum_i m_i \bar{\psi}^i \psi^i = -\sum_i m_i \left(\bar{\psi}^i_L \psi^i_R + \bar{\psi}^i_R \psi^i_L \right).$$

These terms couple the left- and right-handed components. To see that this is true, note that $\bar{\psi}^i = \psi^\dagger \gamma^0$ is the Dirac adjoint of ψ^i, where γ^0 is one of the Gamma matrices in Dirac representation. Then, using $\bar{\psi}^i_{L,R} = \left(\psi^i_{L,R}\right)^\dagger \gamma^0$, it follows that

$$\begin{aligned}
\bar{\psi}^i_L \psi^i_R + \bar{\psi}^i_R \psi^i_L &= \tfrac{1}{4} \left(\psi^i\right)^\dagger \left[\left(1-\gamma^5\right)^\dagger \gamma^0 \left(1+\gamma^5\right) + \left(1+\gamma^5\right) \gamma^0 \left(1-\gamma^5\right) \right] \psi^i \\
&= \tfrac{1}{4} \left(\psi^i\right)^\dagger \left[\left(\gamma^0 - \gamma^5\gamma^0 + \gamma^0\gamma^5 - \gamma^5\gamma^0\gamma^5\right) + \left(\gamma^0 + \gamma^5\gamma^0 - \gamma^0\gamma^5 - \gamma^5\gamma^0\gamma^5\right) \right] \psi^i \\
&= \bar{\psi}^i \psi^i,
\end{aligned}$$

where we used the fact that γ^5 is Hermitian and the identities $\left(\gamma^5\right)^2 = I_{4\times 4}$ and $\gamma^0\gamma^5 = -\gamma^5\gamma^0$. Because of this coupling of left- and right-handed components, the symmetry group is broken down to the group of vector transformations $\mathrm{U}(6)_V$, which act in the same way on the left- and right-handed components. The symmetries are broken down further, because the quark masses m_i are different. Hence, each fermion must be rotated independently, which limits the unbroken symmetry group to $\left(\mathrm{U}(1)_V\right)^6$. Finally, coupling the quarks to electromagnetism,

$$\mathcal{L}_E = -i\bar{\psi}^i \gamma^\mu \left(ieqA_\mu\right) \psi^i,$$

ensures that the symmetry group would be broken to $\left(\mathrm{U}(3)\right)^4$ even in the absence of quark masses. This is due to the different charges of $q = -\frac{1}{3}$ and $q = \frac{2}{3}$. To complete the Lagrangian, we should now add the leptons, but we will not do that here. Instead, we make some simplifying assumptions. In order to discuss pions, we really only need the up and down quarks. In fact, because u and d are so much lighter than the other quarks (and much lighter than the mass scale Λ_{QCD}), it is quite a good approximation if we set $m_u = m_d = 0$. We will also ignore electromagnetism, whose mass scale is of the same order. We then obtain the Lagrangian

$$\mathcal{L} = -\tfrac{1}{4} F^I_{\mu\nu} F^{\mu\nu,I} + i\bar{u}_L \gamma^\mu D_\mu u_L + i\bar{u}_R \gamma^\mu D_\mu u_R + i\bar{d}_L \gamma^\mu D_\mu d_L + i\bar{d}_R \gamma^\mu D_\mu d_R,$$

where the covariant derivative is $D_\mu = \partial_\mu - igA^I_\mu T^I$. The Lagrangian is invariant under independent rotations

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix} \to g_L \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \qquad \begin{pmatrix} u_R \\ d_R \end{pmatrix} \to g_R \begin{pmatrix} u_R \\ d_R \end{pmatrix},$$

where $g_L \in \mathrm{SU}(2)_L$ and $g_R \in \mathrm{SU}(2)_R$. Equivalently, if we let $\psi = (u,d)^{\mathbf{T}}$, we can parametrise the transformations as

$$\psi = \begin{pmatrix} u \\ d \end{pmatrix} \to \exp i \left(\theta_\alpha \tau^\alpha + \gamma^5 \beta_\alpha \tau^\alpha\right) \begin{pmatrix} u \\ d \end{pmatrix},$$

where the τ^α generate SU(2). The transformations with $\beta_\alpha = 0$ form the diagonal (isospin) subgroup. The transformations with $\theta_\alpha = 0$ are the axial transformations. According to Noether's theorem (section 2.2), these symmetries imply the existence of conserved currents:

$$J^{\mu,\alpha} = \bar{\psi}\tau^\alpha\gamma^\mu\psi, \qquad J_5^{\mu,\alpha} = \bar{\psi}\tau^\alpha\gamma^\mu\gamma^5\psi,$$

where $J^{\mu,\alpha}$ is associated with the diagonal transformations and $J_5^{\mu,\alpha}$ with the axial transformations. As we saw in section 2.2, the currents can be integrated over space to yield conserved charges. In this case, we have Q^α and Q_5^α, which generate the diagonal and axial transformations respectively. To relate this to the real world, we should look at the eigenstates of these operators. Consider the three pion states

$$\begin{aligned} \left|\pi^+\right\rangle &= \left|u\bar{d}\right\rangle, \\ \left|\pi^0\right\rangle &= \tfrac{1}{2}\sqrt{2}\left(\left|u\bar{u}\right\rangle - \left|d\bar{d}\right\rangle\right), \\ \left|\pi^-\right\rangle &= -\left|d\bar{u}\right\rangle. \end{aligned} \tag{2.14}$$

It turns out that these states transform under the triplet representation of the isospin group $\mathrm{SU}(2)_V$. In other words, acting with Q^α on one of the three pion states yields another pion state. However, if we act with both Q^α and Q_5^α, then a fourth state $|\sigma\rangle$ must be included, which corresponds to a scalar particle σ. Together, the three pions and the σ transform under the $(\frac{1}{2},\frac{1}{2})$-representation of the full chiral group $\mathrm{SU}(2)\times\mathrm{SU}(2)$. Recall from section 2.2 that conserved charges commute with the Hamiltonian. It follows that σ should have the same mass as the pions. Because this σ-particle is not detected, we surmise that the chiral symmetry must actually be broken in the vacuum state $|\Omega\rangle$ that we observe [45]. Moreover, the unbroken subgroup must be the diagonal group $\mathrm{SU}(2)_V$, since we do observe the triplet of pions. This means that

$$Q^\alpha|\Omega\rangle = 0, \quad \text{but} \quad Q_5^\alpha|\Omega\rangle \neq 0.$$

We could also say that the bilinears $\bar{u}u$ and $\bar{d}d$ acquire a nonzero vacuum expectation value: $\langle\bar{u}u\rangle = \langle\bar{d}d\rangle = V^3 \simeq \Lambda^3_{\mathrm{QCD}}$ and that, as a consequence, there is spontaneous symmetry breaking $\mathrm{SU}(2)\times\mathrm{SU}(2) \to \mathrm{SU}(2)_V$. Without knowing the precise details of the microscopic Lagrangian, it is possible to derive predictions based on the broken symmetry alone. This can be done with an effective Lagrangian. We introduce a 2×2 scalar field matrix $\Sigma(x)$, which transforms under SU(2) × SU(2) as

$$\Sigma \to g_L \Sigma g_R^\dagger, \qquad \Sigma^\dagger \to g_R \Sigma^\dagger g_L^\dagger.$$

This is precisely the $(\frac{1}{2}, \frac{1}{2})$-representation of the three pions and the hypothetical σ-particle, as we shall see. An effective Lagrangian is

$$\mathcal{L} = \mathrm{Tr}\left(\partial_\mu \Sigma\, \partial^\mu \Sigma^\dagger\right) + m^2\, \mathrm{Tr}\left(\Sigma\Sigma^\dagger\right) - \frac{\lambda}{4}\mathrm{Tr}\left(\Sigma\Sigma^\dagger\Sigma\Sigma^\dagger\right).$$

This Lagrangian is known as the linear sigma model and it is invariant under SU(2) × SU(2). As was the case for the simpler sigma model considered before, the vacuum with $\langle\Sigma\rangle = 0$ is not stable. Instead, the potential is minimised for

$$\langle\Sigma\rangle = \frac{v}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad v = \frac{2m}{\sqrt{\lambda}}.$$

Now, the vacuum is no longer invariant under the full group SU(2) × SU(2), but only under the diagonal group $\mathrm{SU}(2)_V$. Decomposing the Lagrangian in terms of a real scalar field σ and an SU(2) triplet $\vec{\pi} = (\pi_1, \pi_2, \pi_3)$, we write

$$\Sigma(x) = \frac{v + \sigma(x)}{\sqrt{2}} \exp\left(2i\frac{\vec{\pi}\cdot\vec{\tau}}{F_\pi}\right),$$

where the constant $F_\pi = 2m/\sqrt{\lambda} = v$ ensures that the $\vec{\pi}$ fields obtain canonical kinetic terms. By using the Baker-Campbell-Hausdorff formula [14],

$$\exp(A)\exp(B) = \exp\left(A + B + \tfrac{1}{2}[A,B] + \tfrac{1}{12}[A,[A,B]] + \tfrac{1}{12}[B,[B,A]] + O\left(A^2B^2\right)\right),$$

we can work out the effect of an infinitesimal transformation $\Sigma \to g_L \Sigma g_R^\dagger$. With $g_L = \exp(i\vec{\theta}_L \cdot \vec{\tau})$ and $g_R = \exp(i\vec{\theta}_R \cdot \vec{\tau})$, we find the action on the pion fields to be

$$\pi^i \to \pi^i + \frac{F_\pi}{2}\left(\theta_L^i - \theta_R^i\right) - \tfrac{1}{2}\epsilon_{ijk}\left(\theta_L^j + \theta_R^j\right)\pi^k + \dots. \tag{2.15}$$

Note that for transformations in the diagonal subgroup, we set $\theta_L^i = \theta_R^i$. Under such transformations, the non-linear term $F_\pi(\theta_L^i - \theta_R^i)$ drops out and we only find the linear transformation $\pi^i \to \pi^i - \frac{1}{2}\epsilon_{ijk}\theta^j\pi^k$. Here, the ϵ_{ijk} are the structure constants of the diagonal group $\mathrm{SU}(2)_V$. We noted before that the pions transform as a triplet of $\mathrm{SU}(2)_V$. Hence, based on these transformation properties, we define $\pi^0 = \pi^3$ and $\pi^\pm = \frac{1}{2}\sqrt{2}\,(\pi^1 \pm i\pi^2)$, relating $\vec{\pi}$ to the physical pion fields (2.14). Meanwhile, the scalar field σ is not involved in the transformations at all and it does not correspond to any physical particle. Therefore, it is best to decouple σ by sending $m, \lambda \to \infty$, whilst keeping F_π fixed, which ensures that the kinetic terms maintain their canonical normalisations. Equivalently, one may set

$$\Sigma\Sigma^\dagger = \left(\frac{v+\sigma}{\sqrt{2}}\right)^2 I = \frac{v^2}{2} I. \tag{2.16}$$

Imposing such a constraint is an application of what we will call the Non-Linearisation Lemma in chapter 3. This lemma states that it is possible to eliminate a field by imposing a constraint, thereby obtaining a consistent non-linear realisation on the space spanned by the remaining fields, provided that the constraint itself is invariant under the transformation group. See section 3.4 for details. Regardless, one finds either by imposing the limit or by imposing the constraint that

$$\Sigma(x) \to U(x) \equiv \frac{v}{\sqrt{2}} \exp\left(2i\frac{\vec{\pi}\cdot\vec{\tau}}{F_\pi}\right),$$

where $U(x)$ is unitary. The remaining degrees of freedom are the three massless non-linearly transforming pion fields. In the next chapter, we shall make more precise many of the ideas used in this chapter. In particular, we will study the theory of linear representations and non-linear realisations.

3

Mathematics of Symmetry Breaking

3.1 Introduction

The previous chapter was a relatively informal discussion of the physics of spontaneous symmetry breaking. In this chapter, we take a more abstract and rigorous point of view. Those readers who are only interested in the physical aspects of non-linear supersymmetry may skip ahead to chapter 5, although it may be worth reading this first section, which is intended as an accessible introduction to the rest of the chapter. In particular, we focus on the role of linear and non-linear representations. A group representation is a structure-preserving map from a group to the group of transformations of a set. See figure 3.1. Usually, one considers transformations of a vector space. We shall refer to representations of this kind as linear representations. Groups can also be realised as transformations of a topological space. To distinguish this from the linear case, we refer to representations of this kind as non-linear realisations. Both linear representations and non-linear realisations are important in spontaneous symmetry breaking.

The basic ingredients for a theory with spontaneous symmetry breaking are: a set of fields $\phi_i(x)$, a group of transformations G acting on those fields, an action

$$S = \int \mathrm{d}^4x \, \mathcal{L}\left(\phi_i, \partial_\mu \phi_i\right),$$

that is invariant under those transformations, and a vacuum state $|\phi_i\rangle$ that is not invariant. Let's make this more precise. We imagine that the fields take values in some abstract space X. We would like the fields to be continuous functions, so we require that X is a topological space. We let the tuple of fields $F(x) = \big(\phi_1(x), \phi_2(x), \dots\big)$ be the continuous function

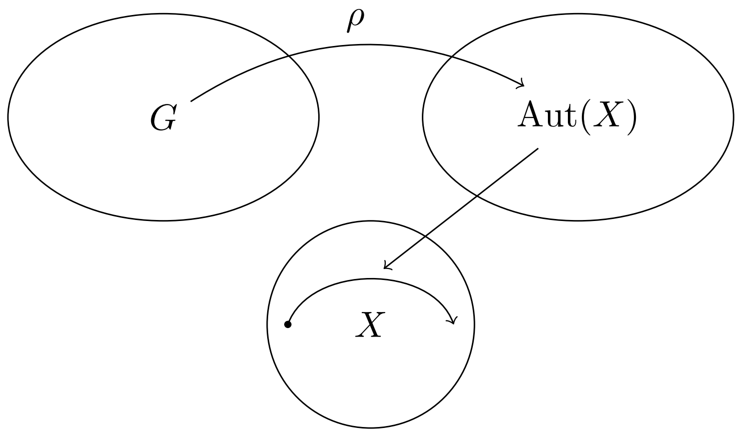


Figure 3.1: The homomorphism ρ maps each group element $g \in G$ to a transformation of X in $\mathrm{Aut}(X)$. The second arrow from the top indicates that elements of $\mathrm{Aut}(X)$ correspond to transformation on X. We refer to ρ as a group representation of G on X.

$F\colon \mathbb{R}^{3,1} \to X$ that maps flat spacetime[1] into X. This is enough structure to consider the role of non-linear realisations. To do this, we associate with each group element $g \in G$ a transformation $\rho(g)$ of X, by which we mean a homeomorphism $\rho(g)\colon X \to X$. Recall that a homeomorphism is a continuous bijection with a continuous inverse. In the next section, we show that a topological group $\tilde{G}$ of such transformations is a *topological transformation group*[2]. Then, a homomorphism $\rho\colon G \to \tilde{G}$ is a non-linear realisation. Recall that a group homomorphism $\rho\colon G \to \tilde{G}$ is a map that satisfies $\rho(g \cdot h) = \rho(g) \circ \rho(h)$, so that it preserves group structure[3]. Clearly, this is consistent with our earlier description of representations as structure-preserving maps from a group to the group of transformations of a set. By abuse of notation, we also refer to X and its elements as non-linear realisations.

Often, the fields are assumed to transform linearly, which is why we need to consider linear representations of the group G. These are constructed in much the same way as non-linear realisations. This time, we require that X is a vector space in addition to being a topological space. Specifically, we assume that X is a topological vector space. We let $\vec{0} \in X$ denote the origin of X, which will represent the vacuum state of the theory. As before, we associate with each element $g \in G$ a transformation of X, but we now require that the transformation is linear. In other words, for each $g \in G$, we have a bijective linear map $\rho(g)\colon X \to X$. Importantly, linear maps leave the origin $\vec{0}$ invariant. This is why unbroken theories can be described by linear representations, but non-linear realisations are necessary for spontaneous symmetry breaking. See figure 3.2.

[1]Spacetime may be curved as well; this is not important for our purposes.
[2]Provided that the map $(f, x) \mapsto f(x)$ is continuous
[3]A group homomorphism between topological groups is also required to be continuous.

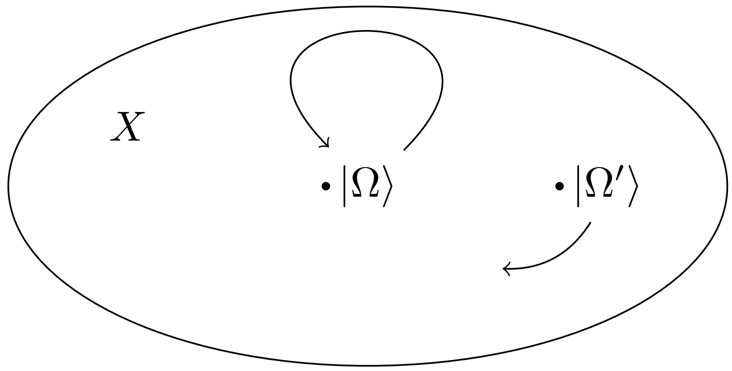


Figure 3.2: The vacuum $|\Omega\rangle$ of the unbroken theory is invariant under the transformation group G, but this is no longer the case when the vacuum is shifted to $|\Omega'\rangle$. If the original transformations were linear with respect to the origin $|\Omega\rangle$, then the transformations are affine with respect to $|\Omega'\rangle$. Only the unbroken theory can be described with linear transformations.

From the previous chapter, we know that spontaneously broken continuous symmetries are associated with massless Goldstone fields. We also saw that those fields generally transform non-linearly, even if the original unbroken theory was linear. The cause of this is a shift in the vacuum state, as illustrated in figure 3.2. If the vacuum is shifted and the fields are expressed as fluctuations about the new vacuum, then the original linear transformations become affine transformations. Affine transformations can be described in terms of their value at one point and a unique linear map. The natural setting for affine transformations is an affine space, which can be understood as a vector space without a designated origin or zero-vector. Thus, if we wish to study linear transformations without choosing a particular origin, it best to consider them as affine transformations on an affine space. This special case, where spontaneous symmetry breaking turns a linear theory into a non-linear, affine theory, is quite common. Therefore, we will study affine transformations in more detail at the end of this section. Before doing that, we consider an example of this special case: the complex scalar field model of section 2.1. However, keep in mind that the results in this chapter also apply to the case where the unbroken theory was already non-linear.

Let us apply the terminology of this chapter to the complex scalar field model. The only field is $\phi(x)$, which takes values in $X = \mathbb{C} \cong \mathbb{R}^2$. With their usual Euclidean topology, in which subsets are open if and only if they contain an open ball around each point, we consider $\mathbb{C}$ as a topological space and $\mathbb{R}^2$ as a topological vector space. The symmetry group is $G = \mathrm{U}(1)$. In the unbroken theory, we have a group action of $\theta \in \mathrm{U}(1)$ on $\phi \in \mathbb{C}$:

$$\phi \to e^{i\theta}\phi = (\cos\theta + i\sin\theta)\,\phi.$$

To write this as a linear representation on $\mathbb{R}^2$, we use the homeomorphism $\phi \mapsto (\mathrm{Im}\ \phi, \mathrm{Re}\ \phi)$, which yields the linear representation

$$\rho(\theta)\begin{pmatrix}\mathrm{Im}\ \phi\\ \mathrm{Re}\ \phi\end{pmatrix}=\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}\mathrm{Im}\ \phi\\ \mathrm{Re}\ \phi\end{pmatrix}.$$

We see that, for each $\theta \in \mathrm{U}(1)$, the origin $\vec{0} = (0,0) \in \mathbb{R}^2$ is invariant. After introducing symmetry breaking and choosing the vacuum $|\phi\rangle = v$, we expand excitations around the vacuum in terms of two real scalar fields $\sigma(x)$ and $\pi(x)$, where $\pi(x)$ is the Goldstone boson. See equation (2.3) and figure 2.3. The transformations are now realised non-linearly:

$$\rho(\theta)(\sigma,\pi) = (\sigma, \pi + F_\pi\theta),$$

as seen in (2.4). This is, in fact, an affine transformation. Clearly, the new vacuum $(0, v)$ is not invariant, which is indicative of spontaneous symmetry breaking. At this point in the discussion, we can proceed in two ways. If we wish to study the Goldstone field, we could decouple the massive scalar field $\sigma(x)$, leaving only the massless field $\pi(x)$. This can be done in a number of ways. In chapter 2, we decoupled $\sigma(x)$ in two ways: by letting the mass of the σ-field increase without bound and by imposing the algebraic constraint $\phi\phi^* = v^2$. At the end of this chapter, in section 3.4, we prove that imposing such a constraint is a consistent way of eliminating a field, provided that the constraint itself is invariant. We call this result the *Non-Linearisation Lemma*. In the scalar field example, the constraint $\phi\phi^* = v^2$ is indeed invariant under $\mathrm{U}(1)$ transformations, so the constraint can be consistently imposed.

On the other hand, we might also want to focus on the remaining unbroken symmetries. In the case of the $\mathrm{U}(1)$ model, the unbroken subgroup is just the trivial group, which is why the transformation $\sigma \to \sigma$ is not particularly interesting. Generally however, the unbroken subgroup is not trivial. For example, recall that the dihedral group D_6 remained unbroken in the snowflake model. In section 3.3, we prove that if we restrict ourselves to the unbroken subgroup $H \leq G$, then the non-linear realisation of the broken model can be turned into a linear representation by means of a suitable coordinate transformation[4]. Such a coordinate transformation should be a homeomorphism from X into a topological vector space V. Of course, the existence of such a homeomorphism requires that X is locally homeomorphic to a topological vector space, at least in a neighbourhood of the vacuum $|\Omega\rangle$. The existence of global coordinates is irrelevant, since the linear approximation is only accurate in a neighbourhood of the vacuum. In general, a manifold satisfies this local condition at

[4]Coleman, Wess, and Zumino speak of "allowed coordinate transformations" of the form $\phi = \chi F(\chi)$, $F(0) = 1$, which guarantees that the on-shell S-matrix is the same, regardless of whether the Lagrangian is expressed in terms of the fields χ_i or ϕ_i [21]. Although this is important, we ignore the point entirely. It is easy to modify our approach to ensure that the coordinate transformation is "allowed", but it requires additional structure on X (multiplication) which we prefer to leave out at this point.

each point. For example, an ordinary real manifold is locally homeomorphic to Euclidean space, whereas a supermanifold is locally homeomorphic to superspace (see chapter 4). It is therefore useful to remain as general as possible and consider arbitrary topological vector spaces. The linearisation procedure is useful, because it allows us to extract a linear theory from a non-linear theory involving both broken and unbroken symmetries. We refer to this result as the *Linearisation Lemma*, which is based on the work of Coleman, Callan, Wess, and Zumino [21, 13] and Bochner [9]. The overall picture is shown in the diagram below:

$$\begin{tikzcd} \{\phi_i\}_{\text{L or NL}} \arrow[r, "\text{SSB}"] & \{\phi_i\}_{\text{NL}} \arrow[r, "\text{NLL}"] \arrow[d, "\text{LL}"] & \{\pi_i\}_{\text{NL}} \\ & \{\sigma_i\}_{\text{L}} \otimes \{\pi_i\}_{\text{NL}} \arrow[ur, "\text{NLL}"] & \end{tikzcd} \tag{3.1}$$

A theory with fields ϕ_i transforming linearly or non-linearly undergoes spontaneous symmetry breaking (SSB), resulting in a theory with non-linearly transforming fields ϕ_i. With the Linearisation Lemma (LL), we recover a theory with the remaining linearly transforming fields σ_i and the non-linearly transforming Goldstone bosons π_i. With the Non-Linearisation Lemma (NLL), the Goldstone fields π_i can be extracted. See also figure 3.3.

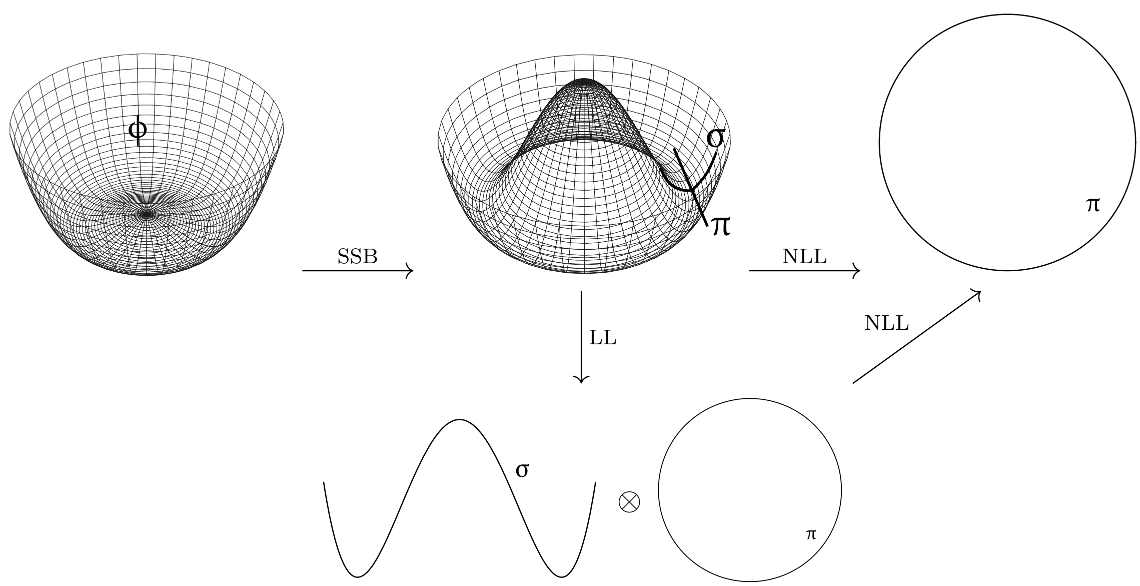


Figure 3.3: Application of the diagram in (3.1) to the U(1) model of chapter 2.

Now, we return to our discussion of affine transformations. Let us briefly review affine spaces and affine maps. For more details, we refer to [6, 72, 31]. Our discussion is mostly based on [31]. As mentioned above, an affine space is, intuitively speaking, a vector space without an origin. This lack of origin has major consequences. In vector spaces, we usually do not distinguish between points and vectors. For instance, the point $(1,1) \in \mathbb{R}^2$ is readily identified with the vector from the origin $(0,0)$ to this point. If there is no origin, then no such identification can be made. In affine spaces, we can only talk about points and vectors between points. This leads us to the following definition [31].

Definition 3.1. An *affine space* $\langle E, \overrightarrow{E}, + \rangle$ is a set E, together with a vector space $\overrightarrow{E}$ and an action $E \times \overrightarrow{E} \to E$ of the vector space on the set, called *translation* and denoted by $+$, which satisfies

1. (Identity): $x + \vec{0} = x$, for all $x \in E$,
2. (Associativity): $(x + \vec{v}) + \vec{w} = x + (\vec{v} + \vec{w})$, for all $x \in E$ and $\vec{v}, \vec{w} \in \overrightarrow{E}$,
3. (Uniqueness): For any $x, y \in E$, there exists a unique $\vec{v} \in \overrightarrow{E}$, such that $x + \vec{v} = y$.

The unique vector from x to y is also denoted as $y - x$ or $\overrightarrow{xy}$.

Geometrically, affine space is a generalisation of Euclidean space. Because points cannot be identified with vectors, there is no conception of angles between points. Nevertheless, one can still talk about parallel lines, or more generally, parallel subsets of an affine space. Results in Euclidean geometry that do not depend on absolute distances and angles can still be proved in affine geometry. One elementary result is Chasles' identity.

Lemma 3.1 (Chasles' identity)**.** *For any points a, b, c in an affine space E, we have*

$$\overrightarrow{ac} = \overrightarrow{ab} + \overrightarrow{ac}.$$

Proof. By the uniqueness property, we have $c = a + \overrightarrow{ac} = b + \overrightarrow{bc}$ and $b = a + \overrightarrow{ab}$. Hence,

$$c = b + \overrightarrow{bc} = (a + \overrightarrow{ab}) + \overrightarrow{bc} = a + (\overrightarrow{ab} + \overrightarrow{bc}),$$

where we used the associativity property in the last step. The conclusion now follows from the uniqueness property. □

This identity will be useful later on. From an algebraic point of view, many concepts in linear algebra also have an affine analogue. For example, corresponding to the notion of linear combinations of vectors, there exists the notion of affine combinations of points. Affine combinations are more restrictive than linear combinations, because linear combinations are not origin-independent unless a certain condition is satisfied. To see this, let $(0,0) \in \mathbb{R}^2$ be the origin of the vector space $\mathbb{R}^2$ and let $\vec{e}_1 = (1,0)$ and $\vec{e}_2 = (0,1)$ be the standard basis. In this basis, we compute a linear combination of the points $a = (1,1)$ and $b = (1,0)$:

$$\lambda_1 a + \lambda_2 b = (\lambda_1 + \lambda_2, \lambda_1). \tag{3.2}$$

If we shift the origin to $(1,1)$, then the points become $a=(0,0)$ and $b=(0,-1)$ with respect to the new origin. The linear combination, with respect to the new origin, becomes

$$\lambda_1 a + \lambda_2 b = (0, -\lambda_2). \tag{3.3}$$

Shifting this back, we see that (3.3) equals $(1, 1-\lambda_2)$ with respect to the old origin. This differs from (3.2) unless $\lambda_1 + \lambda_2 = 1$. Thus, in general, it is impossible to write down an unambiguous linear combination of points. However, as the previous example suggests, it turns out that if the scalars λ_1, λ_2 – or more generally, the family of scalars $(\lambda_i)_{i\in I}$ – sum to unity, then the linear combination is origin-independent. This is the content of the following lemma.

Lemma 3.2. *Let $(a_i)_{i\in I}$ be a family of points in an affine space E, and let $(\lambda_i)_{i\in I}$ be a family of scalars with compact support*[5]*, such that $\sum_{i\in I}\lambda_i = 1$. Then,*

$$a + \sum_{i\in I} \lambda_i \overrightarrow{aa_i} = b + \sum_{i\in I} \lambda_i \overrightarrow{ba_i},$$

for any two points $a, b \in E$.

Proof. Using Lemma 3.1 and the associativity and uniqueness properties, we evaluate the left-hand side as

$$\begin{aligned} a + \sum_{i\in I} \lambda_i \overrightarrow{aa_i} &= a + \sum_{i\in I} \lambda_i \left(\overrightarrow{ab} + \overrightarrow{ba_i}\right) \\ &= a + \sum_{i\in I} \lambda_i \overrightarrow{ab} + \sum_{i\in I} \lambda_i \overrightarrow{ba_i} \\ &= a + \overrightarrow{ab} + \sum_{i\in I} \lambda_i \overrightarrow{ba_i} \\ &= b + \sum_{i\in I} \lambda_i \overrightarrow{ba_i}. \end{aligned}$$

□

Thus, the following definition makes sense.

Definition 3.2 (Affine combination)**.** Let $(a_i)_{i\in I}$ be a family of points in an affine space E, and let $(\lambda_i)_{i\in I}$ be a family of scalars with compact support, such that $\sum_{i\in I}\lambda_i = 1$. Then, we write an *affine combination* of the points $(a_i)_{i\in I}$ as

$$\sum_{i\in I} \lambda_i a_i = o + \sum_{i\in I} \lambda_i \overrightarrow{oa_i},$$

where $o \in E$ is any point in E.

[5]This means that $J = \{i \in I \mid \lambda_i \neq 0\}$ is a finite subset of I.

Now, we can translate concepts in linear algebra to their affine counterparts by replacing linear combinations with affine combinations. For instance, a linear subspace $\overrightarrow{W}$ of a vector space $\overrightarrow{V}$ is a subset of $\overrightarrow{V}$ closed under linear combinations. Thus, it is natural to define affine subspaces as subsets of an affine space closed under affine combinations. For our purposes, we are mostly interested in affine transformations. Noting that a linear map between vector spaces $\overrightarrow{V}$ and $\overrightarrow{V}'$ (over the same field) is a map that preserves linear combinations, it is natural to define affine maps in the same way.

Definition 3.3 (Affine map)**.** Let E and E' be two affine spaces with associated vector spaces over the same field. We say that a map $f\colon E \to E'$ is an *affine map* if it satisfies

$$f\left(\sum_{i\in I} \lambda_i a_i\right) = \sum_{i\in I} \lambda_i f\left(\lambda_i a_i\right),$$

for any affine combination of points $(a_i)_{i\in I}$ in E.

We mentioned before that affine maps can be described in terms of their value at one point and a unique linear map. Let's now prove this statement.

Lemma 3.3. *Let $f\colon E \to E'$ be an affine map between affine spaces E and E'. Then, there exists a unique linear map $\Lambda\colon \overrightarrow{E} \to \overrightarrow{E}'$ between the associated vector spaces, such that*

$$f(a+\vec{v}) = f(a) + \Lambda\vec{v}, \tag{3.4}$$

for any $a \in E$ and $\vec{v} \in \overrightarrow{E}$.

Proof. Let $a \in E$ be arbitrary, then the map $\Lambda\colon \overrightarrow{E} \to \overrightarrow{E}'$, defined by

$$\Lambda\vec{v} = f(a+\vec{v}) - f(a), \tag{3.5}$$

is a linear map. To see this, we note that $\overrightarrow{aa} = \vec{0}$ and $\overrightarrow{a(a+\vec{v})} = \vec{v}$, such that

$$\begin{aligned} a + \alpha\vec{v} + \beta\vec{w} &= a + \alpha\overrightarrow{a(a+\vec{v})} - \alpha\overrightarrow{aa} + \beta\overrightarrow{a(a+\vec{w})} + (1-\beta)\overrightarrow{aa} \\ &= \alpha(a+\vec{v}) - \alpha a + \beta(a+\vec{w}) + (1-\beta)a, \end{aligned}$$

is an affine combination for any scalars α, β and vectors $\vec{v}, \vec{w} \in \overrightarrow{E}$. Hence, since affine maps preserve affine combinations, we see that

$$\begin{aligned} \Lambda\left(\alpha\vec{v} + \beta\vec{w}\right) &= f(a + \alpha\vec{v} + \beta\vec{w}) - f(a) \\ &= \alpha f(a+\vec{v}) - \alpha f(a) + \beta f(a+\vec{w}) + (1-\beta)f(a) - f(a) \\ &= \alpha\left(f(a+\vec{v}) - f(a)\right) + \beta\left(f(a+\vec{w}) - f(a)\right) \\ &= \alpha\Lambda\vec{v} + \beta\Lambda\vec{w}. \end{aligned}$$

To see that Λ is well-defined, we let $b \in E$ be another point. Then,

$$b + \vec{v} = (a + \overrightarrow{ab}) + \vec{v} = a + (\vec{v} + \overrightarrow{ab}) = a + \overrightarrow{a(a+\vec{v})} - \overrightarrow{aa} + \overrightarrow{ab} = (a + \vec{v}) - a + b,$$

which is also an affine combination. Hence, it follows that

$$f(b + \vec{v}) - f(b) = f(a + \vec{v}) - f(a) + f(b) - f(b) = f(a + \vec{v}) - f(a),$$

which shows that definition (3.5) is independent of the point $a \in E$. Uniqueness now follows immediately from (3.4). □

It turns out that the previous lemma implies a version of the Linearisation Lemma for affine realisations. Recall that the Linearisation Lemma states that when a non-linear realisation of a group G on a topological space X is restricted to the unbroken subgroup H, then a linear representation can be obtained by means of a suitable coordinate map from X to a vector space V. The unbroken subgroup consists of those transformations that fix the vacuum $\Omega \in X$. Let us see how this works for affine realisations, where the space $X = E$ is an affine space and $V = \overrightarrow{E}$ its associated vector space.

Definition 3.4 (Affine realisation)**.** Let E be an affine space, such that $\mathrm{Aut}(E)$ is the group of affine transformations of E. Then, we call a group homomorphism $\rho\colon G \to \mathrm{Aut}(E)$ an *affine realisation* of the group G on E.

Definition 3.5 (Linear representation)**.** Let V be a vector space, such that $\mathrm{Aut}(V)$ is the group of bijective linear transformations of V. Then, we call a group homomorphism $\rho\colon G \to \mathrm{Aut}(V)$ a *linear representation* of the group G on V.

Lemma 3.4 (Affine Linearisation Lemma)**.** *Let $\langle E, \overrightarrow{E}, + \rangle$ be an affine space and $\rho\colon G \to \mathrm{Aut}(E)$ an affine realisation on E. Fix a point $a \in E$ and let $H \leq G$ be the subgroup of transformations satisfying $\rho(h)(a) = a$ for all $h \in H$. Then, the homomorphism $\tilde{\rho}\colon H \to \mathrm{Aut}(\overrightarrow{E})$, defined by $\tilde{\rho}(h)(\vec{x}) = \rho(h)(a + \vec{x}) - a$, is a linear representation on $\overrightarrow{E}$.*

Proof. Let $h \in H$ be arbitrary. By Lemma 3.3, the affine transformation can be written as

$$\rho(h)(a + \vec{x}) = \rho(h)(a) + \Lambda(h)\vec{x},$$

for a unique linear map $\Lambda(h)$. Since $\rho(h)(a) = a$, we see that

$$\tilde{\rho}(h)(\vec{x}) = \rho(h)(a + \vec{x}) - a = \Lambda(h)\vec{x},$$

which is a linear transformation of $\overrightarrow{E}$. To see that $\tilde{\rho}$ is a homomorphism, let $h, h' \in H$ be arbitrary. Then, it follows that

$$\begin{aligned}
\tilde{\rho}(hh')(\vec{x}) &= \rho(hh')(a+\vec{x}) - a \\
&= \rho(h) \circ \rho(h')(a+\vec{x}) - a \\
&= \rho(h)(a + \rho(h')(a+\vec{x}) - a) - a \\
&= \tilde{\rho}(h) \circ \tilde{\rho}(h')(\vec{x}).
\end{aligned}$$

We thus have a linear representation of H on $\overrightarrow{E}$. □

This lemma is a special case of the Linearisation Lemma of section 3.3. There, we will prove a similar result for differentiable non-linear transformations. These can be written as

$$\rho(g)(a+\vec{x}) = \rho(g)(a) + \Lambda(g)\vec{x} + \psi(g)(\vec{x}),$$

where $\psi(g)(\vec{x})$ is a super-linear function that rapidly approaches zero near $\vec{x} = 0$. The Linearisation Lemma works by "averaging out" the non-linear part $\psi(g)(\vec{x})$. In order to be able to do this, it is necessary that the group G satisfies several topological properties, so that group integration is possible. Specifically, G must be a locally compact Hausdorff topological group G and the unbroken subgroup H, for which $\rho(h)(a) = a$ for all $h \in H$, must be compact. To our knowledge, it is not known if a similar result holds if these properties are not satisfied, except in special cases such as Lemma 3.4 above. See also [21, p. 2240]. We briefly consider another aspect of this discussion at the end of section 3.3.

This concludes our discussion of affine transformations. Now follows section 3.2, where we formally define the necessary concepts and prove a few lemmas. Then, in section 3.3, we prove the Linearisation Lemma of Bochner and Coleman, Wess, and Zumino in a more general setting. Finally, in section 3.4, we prove the Non-Linearisation Lemma.

3.2 Definitions

In order to define non-linear realisations, we first need the following definitions [57, 52, 40].

Definition 3.6 (Topological group)**.** A group G that is also a topological space is a *topological group* if

1. group multiplication is a continuous map $G \times G \to G$,
2. inversion is a continuous map $G \to G$,

where $G \times G$ has the product topology[6].

[6] The product topology τ of the Cartesian product $X \times Y$ of two topological spaces X and Y consists of all sets $E \times F$, where $E \subset X$ and $F \subset Y$ are both open.

Definition 3.7 (Topological transformation group)**.** Let G be a topological group and X a topological space. One calls G a *topological transformation group* of X if there exists a continuous function $\alpha\colon G \times X \to X$, satisfying

1. $\alpha(e, x) = x$, for the identity $e \in G$,
2. $\alpha(gg', x) = \alpha(g, \alpha(g', x))$, for $g, g' \in G$.

We may write $\alpha(g, x) = gx$ if there is no confusion.

In particular, the group of all self-homeomorphisms of a topological space can be a topological transformation group, as the following lemma shows.

Lemma 3.5. *Let X be a topological space. Then the set* $\mathrm{Homeo}(X)$ *of self-homeomorphisms $f\colon X \to X$ forms a group under composition of functions. Endowing* $\mathrm{Homeo}(X)$ *with a suitable topology, such as the compact-open topology*[7]*, makes it a topological group. Then,* $\mathrm{Homeo}(X)$ *is a topological transformation group if and only if the mapping* $\mathrm{Homeo}(X)\times X \to X$*, defined by $(f, x) \mapsto f(x)$ is continuous.*

Proof. Composition of homeomorphisms yields another homeomorphism and is associative. Furthermore, for each homeomorphism f, there exists an inverse homeomorphism f^{-1} by definition. Also, the identity map $\mathrm{id}_X\colon X \to X$ is a homeomorphism and therefore contained in $\mathrm{Homeo}(X)$. Thus, $\mathrm{Homeo}(X)$ is a group under composition of functions.

For the second part, assume first that $\mathrm{Homeo}(X)$ is a topological transformation group. Then, by Definition 3.7, we have a continuous function $\alpha\colon \mathrm{Homeo}(X) \times X \to X$, satisfying $\alpha(e, x) = \alpha(\mathrm{id}_X, x) = x$ and $\alpha(gg', x) = \alpha(g \circ g', x) = \alpha(g, \alpha(g', x))$. Reinterpreting group multiplication as composition of functions, we see that this is precisely the continuous function $\alpha(f, x) = f(x)$ we needed. The converse is immediate. $\square$

We can now define non-linear group realisations.

Definition 3.8 (Non-linear realisation)**.** A *non-linear realisation* of a group G on a topological space X is a homomorphism $\rho\colon G \to \tilde{G}$, where $\tilde{G}$ is a topological transformation group of X.

It is instructive to restate the definition of linear representations as follows.

Definition 3.9 (Linear representation)**.** A *linear representation* of a group G on a vector space V is a homomorphism $\rho\colon G \to \mathrm{Aut}(V)$, where $\mathrm{Aut}(V)$ is the automorphism group of the vector space V.

---

[7]The compact-open topology τ on $\mathrm{Homeo}(X)$ is defined as follows. Let $K \subset X$ be a compact subset and $U \subset X$ an open subset. Let $V(K, U) = \{f \in \mathrm{Homeo}(X) \mid f(K) \subset U\}$. The collection V of sets $V(K, U)$, for all compact K and open U, is a subbase for τ. In other words, V is a subcollection of τ that generates τ.

We already saw an example of a non-linear realisation in section 3.1, but that was an affine realisation. Let us now consider a non-affine example.

Example 3.1. Endow the subset $X = (0, \infty) \subset \mathbb{R}$ with subspace topology[8]. We note that $G = (X, *, 1)$ forms a topological group under multiplication. We can define a non-linear realisation of G on X by

$$\rho(g)(x) = x^g.$$

These functions are clearly self-homeomorphisms on X. They also satisfy the property

$$\rho(g) \circ \rho(h)(x) = \left(x^h\right)^g = x^{gh} = \rho(gh)(x),$$

which makes $\rho\colon G \to \mathrm{Homeo}(X)$ a homomorphism. □

In order to make sense of the Linearisation Lemma, we need to define what it means to linearise a non-linear realisation. This can be done by endowing a topological space with vector space structure. Let us recall a few facts about topological vector spaces over topological fields. Topological vector spaces and topological fields are defined as follows.

Definition 3.10 (Topological field)**.** Let F be a field that is also a topological space. We call F a *topological field* if

1. addition is a continuous map $F \times F \to F$,
2. multiplication is a continuous map $F \times F \to F$,
3. (multiplicative) inversion is a continuous map $F \backslash \{0\} \to F \backslash \{0\}$,

where $F \times F$ has the product topology.

Definition 3.11 (Topological vector space)**.** A vector space V over a topological field K is a *topological vector space* if

1. for every $\vec{x} \in V$, the singleton $\{\vec{x}\}$ is closed,
2. vector addition is a continuous map $V \times V \to V$,
3. scalar multiplication is a continuous map $K \times V \to V$,

where the domains of these functions have the product topology.

[8] The subspace topology τ_S of a subset S of a topological space X consists of all sets $E \subset S$ that can be written as an intersection $E = F \cap S$, where F is open in X.

The first condition of Definition 3.11 implies that distinct points have distinct neighbourhoods. To see this, let $\vec{x}, \vec{y}$ be distinct. Then, $\{\vec{x}\}$ is closed, so its complement $\{\vec{x}\}^{\mathrm{c}}$ is an open neighbourhood of $\vec{y}$. Similarly, $\{\vec{y}\}^{\mathrm{c}} \neq \{\vec{x}\}^{\mathrm{c}}$ is an open neighbourhood of $\vec{x}$. We follow [62, 48] by including this condition in the definition, because it allows us to prove the even stronger statement that topological vector spaces are Hausdorff spaces, where distinct points have disjoint neighbourhoods. Here, we are using the following definition of neighbourhoods.

Definition 3.12 (Neighbourhood)**.** Let X be a topological space. Then, a subset $V \subset X$ is a *neighbourhood* of a point p if V contains an open set U, such that $p \in U$. If V is itself open, then V is an *open neighbourhood*.

Before proving that topological vector spaces are Hausdorff, we first need to introduce some more definitions and prove another lemma. In doing this, we always assume that the topological field has an associated absolute value function, which is defined as follows.

Definition 3.13 (Absolute value)**.** Let K be a topological field. Then, we can define a continuous map $|\cdot|\colon K \to \mathbb{R}$, called the *absolute value function*, satisfying for all $x, y \in K$:

1. (Non-negativity): $|x| \geq 0$,
2. (Non-degenerateness): $|x| = 0$ if and only if $x = 0$,
3. (Multiplicativity): $|xy| = |x||y|$,
4. (Sub-additiveness): $|x + y| \leq |x|+|y|$.

It follows from the multiplicativity and non-negativity properties that, if $1 \in K$ is the multiplicative identity, then $|1| = 1$.

Using the absolute value function, we can define balanced sets.

Definition 3.14 (Balanced set)**.** Let V be a topological vector space over a topological field K. A subset $E \subset V$ is *balanced* if

$$\alpha E \subset E, \qquad \forall \alpha \in K : |\alpha| \leq 1.$$

To prove that topological vector spaces are Hausdorff, we will also use symmetric sets.

Definition 3.15 (Symmetric set)**.** Let V be a topological vector space over a topological field K. A subset $E \subset V$ is *symmetric* if $E = -E$. In other words, E is symmetric if

$$\vec{v} \in E \implies -\vec{v} \in E.$$

With these definitions, let us prove some general properties of topological vector spaces. The next two lemmas are standard results [62, 48].

Lemma 3.6. *Let V be a topological vector space. For any neighbourhood U of $\vec{0}$, there exists*

1. *a balanced neighbourhood A of $\vec{0}$, such that $A \subset U$,*
2. *a symmetric neighbourhood A of $\vec{0}$, such that $A + A \subset U$,*
3. *a balanced neighbourhood A of $\vec{0}$, such that $aA + bA \subset U$, for any scalars $a, b \in K$ with $|a| \leq 1$ and $|b| \leq 1$.*

Proof. Let U be a neighbourhood of $\vec{0}$, so that there exists an open set $W \subset U$ containing $\vec{0}$. We treat each case separately:

1. Note that $0 \cdot \vec{0} = \vec{0}$ and that scalar multiplication is continuous. Hence, the pre-image of W under scalar multiplication, which we write as $D \times X$, is open. Since $K \times V$ has the product topology, both D and X must be open. It follows that, for some $0 < \delta \in \mathbb{R}$,

   $$E = \{x \in K : |x| < \delta\} \subset D.$$

   Indeed, suppose that this is false for all δ, then there must be a point $x \notin D$, such that $|x| < \delta$ for all $\delta > 0$. By the non-degenerateness property of the absolute value function, we must have $x = 0$, but this is a contradiction, since $0 \in D$.

   To see that $E \cdot X$ is a balanced set, let $\vec{v} \in \alpha E \cdot X$ for $|\alpha| \leq 1$. Then, $\vec{v} = \alpha e \vec{x}$ for $e \in E$ and $\vec{x} \in X$. Clearly, $|\alpha e| < \delta$, such that $\vec{v} \in E \cdot X$. Hence, $E \cdot X$ is balanced.

   Since the non-zero scalar multiplication $\vec{x} \mapsto \alpha \vec{x}$ has a continuous inverse (namely $\vec{x} \mapsto \alpha^{-1} \vec{x}$), it is a homeomorphism. Hence, $\alpha \cdot X$ is open by continuity. The union

   $$\bigcup_{0<|\alpha|<\delta} \alpha \cdot X = E \cdot X,$$

   is also open. Therefore, $E \cdot X \subset D \cdot X = W$ is a balanced neighbourhood of $\vec{0}$.

2. Next, observe that $\vec{0} + \vec{0} = \vec{0}$ and that the pre-image $X = \{(\vec{x}, \vec{y}) \in V \times V \ : \ \vec{x} + \vec{y} \in W\}$ of W under vector addition is open by continuity. Since $V \times V$ has the product topology, there must be two open sets X_1 and X_2, both containing $\vec{0}$, such that $X_1 \times X_2 \subset X$. Let $A = X_1 \cap (-X_1) \cap X_2 \cap (-X_2)$. Clearly, A is symmetric. Furthermore, as the intersection of four open sets containing $\vec{0}$, A is an open neighbourhood of $\vec{0}$.

3. Finally, part 3 follows from the first two parts. □

Lemma 3.7. *Every topological vector space V is Hausdorff.*

Proof. We must prove that distinct points have disjoint neighbourhoods. By definition, the distinct singletons $\{\vec{x}\}$ and $\{\vec{y}\}$ are both closed. Let $C = \{\vec{y}\}$. It is clear that the complement C^{c} is an open neighbourhood of $\vec{x}$. Furthermore, $(C^{\mathrm{c}} - \vec{x})$ is an open neighbourhood of $\vec{0}$, because the map $\vec{v} \mapsto \vec{v} + \vec{x}$ is a homeomorphism. By applying part 2 of Lemma 3.6 twice, we see that there exists a symmetric sub-neighbourhood W of $\vec{0}$, such that

$$W + W + W \subset W + W + W + W \subset (C^{\mathrm{c}} - \vec{x}),$$

which implies that

$$(\vec{x} + W + W + W) \cap C = \varnothing.$$

Using the fact that W is symmetric, we find that

$$(\vec{x} + W + W) \cap (C + W) = \varnothing.$$

The conclusion now follows from the fact that $\vec{x} + W + W$ is a neighbourhood of $\vec{x}$ and $\vec{y} + W = C + W$ is a neighbourhood of $\vec{y}$. □

For the Linearisation Lemma, differentiability is crucial. Differentiable functions on topological vector spaces can be defined using the notion of tangency to $\vec{0}$ [50].

Definition 3.16 (Tangent to $\vec{0}$)**.** Let V and W be topological vector spaces. A function $\phi\colon V \to W$ is said to be *tangent to* $\vec{0}$ if, given a neighbourhood B of $\vec{0}_W$ in W, there exists a neighbourhood A of $\vec{0}_V$ in V, such that

$$\phi(tA) \subset o(t)B,$$

for some function $o(t)\colon K \to K$ satisfying $\lim\limits_{t\to 0} o(t)/t = 0$.

Definition 3.17 (Differentiability)**.** Let V and W be topological vector spaces. A continuous function $f\colon V \to W$ is differentiable at $\vec{x_0} \in V$ if there exists a continuous linear map $\Lambda\colon V \to W$, such that

$$f(\vec{x}_0 + \vec{y}) = f(\vec{x_0}) + \Lambda\vec{y} + \phi(\vec{y}),$$

implies that ϕ is tangent to $\vec{0}$. It is easy to prove that Λ is unique if it exists. We call Λ the derivative of f at $\vec{x}_0$.

Next, we state the following lemma for future reference.

Lemma 3.8. *Let U, V, and W be topological vector spaces. Let $\phi\colon V \to W$ and $\psi\colon V \to W$ be tangent to $\vec{0}$, let $f\colon U \to V$ be a continuous function that satisfies $f(\vec{0}_U) = \vec{0}_V$ and the condition that for every neighbourhood B of $\vec{0}_V$, there exists a neighbourhood A of $\vec{0}_U$, such that $f(tA) \subset tB$. Finally, let $\Lambda\colon W \to U$ be a continuous linear map. Then,* *1.* $\phi \circ f(\vec{x})$ *is tangent to* $\vec{0}$*,*

2. $\Lambda \circ \phi(\vec{x})$ *is tangent to* $\vec{0}$*,*

3. $\phi(\vec{x}) + \psi(\vec{x})$ *is tangent to* $\vec{0}$*.*

Proof. We treat each case separately:

1. Let $B \subset W$ be a neighbourhood of $\vec{0}_W$. Because ϕ is tangent to $\vec{0}$, there exists a neighbourhood $A \subset V$ of $\vec{0}_V$, such that $\phi(tA) \subset o(t)B$ for some function $o(t)$, satisfying $\lim_{t\to 0} o(t)/t = 0$. Furthermore, by assumption, there exists a neighbourhood A' of $\vec{0}_V$, such that $f(tA') \subset tA$. Then,
$$\phi \circ f(tA') \subset \phi(tA) \subset o(t)B.$$
2. Let $B \subset U$ be a neighbourhood of $\vec{0}_U$. Because Λ is continuous and $\Lambda(\vec{0}_W) = \vec{0}_U$, we see that $A' = \Lambda^{-1}(B)$ is a neighbourhood of $\vec{0}_W$. Since ϕ is tangent to $\vec{0}$, there exists a neighbourhood A'' of $\vec{0}_V$, such that $\phi(tA'') \subset o(t)A'$ for some function that satisfies $o(t)/t \to 0$ as $t \to 0$. We then have
$$\Lambda \circ \phi(tA'') \subset \Lambda o(t)A' = o(t)B.$$
3. Let $B \subset W$ be a neighbourhood of $\vec{0}_W$. By Lemma 3.6, there exists a balanced sub-neighbourhood $B' \subset B$ of $\vec{0}_W$, such that $aB' + bB' \subset B$ for any scalars $|a| \leq 1$ and $|b| \leq 1$. Because ϕ, ψ are both tangent to $\vec{0}$, there exist neighbourhoods $A, A' \subset V$ of $\vec{0}_V$, such that $\phi(tA) \subset o(t)B'$ and $\psi(tA') \subset o'(t)B'$ for some functions satisfying $\lim_{t\to 0} o(t)/t = \lim_{t\to 0} o'(t)/t = 0$. Now let $A'' = A \cap A'$ and note that A'' is also a neighbourhood of $\vec{0}_V$. Then, if we also let $o''(t) = 2\max\{o(t), o'(t)\}$, it follows that
$$\{\phi(\vec{x}) + \psi(\vec{x}) : \vec{x} \in A''\} \subset \phi(tA) + \psi(tA') \subset o(t)B' + o'(t)B' \subset o''(t)B.$$
Here, we made use of the fact that $A \subset B \implies f(A) \subset f(B)$ for any map f and sets A, B. We also used $A_1 \subset B_1, A_2 \subset B_2 \implies A_1 + A_2 \subset B_1 + B_2$ for any subsets A_1, A_2, B_1, B_2 of a vector space, which is easy to verify. □

Next, we prove the chain rule for differentiable functions on topological vector spaces. We only need the specific case of differentiability at $\vec{0}$, which simplifies matters.

Lemma 3.9 (Chain rule)**.** *Let* U, V, W *be topological vector spaces and let* $g\colon U \to V$ *and* $f\colon V \to W$ *be two functions that are differentiable* $\vec{0}_U$ *and* $\vec{0}_V$*, respectively. Assume also that* $g(\vec{0}_U) = \vec{0}_V$ *and* $f(\vec{0}_V) = \vec{0}_W$*. Then, the composition* $f \circ g$ *is also differentiable at* $\vec{0}_U$ *and the derivative is the composition* $M\Lambda$ *of linear maps, where* M *is the derivative of* f *at* $\vec{0}_V$ *and* Λ *the derivative of* g *at* $\vec{0}_U$*.*

Proof. We expand f and g into their familiar derivative forms:

$$\begin{aligned}(f \circ g)(\vec{x}) &= f\left(\Lambda\vec{x} + \phi(\vec{x})\right) \\ &= M\Lambda\vec{x} + M\phi(\vec{x}) + \psi\left(\Lambda\vec{x} + \phi(\vec{x})\right) \\ &\equiv M\Lambda\vec{x} + \rho(\vec{x}),\end{aligned}$$

where Λ, M are continuous linear maps and ϕ, ψ are tangent to $\vec{0}$. Composition preserves continuity and linearity, so $M\Lambda$ is also a continuous linear map. We need only show that $\rho(\vec{x})$ is tangent to $\vec{0}$. To do this, observe that $\Lambda\vec{x}+\phi(\vec{x})$ is continuous and that $\Lambda\vec{0}+\phi(\vec{0}) = \vec{0}$. It is also easy to confirm that for every neighbourhood B of $\vec{0}_V$, there exists a neighbourhood A of $\vec{0}_U$, such that $\Lambda(tA) + \phi(tA) \subset tB$. The result then follows if we apply parts 1, 2, and 3 of Lemma 3.8. □

Unlike the chain rule, most results from calculus in $\mathbb{R}^n$ do not easily generalise to arbitrary topological vector spaces [16, 34]. However, things become easier when we restrict ourselves to normed vector spaces. In light of our discussion of affine spaces in section 3.1, it is interesting to note that these results can be extended without much difficulty to affine spaces where the associated vector space is normed [58], but we will not do that here. Note that every normed vector space V with norm $\|\cdot\|$ is associated with a topological vector space, where the topology is induced by the metric $d(\vec{x},\vec{y}) = \|\vec{x}-\vec{y}\|$. Indeed, it is possible to show that such a vector space satisfies Definition 3.11 using the ϵ, δ-definition of continuity and the triangle inequality. We now define tangency to $\vec{0}$ in normed vector spaces as follows.

Definition 3.18. Let V and W be normed vector spaces. If a function $\phi\colon V \to W$ satisfies $\phi(\vec{0}_V) = \vec{0}_W$, and the domain of ϕ is a neighbourhood of $\vec{0}_V$, and

$$\lim_{\vec{x}\to\vec{0}_V} \frac{\|\phi(\vec{x})\|}{\|\vec{x}\|} = 0,$$

then we say that ϕ is *tangent to* $\vec{0}$.

It is not difficult to show that this is equivalent to Definition 3.16. For completeness, we prove this in the following lemma.

Lemma 3.10. *Let V and W be normed vector spaces. The map $\phi\colon V \to W$ is tangent to $\vec{0}$ according to Definition 3.18 if and only if it is tangent to $\vec{0}$ according to Definition 3.16.*

Proof. First assume Definition 3.18. We then know that $\phi(\vec{0}_V) = \vec{0}_W$ and that for any $\epsilon > 0$, there exists a $\delta > 0$, such that for any $\vec{x}$ in the domain of ϕ, we have

$$0 < \|t\vec{x}\| < \delta \quad\Longrightarrow\quad \frac{\|\phi(t\vec{x})\|}{\|t\vec{x}\|} < \epsilon.$$

We can write this as $\|\phi(t\vec{x})\| \leq o(t)\|\vec{x}\| < \epsilon\|t\vec{x}\|$ for some function $o(t)$ for which there exists an $\eta > 0$ so that $o(t)/t < \epsilon$ whenever $0 < |t| < \eta$. In other words, $\lim_{t\to 0} o(t)/t = 0$. Now, let B be a neighbourhood of $\vec{0}_W$, so that there exists an open ball $B_\epsilon(\vec{0}_W)$ with radius ϵ and centre $\vec{0}_W$, entirely contained in B. We also let $A = B_\epsilon(\vec{0}_V)$ be the open ball with radius ϵ and centre $\vec{0}_V$, contained in V. Then, for any $\vec{y} \in A$, we have $\|\vec{y}\| < \epsilon$, so that $\|\phi(t\vec{y})\| \leq o(t)\|\vec{y}\| < o(t)\epsilon$, which implies that $\phi(t\vec{y}) \in o(t)B$.

Next, we assume Definition 3.16. It is clear that $\phi(\vec{0}_V) = \vec{0}_W$. Let $\epsilon > 0$ and define the open ball $B = B_\epsilon(\vec{0}_W)$ with centre $\vec{0}_W$. Then, there exists a neighbourhood A of $\vec{0}_V$, such that $\phi(tA) \subset o(t)B$. Clearly, tA is contained in the domain of ϕ, which makes the domain a neighbourhood of $\vec{0}_V$. Moreover, because A is a neighbourhood of $\vec{0}_V$, there exists an open ball $B_\delta(\vec{0}_V)$, so that for any $\vec{x} \in V$, it is true that $\|t\vec{x}\| < \delta$ implies that $t\vec{x} \in tA$. Then also, $\phi(t\vec{x}) \in o(t)B$, so $\|\phi(t\vec{x})\| < o(t)\epsilon$. In other words, for arbitrary $\vec{y}$ in the domain of ϕ,

$$\lim_{\vec{y}\to\vec{0}_V} \frac{\|\phi(\vec{y})\|}{\|\vec{y}\|} = \lim_{t\to 0} \frac{\|\phi(t\vec{x})\|}{\|t\vec{x}\|} < \lim_{t\to 0} \frac{o(t)}{t}\frac{\epsilon}{\|\vec{x}\|} = 0. \qquad \square$$

With this definition of tangency to $\vec{0}$ in place, the definition of differentiability carries over from the arbitrary topological vector space case without any alterations. We can now prove the following inverse mapping theorem [58].

Lemma 3.11 (Inverse function theorem). *Let $E, F \subset V$ be neighbourhoods of $\vec{0}$ in a normed vector space V. The inverse f^{-1} of a homeomorphism $f\colon E \to F$ is differentiable at $\vec{0}$ if:*

1. *$f(\vec{0}) = \vec{0}$,*
2. *f is differentiable at $\vec{0}$,*
3. *the derivative Λ of f at $\vec{0}$ is a homeomorphism.*

Proof. By the chain rule (Lemma 3.9), it is sufficient to prove the result for $g = \Lambda^{-1} \circ f$. Now, because g is differentiable at $\vec{0}$, we have for $\vec{x} \in E$ that

$$g(\vec{x}) = \vec{x} + \phi(\vec{x}),$$

where ϕ is tangent to $\vec{0}$. Letting $\vec{y} = g(\vec{x})$, we write the inverse as

$$g^{-1}(\vec{y}) = \vec{y} - \phi\left(g^{-1}(\vec{y})\right).$$

We must show that the function $\psi(\vec{y}) \equiv -\phi\left(g^{-1}(\vec{y})\right)$ is tangent to $\vec{0}$. Using Definition 3.18, this amounts to showing that the domain of ψ is a neighbourhood of $\vec{0}$, that $\psi(\vec{0}) = \vec{0}$, and

$$\lim_{\vec{y}\to\vec{0}} \frac{\|\psi(\vec{y})\|}{\|\vec{y}\|} = \lim_{\vec{y}\to\vec{0}} \frac{\|\vec{y} - g^{-1}(\vec{y})\|}{\|\vec{y}\|} = 0.$$

By assumption, the first two conditions are already satisfied, so we proceed with proving the limit. Because ϕ is tangent to $\vec{0}$, we know that for any $\epsilon > 0$, there exists a $\delta > 0$, so that

$$0 < \|\vec{x}\| < \delta \quad \Longrightarrow \quad \frac{\|\phi(\vec{x})\|}{\|\vec{x}\|} = \frac{\|g(\vec{x}) - \vec{x}\|}{\|\vec{x}\|} < \tfrac{1}{2}\epsilon.$$

By the reverse triangle inequality, this means that $\|\vec{x}\| - \|g(\vec{x})\| < \frac{1}{2}\epsilon\|\vec{x}\| < \frac{1}{2}\|\vec{x}\|$. In fact, we have that $\frac{1}{2}\|\vec{x}\| < \|g(\vec{x})\|$. But this means that

$$0 < \|\vec{x}\| < \delta \quad \Longrightarrow \quad \|g^{-1}(\vec{y}) - \vec{y}\| = \|\vec{x} - g(\vec{x})\| < \tfrac{1}{2}\epsilon\|\vec{x}\| < \epsilon\|g(\vec{x})\| = \epsilon\|\vec{y}\|,$$

which is what we needed to show. □

Corollary 3.12. *Let $E, F \subset V$ be neighbourhoods of $\vec{0}$ in a finite-dimensional normed vector space V. The inverse f^{-1} of a homeomorphism $f\colon E \to F$ is differentiable at $\vec{0}$ if $f(\vec{0}) = \vec{0}$ and f is differentiable at $\vec{0}$.*

Proof. In a finite-dimensional normed (or topological) vector space, each linear map is continuous. Therefore, we can forget about condition 3 in Lemma 3.11. □

3.3 Linearisation Lemma

Let $\rho\colon G \to \tilde{G}$ be a non-linear realisation of a locally compact Hausdorff topological group G as a topological transformation group $\tilde{G}$ of a topological space X. As mentioned at the end of section 3.1, the proof of the Linearisation Lemma makes use of group integration, which requires that G is locally compact and Hausdorff. Nevertheless, the Affine Linearisation Lemma (Lemma 3.4) demonstrates that similar results are possible for groups that do not have these properties. We have only considered one specific example thus far: the complex scalar model, for which $G = \mathrm{U}(1)$. In this case, we could already apply the Affine Linearisation Lemma. However, because $\mathrm{U}(1)$ is a compact group, we can also apply the general Linearisation Lemma. The goal of this section is to state and prove the Linearisation Lemma and several related results.

Linearisation also requires that X is locally homeomorphic to a topological vector space V, at least in some neighbourhood $E \subset X$ of the vacuum $\Omega \in X$, where the linear approximation is accurate. We want our results to be independent of the particular homeomorphism $E \to V$ that we choose. Hence, we use a collection of suitable coordinate maps, which we define by analogy with the chart definition of manifolds.

Definition 3.19. A *collection of suitable coordinate maps* is a set of homeomorphisms $\phi\colon E \to V$, such that the transition functions $\phi \circ \psi^{-1}$, where ϕ and ψ are any two suitable coordinate maps, satisfy

1. $\phi \circ \psi^{-1}(\vec{0}) = \vec{0}$,

2. $\phi \circ \psi^{-1}$ is differentiable at $\vec{0}$.

We shall see that if our results hold for one suitable coordinate map, then they hold for the entire collection. Suppose now that we are given a suitable coordinate map $\phi\colon E \to V$. Then, the restriction to E of the non-linear realisation ρ on X naturally induces a non-linear realisation $\tilde{\rho}$ on V through

$$\tilde{\rho}(g) = \phi \circ \rho(g)|_E \circ \phi^{-1},$$

which is well-defined if $\rho(g)(E) \subset E$. This condition is necessary, because ϕ is only defined on E. If the condition is not satisfied, then there exists an $\vec{x} \in V$, such that $\rho(g) \circ \phi^{-1}(\vec{x}) \notin E$, in which case $\phi \circ \rho(g) \circ \phi^{-1}(\vec{x})$ does not exist. In the trivial case that $X = E$, this condition is satisfied by definition. The goal of this section is to derive a linear representation on V. This can be done by constructing a coordinate map ϕ, such that the transformations $\tilde{\rho}(g)$ are linear. It turns out that the concept of differentiability is crucial.

Definition 3.20 (Differentiability)**.** Let ϕ be a suitable coordinate map. We say that a group element $g \in G$ is *ϕ-differentiable* at $\vec{x} \in V$ if its realisation $\tilde{\rho}(g)\colon V \to V$ on the vector space V is differentiable at $\vec{x}$.

Definition 3.21 (Linearisability)**.** Let ϕ be a suitable coordinate map. We say that $g \in G$ is *ϕ-linearisable* if $\tilde{\rho}(g)(\vec{0}) = \vec{0}$ and both g and g^{-1} are ϕ-differentiable at $\vec{0}$.

Lemma 3.13. *Assume that V is a finite-dimensional normed vector space. If $g \in G$ is ϕ-differentiable and $\tilde{\rho}(g)(\vec{0}) = \vec{0}$, then g is ϕ-linearisable.*

Proof. This is a direct consequence of Corollary 3.12. □

The following lemma shows that ϕ-differentiability at $\vec{0}$ and ϕ-linearisability are independent of the suitable coordinate map ϕ. In the context of a given collection of suitable coordinate maps, we can therefore simply speak of differentiability at $\vec{0}$ and linearisability.

Lemma 3.14. *Let ϕ and ψ be two suitable coordinate maps. If a group element $g \in G$ is ϕ-differentiable at $\vec{0}$, then it is ψ-differentiable at $\vec{0}$. If g is ϕ-linearisable, then it is ψ-linearisable.*

Proof. If g is ϕ-differentiable at $\vec{0}$, then the realisation $\phi \circ \rho(g) \circ \phi^{-1}$ on V is differentiable at $\vec{0}$. By the chain rule for topological vector spaces (Lemma 3.9), we also have that

$$\psi \circ \rho(g) \circ \psi^{-1} = \psi \circ \phi^{-1} \circ \phi \circ \rho(g) \circ \phi^{-1} \circ \phi \circ \psi^{-1}$$

is differentiable at $\vec{0}$. Hence, g is ψ-differentiable at $\vec{0}$. For the second part, assume that g is ϕ-linearisable, such that g and g^{-1} are ϕ-differentiable at $\vec{0}$ and $\phi \circ \rho(g) \circ \phi^{-1}(\vec{0}) = \vec{0}$. By the first part, we know that g and g^{-1} are also ψ-differentiable at $\vec{0}$. Furthermore,

$$\psi \circ \rho(g) \circ \psi^{-1}(\vec{0}) = \psi \circ \phi^{-1} \circ \phi \circ \rho(g) \circ \phi^{-1} \circ \phi \circ \psi^{-1}(\vec{0}) = \vec{0}.$$

Hence, g is also ψ-linearisable. □

We can now prove the following without reference to a specific coordinate map.

Lemma 3.15. *Let $H \subseteq G$ be the set of linearisable elements $g \in G$. Then H is a subgroup, which we call the* linearisable subgroup.

Proof. Let $g, h \in H$. Then g, h are differentiable at $\vec{0}$ and $\tilde{\rho}(g)(\vec{0}) = \tilde{\rho}(h)(\vec{0}) = \vec{0}$. By definition, g^{-1} and h^{-1} are also differentiable at $\vec{0}$. By the chain rule (Lemma 3.9), it follows that $\tilde{\rho}(gh^{-1}) = \tilde{\rho}(g) \circ \tilde{\rho}(h^{-1})$ and $\left(\tilde{\rho}(gh^{-1})\right)^{-1} = \tilde{\rho}(h) \circ \tilde{\rho}(g^{-1})$ are differentiable at $\vec{0}$. Finally, $\tilde{\rho}(gh^{-1})(\vec{0}) = \tilde{\rho}(g) \circ (\tilde{\rho}(h))^{-1}(\vec{0}) = \tilde{\rho}(g)(\vec{0}) = \vec{0}$, so gh^{-1} is linearisable. We conclude that $H \subseteq G$ is closed under multiplication and inverses, so H is a subgroup. □

Lemma 3.16. *Let $K \subseteq G$ be the set of elements g, so that both g and g^{-1} are differentiable at $\vec{0}$. Then K is a subgroup, called the* differentiable subgroup. *If V is a finite-dimensional normed vector space, then K is the set of elements that are differentiable at $\vec{0}$.*

Proof. Essentially the same argument can be used, in combination with Corollary 3.12. □

We can say a lot more about the subgroups H and K. First of all, in many applications, it is true that $K = G$, such as in Example 3.1. However, this is not necessary for the Linearisation Lemma. Secondly, it is possible to show that H is a closed subset of G, as we will see. Thirdly, if we assume that $K = G$ and that the vector space V is finite-dimensional and therefore locally compact, then there is a nice characterisation of the coset space G/H. Specifically, we will see that G/H is homeomorphic to the orbit $G \cdot \Omega$ of the vacuum Ω. This homeomorphism can also be used to test whether H is nontrivial, without explicitly constructing the linearisable subgroup. Most importantly, the homeomorphism allows us to prove an important corollary of the Linearisation Lemma. Before doing all that, let us first prove the Linearisation Lemma itself, which also holds in the case that V is infinite-dimensional.

The proof of the Linearisation Lemma makes use of group integration. Specifically, we will use a right-invariant Haar measure μ on the subsets of G. Right-invariance means that $\mu(E) = \mu(Eg)$ for all $g \in G$ and measurable subsets $E \subset G$. Let us say a few things about Haar measures in general. First of all, a Haar measure μ is left-invariant if and only if the measure μ', defined by $\mu'(A) = \mu\left(A^{-1}\right)$ is right-invariant. According to Haar's theorem [38, 79], there exists a left-invariant Haar measure for every locally compact Hausdorff topological group. Furthermore, this measure is unique up to a positive multiplicative constant. In our case, we may choose to normalise the measure, such that $\int_H \mathrm{d}\mu(g) = 1$, but this is not necessary for the proof.

We briefly sketch how a Haar measure might be constructed. For details, refer to [19, 33]. Let S and T be two subsets of G, where S is compact and the interior $\operatorname{int}(T)$ of T is non-empty. We write the left translate of $\operatorname{int}(T)$ by g as $g\operatorname{int}(T) = \{gx \in G : x \in \operatorname{int}(T)\}$. Then, $\{g\operatorname{int}(T) : g \in G\}$ is an open cover of S. Because S is compact, there must be a finite sequence $(g_i)_{i=1}^n$, such that $\{g_i\operatorname{int}(T) : 1 \leq i \leq n\}$ covers S. We denote the smallest integer n for which this is possible by $(S : T)$. Let $\mathcal{S}$ be the collection of compact subsets of G. Also let $\mathcal{U}$ be the collection of open neighbourhoods of the identity $e \in G$. Using the local compactness of G, we see that there exists a compact subset S_0 of G for which $\operatorname{int}(S_0) \neq \varnothing$. Now, for each $U \in \mathcal{U}$, we define the map $\mu_U : \mathcal{S} \to \mathbb{R}$ by

$$\mu_U(S) = \frac{(S : U)}{(S_0 : U)}.$$

Because S_0 is non-empty, $(S_0 : U) > 0$, so this is well-defined. Now, for any $S, T \in \mathcal{S}$, it is easy to see that $\mu_U(S \cup T) \leq \mu_U(S) + \mu_U(T)$, as the union of an open cover of S and an open cover of T is also an open cover of $S \cup T$. Moreover, it can actually be shown that if $S \cap T = \varnothing$ and if U is small enough, then $\mu_U(S \cup T) = \mu_U(S) + \mu_U(T)$. In order to obtain a measure that is countably additive on all subsets, we should thus take a limit of μ_U as U becomes smaller. This can be done as follows. First, we note that $0 \leq \mu_U(S) \leq (S : S_0)$. Using this, we may consider each μ_U as a point in $Y = \prod_{S \in \mathcal{S}} [0, (S : S_0)]$. Now, for each $V \in \mathcal{U}$, define the family $C(V) = Y \backslash \{\mu_U : U \in \mathcal{U}, U \subset V\}$. The approach is then to show that the intersection $\bigcap_{V \in \mathcal{U}} C(V)$ is non-empty, so that we can pick some $\mu \in \bigcap_{V \in \mathcal{U}} C(V)$. It can then be shown that such a μ, when extended to all subsets of G, is countably additive. This is the unique left-invariant Haar measure for G, up to a multiplicative constant.

Theorem 3.17 (Linearisation Lemma)**.** *Let $H \leq G$ be the linearisable subgroup of G. If H is compact, then there exists a suitable coordinate map $\bar{\phi}$, such that the induced representation $\bar{\rho}(g) = \bar{\phi} \circ \rho(g) \circ \bar{\phi}^{-1}$ on V, is a linear representation when restricted to H.*

Proof. Having accounted for differentiability and made the necessary adjustments for topological vector spaces, the proof of Bochner [9] and Coleman, Wess, and Zumino [21] mostly carries over. Thus, let $\phi \colon E \to V$ be an arbitrary suitable coordinate map, and let $\tilde{f}_g = \tilde{\rho}(g) = \phi \circ \rho(g) \circ \phi^{-1}$ be the induced representation on V, for any $g \in G$. Now, let $h \in H$ be an arbitrary linearisable group element. Then $\tilde{f}_h$ is differentiable at $\vec{0}$ and satisfies $\tilde{f}_h(\vec{0}) = \vec{0}$. Hence, we can write

$$\tilde{f}_h(\vec{x}) = \Lambda_h \vec{x} + \chi(\vec{x}),$$

where $\chi(\vec{x})$ is tangent to $\vec{0}$. Now define a homeomorphism $\psi\colon V \to V$ by

$$\vec{y} = \psi(\vec{x}) = \int_H \mathrm{d}\mu(g)\ \Lambda_g^{-1}\tilde{f}_g(\vec{x}),$$

where μ is a right-invariant Haar measure on G, such that $\mu(gg_0) = \mu(g)$. As noted above, we may choose to normalise the measure, such that $\int_H \mathrm{d}\mu(g) = 1$, but this is not necessary for the proof. Now, it is easy to see that $\bar{\phi} \equiv \psi \circ \phi$ is a suitable coordinate map. Then, for constant $h \in H$, we obtain

$$\begin{aligned}
\psi\left(\tilde{f}_h(\vec{x})\right) &= \int_H \mathrm{d}\mu(g)\ \Lambda_g^{-1}\tilde{f}_g(\tilde{f}_h(\vec{x})) \\
&= \int_H \mathrm{d}\mu(g)\ \Lambda_h\Lambda_h^{-1}\Lambda_g^{-1}\tilde{f}_g(\tilde{f}_h(\vec{x})) \\
&= \Lambda_h \int_H \mathrm{d}\mu(gh)\ \Lambda_{gh}^{-1}\tilde{f}_{gh}(\vec{x}) \\
&= \Lambda_h\psi(\vec{x}).
\end{aligned}$$

This implies that $\psi \circ \tilde{f}_h \circ \psi^{-1}(\vec{y}) = \Lambda_h\vec{y}$. Observe that Λ_h is a linear homeomorphism, such that $\Lambda_h \in \mathrm{Aut}(V)$. Hence, the homomorphism $\bar{\rho}\colon G \to \mathrm{Aut}(V)$, defined by $\bar{\rho} : g \mapsto \bar{\phi} \circ \rho(g) \circ \bar{\phi}^{-1}$ restricted to H is a linear representation on V. Of course, this is also the representation induced by the suitable coordinate map $\bar{\phi}$. □

Let us now consider two examples.

Example 3.2. We return to Example 3.1. Recall that we had endowed the subset $X = (0, \infty) \subset \mathbb{R}$ with subspace topology. We also introduced $G = (X, *, 1)$ as a topological group under multiplication. Furthermore, we defined a non-linear realisation of G on X by

$$\rho(g)(x) = x^g.$$

Observe that the functions $\rho(g)$ are differentiable for all $g \in G$. Moreover, the entire group fixes the point $1 \in X$. It follows that the entire group is linearisable. Unfortunately, X is not compact, so the Linearisation Lemma cannot be applied. □

Example 3.3. Let S^1 be the circle group. The elements of S^1 can be written as $g = e^{2i\pi a}$ with $a \in [0, 1) \subset \mathbb{R}$. Next, define the set $X = [a+bi \in \mathbb{C} : 0 \leq b < 2\pi]$. Then, the restriction of the exponential map to X is injective and surjective onto its image $\mathbb{C}\backslash\{0\}$. Its inverse is given by the principal branch of the complex logarithm, which is analytic everywhere on the complex plane except on the negative real axis [67]. With this in mind, we attempt to define a non-linear realisation of S^1 on $\mathbb{C}\backslash\{0\}$ by

$$\rho(g)(z) = z^g = \exp\left(g\log(z)\right),$$

where we use the principal branch of the logarithm. Note first of all that these functions fix the point $1 \in \mathbb{C}\backslash\{0\}$. By the preceding arguments, the map $\rho(g)$ is also a bijection. Furthermore, $\rho(g)$ and its inverse are both continuous everywhere on the complex plane, except on the negative real axis. This means that $\rho(g)$ is not a homeomorphism. Still, it is instructive to ignore this problem for the moment. After all, we are only interested in the local behaviour of these functions around the fixed point 1.

With the understanding that $X = \mathbb{C}\backslash\{0\}$ is a topological space and $V = \mathbb{C}\backslash\{-1\}$ a topological vector space, we let the map $\phi : x \mapsto x - 1$ be a suitable coordinate map between X and V. We pick the point $\vec{0} = \phi^{-1}(1)$ as our vacuum. We find that the induced realisation on V can be written as

$$\begin{aligned}\tilde{\rho}(g)(z) &= \rho(g)(z+1) - 1 \\ &= \rho(g)(1) - 1 + \left.\frac{\mathrm{d}\rho(g)(z)}{\mathrm{d}z}\right|_{z=1} z + o\left(z^2\right) \\ &= 0 + gz + o\left(z^2\right).\end{aligned}$$

Finally, noting that S^1 is compact, we apply the Linearisation Lemma to conclude that $\tilde{\rho}(g) = g$ is a linear representation on V, with the caveat that the functions $\rho(g)$ suffer a discontinuity along the negative real axis. □

As mentioned above, the Linearisation Lemma holds even for infinite-dimensional vector spaces. Nevertheless, the Linearisation Lemma only implies the existence of a linear representation of H, but not of the entire group G. If we assume that V is finite-dimensional, then it is possible to show that there exists such a representation of G on a subspace of V. This is an extension of a similar result in [21]. Before we can do that, we first need to prove a number of lemmas about topological groups and topological spaces in general. In the meantime, we will be able to give the characterisation of the coset space G/H, alluded to above. Let us begin with two lemmas on topological groups.

Lemma 3.18. *Let G be a second-countable topological group. Then, given a neighbourhood U of the identity $e \in G$, there exists a sequence $(h_n) \subset G$, such that*

$$G = \bigcup_n h_n U. \tag{3.6}$$

Proof. Because U is a neighbourhood, by definition there exists an open set $Y \subset U$, with $e \in Y$. Let (B_n) be a countable open basis for G. Because G is second-countable, it is also separable. In other words, there exists a sequence (g_n), such that every non-empty open subset of G contains at least one element of the sequence (g_n). We will show that the sequence $(h_n) = (g_n^{-1})$ satisfies (3.6). To do this, we let $x \in G$ be arbitrary and we proceed by showing that $x \in h_j U$ for some $h_j \in (h_n)$.

Let $\mu\colon G \times G \to G$ denote the continuous multiplication map. Then $\mu^{-1}(Y) \subset G \times G$ is open by continuity. We note that $\mu(x^{-1}, x) = e \in Y$. Hence, by the product topology, there exist open subsets V, W, such that $V \times W \subset \mu^{-1}(Y)$ and $(x^{-1}, x) \in V \times W$. Now, pick an element $g_j \in (g_n)$, so that $g_j \in V$. This implies that $\{g_j\} \times W \subset V \times W$, so that $g_j W \subset Y$. Putting everything together, we see that $x \in W \subset g_j^{-1} Y = h_j Y \subset h_j U$, as desired. □

Our second lemma on topological groups is very reminiscent of part 2 of Lemma 3.6 concerning topological vector spaces. This is of course not surprising, because a vector space is also a group under addition. Again, the following is a standard result [57, 76].

Lemma 3.19. *Let G be a topological group and U a neighbourhood of the identity $e \in G$. Then, U contains an open neighbourhood $W \subset U$ of e, satisfying $W^2 \subset U$ and $W = W^{-1}$.*

Proof. Because U is a neighbourhood of e, there exists an open set $Y \subset U$ with $e \in Y$. Let $\mu\colon G \times G \to G$ denote the multiplication map. Then, $\mu^{-1}(Y)$ is open by continuity and $(e, e) \in \mu^{-1}(Y)$. By the product topology, there exist open sets V_1, V_2 with $(e, e) \in V_1 \times V_2$. Now, let $V = V_1 \cap V_2$. Then, $V \times V \subset \mu^{-1}(Y)$, so $VV \subset U$. Next, let $W = V \cap V^{-1}$, so that $W = W^{-1}$. Clearly, we still have that $WW \subset U$. Finally, observe that W is an open neighbourhood of e. □

Now, let us prove two lemmas about locally compact Hausdorff spaces. The following lemma is based on [41].

Lemma 3.20. *Let X be a locally compact Hausdorff topological space. For any open subset $U \subset X$ and any point $p \in U$, there exists an open set $V \subset U$ with $p \in V$, such that the closure* $\mathrm{cl}(V) \subset U$.

Proof. We make use of the fact that the space X is regular, which is true because X is Hausdorff and locally compact. Regularity means that, for any closed non-empty subset $C \subset X$ and any point $x \notin C$, there exists an open neighbourhood O_1 of x and an open set O_2 with $C \subset O_2$, such that $O_1 \cap O_2 = \varnothing$. Now let $U \subset X$ be an open set and $p \in U$ some point. Find an open subset $O \subset U$ with $p \in O$. If $O = X$, then we are done, because X is closed and open. Otherwise, O^{c} is closed and non-empty and $p \notin O^{\mathrm{c}}$. Hence, by regularity, there exist open sets O_1, O_2, such that

$$p \in O_1, \qquad O^{\mathrm{c}} \subset O_2, \qquad O_1 \cap O_2 = \varnothing.$$

The last relation implies that $O_1 \subset O_2^{\mathrm{c}}$. Furthermore, because O_2^{c} is closed, we also have $\mathrm{cl}(O_1) \subset O_2^{\mathrm{c}}$. From the second relation, we find $O_2^{\mathrm{c}} \subset O$. Letting $V = O_1$ and putting everything together, we find

$$p \in V \subset \mathrm{cl}(V) = \mathrm{cl}(O_1) \subset O_2^{\mathrm{c}} \subset O \subset U.$$ □

The second result is also known as the Category Theorem [40, p. 110].

Lemma 3.21. *Let X be a locally compact Hausdorff topological space that can be written as a countable union of closed subsets (X_n),*

$$X = \bigcup_n X_n. \tag{3.7}$$

Then at least one X_n contains an open subset of X.

Proof. We proceed by contradiction. Suppose that none of the (X_n) contain an open subset. By local compactness, we can find an open set $U_1 \subset X$, such that its closure $\mathrm{cl}(U_1)$ is compact. We shall now iteratively "subtract" the closed subsets X_n from U_1 and consider what is left. First, we choose a point $a_1 \in U_1 \backslash X_1$, which is possible because we assumed that X_1 does not contain any open sets. By Lemma 3.20, we can also find an open neighbourhood U_2 of a_1, such that $\mathrm{cl}(U_2) \subset U_1 \backslash X_1$. Note that $\mathrm{cl}(U_2)$ is compact, because $\mathrm{cl}(U_1)$ is compact. In the same way, we pick $a_2 \in U_2 \backslash X_2$ and a neighbourhood U_3 of a_2, such that $\mathrm{cl}(U_3) \subset U_2 \backslash X_2$. Proceeding in this way, we find a decreasing sequence $\mathrm{cl}(U_1), \mathrm{cl}(U_2), \ldots$ of compact non-empty sets. Hence, there exists a point $b \in \bigcap_n \mathrm{cl}(U_n)$. However, this implies that $b \notin X_n$ for all n, which is in contradiction with (3.7). □

Now, we can finally prove our statement about the coset space G/H. This lemma corresponds to Theorem 3.2 of [40, p. 111].

Lemma 3.22. *Let H be the linearisable subgroup and K the differentiable subgroup of G. We assume that $G = K$, so that all transformations are differentiable. Denote with $G \cdot \Omega$ the orbit of $\Omega \equiv \phi^{-1}(\vec{0})$, where ϕ is any suitable coordinate map. If V is finite-dimensional and G is secound-countable, then the map $\Phi : gH \mapsto \rho(g)(\Omega)$ is a homeomorphism $G/H \xrightarrow{\sim} G \cdot \Omega$.*

Proof. First of all, let us recall that the coordinate map ϕ is a homeomorphism $E \xrightarrow{\sim} V$ from E to the topological vector space V. We proved in section 3.2 that every topological vector space is Hausdorff. It is also true that finite-dimensional topological vector spaces are locally compact. Through ϕ, these same properties hold for E. Furthermore, because singletons in V are closed, we also note that $\{\Omega\}$ is closed. Hence, because the map $\Psi : g \mapsto \rho(g)(\Omega)$ is continuous, we see that $H = \Psi^{-1}(\Omega)$ is closed[9]. It is easy to see that the map $\Phi : gH \mapsto \rho(g)(\Omega)$ is a bijection, because $\rho(h)(\Omega) = \Omega$ if and only if $h \in H$. We also see that Φ is continuous, because the canonical map $\pi \colon G \to G/H$ is continuous. To establish that Φ is a homeomorphism, it is therefore sufficient to show that Ψ is an open map. We do this by showing that each point in $\Psi(U)$, for every open set $U \subset G$, is an interior point.

[9] Of course, this also holds if V is infinite-dimensional, G not second-countable, and if $G \neq K$.

Thus, let V be an open neighbourhood of some point $g \in G$. Then, $g^{-1}V$ is an open neighbourhood of the identity $e \in G$. By Lemma 3.19, there exists a neighbourhood $U_1 \subset g^{-1}V$ of e, such that $U_1 = U_1^{-1}$ and $U_1^2 \subset g^{-1}V$. By local compactness, there exists a compact neighbourhood W of e. Using Lemma 3.20, we see that there exists an open neighbourhood U_2 of e, with $\mathrm{cl}(U_2) \subset U_1$. Then, $\mathrm{cl}(U_2) \cap W$ is also a neighbourhood of e. By the Hausdorff property, W is closed, so $\mathrm{cl}(U_2) \cap W$ is a closed subset of W. It follows that $U_3 = \mathrm{cl}(U_2) \cap W$ is a compact neighbourhood of e. Finally, let $U = U_3 \cap U_3^{-1}$ be another neighbourhood of e. Evidently, $U = U^{-1}$, $U^2 \subset g^{-1}V$ and U is compact. By Lemma 3.18, there exists a sequence $(h_n) \subset G$, such that $G = \bigcup_n h_n U$. Clearly, we also have

$$G \cdot \Omega = \bigcup_n h_n U \cdot \Omega.$$

The continuous map $\Psi : g \mapsto \rho(g)(\Omega)$ preserves compactness, so each of the $h_n U \cdot \Omega$ is compact and therefore closed, by the Hausdorff property of E. It follows, by Lemma 3.21, that there is some n, such that $h_n U \cdot \Omega$ contains an open subset of $G \cdot \Omega$. By continuity, the same is true for $U \cdot \Omega$. This set therefore contains some interior point $u \cdot \Omega \in U \cdot \Omega$. It follows that Ω is an interior point of $u^{-1} U \cdot \Omega \subset U^2 \cdot \Omega \subset g^{-1} V \cdot \Omega$. Hence, $g \cdot \Omega$ is an interior point of $V \cdot \Omega = \Psi(V)$. Because $g \in V$ is an arbitrary point in an arbitrary open subset $V \subset G$, we have shown that Ψ is open, which concludes our proof. □

As promised, we have shown that G/H is homeomorphic to the orbit $G \cdot \Omega$. We immediately see that H is trivial if and only if $G \cong G \cdot \Omega$. In the case that the group action of G on X is transitive, such that $G \cdot \Omega = X$, this means that H is trivial if and only if $G \cong X$. With all of this out of the way, we can now prove that there exists a linear representation of G on a subspace of V. This is a generalisation of [21].

Corollary 3.23. *Let H be the linearisable subgroup and K the differentiable subgroup of G. We assume that $G = K$, so that all transformations are differentiable. If V is finite-dimensional, G is second-countable, and H is compact, then there exists a suitable coordinate map ϕ, such that if we let $W \subset V$ denote the linear subspace spanned by the orbit $\phi(G \cdot \Omega)$, where $\Omega = \phi^{-1}(\vec{0})$, and if we write each vector $\vec{x} \in V$ as*

$$\vec{x} = (\vec{\pi}, \vec{\sigma}), \qquad \vec{\pi} \in W, \quad \vec{\sigma} \in V/W,$$

then there exists a continuous map $\psi \colon \phi(G \cdot \Omega) \times V/W \to V$, such that the induced realisation $\tilde{\rho}(g)$ satisfies

$$\tilde{\rho}(g) \circ \psi(\vec{\pi}, \vec{\sigma}) = \psi\left(\rho'(g)(\vec{\pi}), \Lambda_g \vec{\sigma}\right),$$

where $\rho'(g) \colon G \to \tilde{G}$ is a non-linear realisation of G on $\phi(G \cdot \Omega)$ and $\Lambda_g \in \mathrm{Aut}(V/W)$ is a linear representation of G on V/W.

Proof. By Theorem 3.17, there exists a suitable coordinate map $\phi\colon E \to V$, such that the induced realisation $\tilde{\rho}(g)$ is a linear representation on V, when restricted to H. By Lemma 3.22, the orbit $G \cdot \Omega$ of $\Omega = \phi^{-1}(\vec{0})$ is homeomorphic to the coset space G/H. In fact, we have the homeomorphism $G/H \to G \cdot \Omega$ defined by $gH \mapsto \rho(g)(\Omega)$. We then also have the homeomorphism $\zeta\colon G/H \to \phi(G \cdot \Omega)$, defined by

$$\zeta(gH) = \phi\left(\rho(g)(\Omega)\right) = \tilde{\rho}(g)(\vec{0}).$$

Now, let $W = \operatorname{span}\{\phi(G \cdot \Omega)\}$ be the smallest linear subspace of V that contains the orbit of $\vec{0}$. It is important to note that $\phi(G \cdot \Omega)$ is a subset of V that is generally not closed under vector addition or scalar multiplication, precisely because we are dealing with a non-linear realisation. Let $\dim(V) = n$ and $\dim(W) = m$. We can erect a set of basis vectors $\{\vec{e}_1, \ldots, \vec{e}_m\}$ of W and $\{\vec{e}_{m+1}, \ldots, \vec{e}_n\}$ of V/W, which together span V. An arbitrary vector $\vec{x} \in V$ is then:

$$\vec{x} = (\vec{\pi}, \vec{\sigma}), \qquad \vec{\pi} \in W, \quad \vec{\sigma} \in V/W.$$

Fixing coset representatives allows us to distinguish a unique transformation $\tilde{\rho}(\zeta^{-1}(\vec{\pi}, \vec{0}))$ for each $\vec{\pi} \in \phi(G \cdot \Omega)$. We use this to define the continuous[10] map $\psi\colon \phi(G \cdot \Omega) \times V/W \to V$ by:

$$\psi(\vec{\pi}, \vec{\sigma}) = \tilde{\rho}(\zeta^{-1}(\vec{\pi}, \vec{0}))(\vec{0}, \vec{\sigma}).$$

We are then able to consider the effect of an arbitrary group element $g \in G$ on an arbitrary vector $\vec{x} \in \phi(G \cdot \Omega) \times V/W$:

$$\begin{aligned} \tilde{\rho}(g) \circ \psi(\vec{\pi}, \vec{\sigma}) &= \tilde{\rho}(g) \circ \tilde{\rho}(\zeta^{-1}(\vec{\pi}, \vec{0}))(\vec{0}, \vec{\sigma}) \\ &= \tilde{\rho}(g\zeta^{-1}(\vec{\pi}, \vec{0}))(\vec{0}, \vec{\sigma}). \end{aligned}$$

Using the coset representatives, there is a unique choice of $a, b \in G$, such that $aH = \zeta^{-1}(\vec{\pi}, \vec{0})$ and $bH = g\zeta^{-1}(\vec{\pi}, \vec{0}) = gaH$. Then, $ga = bh$ for some unique $h \in H$. Hence,

$$\begin{aligned} \tilde{\rho}(g) \circ \psi(\vec{\pi}, \vec{\sigma}) &= \tilde{\rho}(bh)(\vec{0}, \vec{\sigma}) \\ &= \tilde{\rho}(b)\tilde{\rho}(h)(\vec{0}, \vec{\sigma}) \\ &= \tilde{\rho}(b)(\vec{0}, \Lambda_h \vec{\sigma}), \end{aligned}$$

where Λ_h is linear by virtue of Theorem 3.17. We also have $\zeta(bH) = \tilde{\rho}(b)(\vec{0})$, so that

$$\tilde{\rho}(g) \circ \psi(\vec{\pi}, \vec{\sigma}) = \psi(\zeta(bH), \Lambda_h \vec{\sigma}).$$

[10] Recall that homomorphisms between topological groups are continuous group homomorphisms.

It is easy, but tedious, to verify that $\rho'(g)(\vec{\pi}) \equiv \zeta(g\zeta^{-1}(\vec{\pi})) = \zeta(bH)$ is a non-linear realisation of G on $\phi(G \cdot \Omega)$ and that $\rho''(g)(\vec{\sigma}) \equiv \Lambda_{b^{-1}ga}\vec{\sigma} = \Lambda_h\vec{\sigma}$ is a linear representation of G on V/W. We only state the result that

$$\tilde{\rho}(w) \circ \tilde{\rho}(g) \circ \psi(\vec{\pi}, \vec{\sigma}) = \psi(\zeta(wg\zeta^{-1}(\vec{\pi})), \Lambda_{h'}\vec{\sigma}) = \tilde{\rho}(wg) \circ \psi(\vec{\pi}, \vec{\sigma}),$$

where $h' = c^{-1}wga \in H$ and $c \in G$ is the unique choice for $wgaH = cH$. □

Unfortunately, even the generalised version of the Linearisation Lemma, that we have proved in this section, does not help us directly in the study of non-linear realisations of supersymmetry. This is due to the fact that, for spontaneous supersymmetry breaking, the unbroken subgroup H is the Poincaré group $\mathrm{ISO}(3,1) = \mathbb{R}^{3,1} \rtimes \mathrm{SO}(3,1)$, which is not compact because the Lorentz group $\mathrm{SO}(3,1)$ is not compact. A possible resolution lies in Weyl's unitarian trick, which allows one to prove facts about non-compact groups using related compact groups. In this case, we might apply the Linearisation Lemma to the group $\mathrm{SO}(4)$ of rotations in 4-dimensional Euclidean space. Recall from section 3.1 that the tuple of fields $F\colon \mathbb{R}^{3,1} \to X$ mapped flat spacetime into X. To apply Weyl's unitarian trick, we would have to relate F to a similar function $\mathbb{R}^4 \to X$ through a Wick rotation. We will not consider the details of this potential approach here. Fortunately, diagram (3.1) indicates another possible route for studying the Goldstone fields associated with supersymmetry breaking, namely the Non-Linearisation Lemma.

3.4 Non-Linearisation Lemma

In this section, we consider a related problem, in which we wish to eliminate the non-Goldstone fields using a constraint. This corresponds to the route labelled NLL in diagram (3.1). To formulate the problem more precisely, we start with a non-linear realisation $\rho\colon G \to \mathrm{Homeo}(V)$ of a topological group G on an n-dimensional topological vector space V over a topological field K. Given a basis $\{\vec{e}_i\}$ with $i = 1, \ldots, n$, we parametrise each point $\vec{x} \in V$ by a set of coordinates $\{x_i\}$. We may impose a constraint $F(x_1, \ldots, x_n) = 0$ on the coordinates, allowing us to eliminate one coordinate, say $x_n = H(x_1, \ldots, x_{n-1})$. We let W denote the subset on which the constraint is true. The question is: given the non-linear realisation $\rho(g)(x_1, \ldots, x_n)$ on V, how do we define a non-linear realisation $\tilde{\rho}(g)(x_1, \ldots, x_{n-1})$ on W, which doesn't involve the nth coordinate x_n? A possible answer is

$$\tilde{\rho}(g)(\tilde{x}) \equiv \rho(g)(\tilde{x}, H(\tilde{x})), \tag{3.8}$$

where $\tilde{x} = (x_1, \ldots, x_{n-1})$. The Non-Linearisation Lemma asserts that this is indeed a valid non-linear realisation, provided that $\rho(g)(W) \subset W$ for each $g \in G$. In other words, the induced realisation (3.8) is valid if the constraint $F(\vec{x}) = 0$ is invariant. We shall first prove the following lemma, which is a generalisation of a similar result due to Casalbuoni et al. [15]. In what follows, we denote the projection of $\rho(g)(\vec{x})$ onto its ith coordinate by $\rho(g)(\vec{x})_i$.

Lemma 3.24 (Casalbuoni et al.)**.** *Let $F\colon V \to K$ be differentiable at every point in V. Assume that the function $H\colon \mathrm{span}(W) \to K$ is differentiable at every point in $\mathrm{span}(W)$ and satisfies $F(\tilde{x}, H(\tilde{x})) = 0$, so that $x_n = H(\tilde{x})$ solves the constraint $F(\vec{x}) = 0$. Then, for any $g \in G$ and $\vec{x} \in W$,*

$$\rho(g)(\vec{x})_n = H(\rho(g)(\vec{x})_1, \ldots, \rho(g)(\vec{x})_{n-1}),$$

if and only if $\rho(g)(\vec{x}) \in W$.

In other words, the content of this lemma is that the action takes $\vec{x} \in W$ to another point in W if and only if the projection of $\rho(g)(\vec{x})$ onto its nth coordinate yields the same answer as solving for x_n using the projection of $\rho(g)(\vec{x})$ onto its other coordinates.

Proof. Because $F\colon V \to K$ is differentiable, we have

$$F(\vec{x}_0 + \vec{\delta}) = F(\vec{x}_0) + M(\vec{x}_0)\vec{\delta} + \phi(\vec{\delta}), \qquad \forall \vec{x}_0, \vec{\delta} \in V,$$

where $M(\vec{x}_0)$ is a $1 \times n$ linear operator and where ϕ is tangent to $\vec{0}$. We may write the derivative as $M(\vec{x}_0) = (\partial F/\partial x_1, \ldots, \partial F/\partial x_n)|_{\vec{x}=\vec{x}_0}$. We will consistently use this more standard notation in this proof. Now let $\vec{y} \in W$, such that $F(\vec{y}) = 0$ and consider a nearby point $\vec{z} \in W$. We then find that

$$F(\vec{z}) = F(\vec{y} + \vec{z} - \vec{y}) = 0 + \sum_{i=1}^{n} (z_i - y_i) \left.\frac{\partial F}{\partial x_i}\right|_{\vec{x}=\vec{y}} + \phi(\vec{z} - \vec{y}) = 0.$$

It follows that

$$z_n - y_n = -\sum_{i=1}^{n-1} (z_i - y_i) \left.\frac{\partial F/\partial x_i}{\partial F/\partial x_n}\right|_{\vec{x}=\vec{y}} - \left.\frac{\phi(\vec{z} - \vec{y})}{\partial F/\partial x_n}\right|_{\vec{x}=\vec{y}}.$$

Then, neglecting the term proportional to ϕ, which is tangent to $\vec{0}$, we find

$$\left.\frac{\partial H(\tilde{x})}{\partial x_i}\right|_{\vec{x}=\vec{y}} = \left.\frac{\partial x_n}{\partial x_i}\right|_{\vec{x}=\vec{y}} = -\left.\frac{\partial F/\partial x_i}{\partial F/\partial x_n}\right|_{\vec{x}=\vec{y}}. \tag{3.9}$$

Meanwhile, for an arbitrary $\vec{v} \in W$ let $\vec{w} = \rho(g)(\vec{v})$, where $\vec{w}$ is not necessarily contained in W. We also let $\tilde{v} = (v_1, \ldots, v_{n-1})$ and $\tilde{w} = (w_1, \ldots, w_{n-1})$ be the first $n-1$ coordinates of $\vec{v}$ and $\vec{w}$, respectively. Evaluating the constraint function at $\vec{w}$, we find that

$$F(\vec{w}) = F(\vec{v} + \vec{w} - \vec{v}) = 0 + \sum_{i=1}^{n} (w_i - v_i) \left.\frac{\partial F}{\partial x_i}\right|_{\vec{x}=\vec{v}} + \phi(\vec{w} - \vec{v}). \tag{3.10}$$

Using the fact that H is differentiable, we can write

$$H(\tilde{w}) = H(\tilde{v}) + \sum_{i=1}^{n-1}(w_i - v_i) \left.\frac{\partial H}{\partial x_i}\right|_{\vec{x}=\vec{v}} + \psi\left(\tilde{w} - \tilde{v}\right), \tag{3.11}$$

where ψ is tangent to $\vec{0}$. Let us now neglect the terms proportional to ψ and ϕ, which are tangent to $\vec{0}$, and evaluate (3.11) using (3.9) and (3.10):

$$\begin{aligned} H(\tilde{w}) &= H(\tilde{v}) - \sum_{i=1}^{n-1}(w_i - v_i) \left.\frac{\partial F/\partial x_i}{\partial F/\partial x_n}\right|_{\vec{x}=\vec{v}} \\ &= H(\tilde{v}) + w_n - v_n - \left.\frac{F\left(\vec{w}\right)}{\partial F/\partial x_n}\right|_{\vec{x}=\vec{v}} \\ &= w_n - \left.\frac{F\left(\vec{w}\right)}{\partial F/\partial x_n}\right|_{\vec{x}=\vec{v}}. \end{aligned}$$

Hence, it follows that for any $\vec{v} \in W$ and $\vec{w} = \rho(g)(\vec{v})$, we have

$$w_n = H(\tilde{w}) \quad \Longleftrightarrow \quad F\left(\vec{w}\right) = 0 \quad \Longleftrightarrow \quad \vec{w} \in W,$$

as desired. □

Now we are able to prove that (3.8) defines a non-linear realisation.

Theorem 3.25 (Non-Linearisation Lemma)**.** *Let $\rho\colon G \to \mathrm{Homeo}(V)$ be a non-linear realisation of a topological group G on an n-dimensional topological vector space V over a topological field K. Let $F\colon V \to K$ be differentiable at every point in V. Define $W \subset V$ to be the subset of points $\vec{x}$ for which $F(\vec{x}) = 0$. Assume that $H\colon \mathrm{span}(W) \to K$ is differentiable at every point in $\mathrm{span}(W)$ and satisfies $F(\tilde{x}, H(\tilde{x})) = 0$, so that $x_n = H(\tilde{x})$ solves the constraint $F(\vec{x}) = 0$. Then,*

$$\tilde{\rho}(g)(\tilde{x}) \equiv \rho(g)(\tilde{x}, H(\tilde{x})),$$

defines a non-linear realisation on W if and only if $\rho(g)(W) \subset W$.

Proof. It is clear that the transformations $\tilde{\rho}(g)\colon W \to W$ are continuous and that, if $e \in G$ is the identity, then $\tilde{\rho}(e)(\tilde{x}) = \tilde{x}$. It remains to show that for any $g, h \in G$, we have

$$\tilde{\rho}(g) \circ \tilde{\rho}(h) = \tilde{\rho}(gh). \tag{3.12}$$

If this is true, then the transformations are also homeomorphisms. Let $\vec{x} \in W$ and $g, h \in G$. For ease of writing, we define the following quantities:

$$\tilde{x} = \left(x_1, \ldots, x_{n-1}\right), \qquad \vec{y} = \rho(h)\left(\tilde{x}, H(\tilde{x})\right), \qquad \tilde{y} = \left(y_1, \ldots, y_{n-1}\right).$$

Then, we evaluate the left-hand side of (3.12) at the point $\vec{x}$:

$$\begin{aligned}\tilde{\rho}(g) \circ \tilde{\rho}(h)(\tilde{x}) &= \tilde{\rho}(g) \circ \rho(h)(\tilde{x}, H(\tilde{x})) \\ &= \rho(g)\left(\tilde{y}, H(\tilde{y})\right).\end{aligned}$$

We also evaluate the right-hand side of (3.12) at $\vec{x}$:

$$\begin{aligned}\tilde{\rho}(gh)(\tilde{x}) &= \rho(gh)(\tilde{x}, H(\tilde{x})) \\ &= \rho(g) \circ \rho(h)(\tilde{x}, H(\tilde{x})) \\ &= \rho(g)(\tilde{y}, y_n).\end{aligned}$$

The maps $\rho(g)\colon V \to V$ are bijective, so $\tilde{\rho}(g) \circ \tilde{\rho}(h)(\vec{x}) = \tilde{\rho}(gh)(\vec{x})$ if and only if

$$H(\tilde{y}) = y_n.$$

By Lemma 3.24, this is equivalent to $\vec{y} \in W$. All of the above holds for every $\vec{x} \in W$. Hence, the trivial consideration

$$\rho(g)(W) \subset W \Longleftrightarrow \forall \vec{x} \in W : \rho(g)(\vec{x}) \in W,$$

completes the proof. □

Lemma 3.25 gives the necessary and sufficient condition for this approach to be consistent, namely the constraint $F(\vec{x}) = 0$ must be invariant. An example of this approach in action can be found in section 6.2, where a linear representation involving the tuple of fields $\Phi = (\phi, \psi, F)$ induces the non-linear Volkov-Akulov realisation of supersymmetry through a nilpotency constraint $\Phi^2 = 0$, whereby the field $\phi = \psi^2/2F$ is eliminated in favour of the other two fields. The following dictionary illustrates the analogy:

$$\begin{array}{ccc}\Phi & \leftrightarrow & \vec{x}, \\ (\phi, \psi, F) & \leftrightarrow & (x_1, x_2, x_3), \\ \Phi^2 = 0 & \leftrightarrow & F(\vec{x}) = 0, \\ \phi = \psi^2/2F & \leftrightarrow & x_3 = H(x_1, x_2).\end{array}$$

Let's make the following important observation. We have introduced the Non-Linearisation Lemma in section 3.1 as a way of isolating the Goldstone fields. Of course, on the basis of the considerations in this section alone, the Non-Linearisation Lemma can be applied to eliminate any field from the spectrum. However, physical considerations often imply that it makes sense to eliminate one field or another. For example, usually some fields are much more massive than others. In the low-energy limit of such a theory, it makes sense to neglect the kinetic energy of the massive fields. These fields are then stuck at their vacuum expectation value, so that they can be eliminated with an algebraic (non-dynamic) constraint. Of course, this could never be done for massless fields such as Goldstone bosons. This is precisely how the constraint $\Phi^2 = 0$ is used in section 6.2, where the massive scalar partner ϕ of the goldstino ψ is eliminated, leaving the massless goldstino ψ (and an auxiliary field F).

We have spent quite some time studying spontaneous symmetry breaking in general. In the next chapter, we finally start our study of representations of supersymmetry.

4

Mathematics of Supersymmetry

4.1 Superalgebras

In chapter 5, we will study the physics of supersymmetry in detail. This chapter is devoted to developing the mathematical tools necessary for a formal treatment of supersymmetry. Fundamental to theory of supersymmetry is the notion of a superalgebra. We have the following definition [75].

Definition 4.1. A *superalgebra* V over a field K is a vector space with a direct sum decomposition $V = V_0 \oplus V_1$ into even and odd parts, along with a bilinear multiplication operation $\mu\colon V \times V \to V$ satisfying

$$\mu(V_i, V_j) \subset V_{i+j}, \qquad \text{for} \quad i, j \in \{0, 1\}.$$

In addition, we say that V is a *commutative superalgebra* if V satisfies the graded commutativity property

$$\mu(x, y) = -(-1)^{|x||y|}\mu(y, x), \qquad \forall x, y \in V, \tag{4.1}$$

where $|x| = 0$ if $x \in V_0$ and $|x| = 1$ if $x \in V_1$. If the multiplication operation μ is associative, we call V an *associative superalgebra*. Finally, we say that V is a *unital superalgebra* if there exists a multiplicative identity $\mathcal{I}$, for which it then necessarily holds that $\mathcal{I} \in V_0$.

The prototypical example of a commutative superalgebra is the Grassmann algebra (or exterior algebra), which we will use exclusively. An excellent mathematical overview of the use of Grassmann algebras in supersymmetric field theories can be found in [29]. We broadly follow their line of construction, which in turn is based on the work of Rogers on supermanifolds [61].

Definition 4.2. A *Grassmann algebra* $\Lambda(V)$ over a vector space V over a field K is the quotient algebra of the tensor algebra $T(V)$ modulo the ideal I generated by all element of the form $x \otimes x \in T(V)$. The bilinear multiplication map on $\Lambda(V)$ is defined by

$$x \wedge y = x \otimes y \mod I,$$

and called the *exterior product*. Henceforth, we just write $xy = x \wedge y$ if there is no confusion.

Remark 1. In what follows, we always assume that the field K has characteristic 0, and that the Grassmann algebra $\Lambda(V)$ is associative and unital.

Concretely, $\Lambda(V)$ consists of all tensor products $x_1 \otimes x_2 \otimes \cdots \otimes x_k \in T(V)$ with $x_1, \ldots, x_k \in V$, identifying all products of the form $x \otimes x$ with 0. It follows that for $x, y \in V$, the exterior product is anticommutative:

$$0 = (x + y)(x + y) = xx + yx + xy + yy = yx + xy.$$

Of course, this does not necessarily hold for any $x, y \in \Lambda(V)$. Indeed, as a superalgebra, $\Lambda(V)$ consists of both commuting and anticommuting elements. We return to this shortly. First, we introduce the notion of the kth exterior power of V, which is the vector subspace $\Lambda^k(V) \subset \Lambda(V)$ of all elements of the form

$$x_1 x_2 \cdots x_k, \quad \text{where} \quad x_1, \ldots, x_k \in V,$$

If $n = \dim(V)$ is finite, then a basis of $\Lambda^k(V)$ can always be generated from basis vectors of V. In particular if $\{e_1, \ldots, e_n\}$ is a basis of V, then

$$\{e_{i_1} e_{i_2} \cdots e_{i_k} \mid 1 \leq i_1 < i_2 < \cdots < i_k \leq n, \quad \text{with} \quad i_1, \ldots, i_k \in \mathbb{N}\} \tag{4.2}$$

is a basis for $\Lambda^k(V)$. This is true, because the bilinearity of the exterior product ensures that any exterior product can be expanded into a linear combination of exterior products of basis vectors of V. Furthermore, all terms can be re-ordered to obtain the canonical ordering in (4.2) and any term involving repeating factors $e_i e_i = 0$ vanishes. Another consequence of that last fact is that $\Lambda^k = \{0\}$ for $k > n$. Hence, $\Lambda(V)$ can be written as a direct sum of the first n exterior powers of V:

$$\Lambda(V) = K \oplus \Lambda^1(V) \oplus \cdots \oplus \Lambda^n(V).$$

Therefore, it is also possible to generate a basis for $\Lambda(V)$ using the n basis vectors $\{e_i\}$ of V. Next, observe that $e_i e_j = -e_j e_i$. More generally, if a basis for $\Lambda(V)$ is generated by L generators ξ^i satisfying

$$\xi^i \xi^j = -\xi^j \xi^i,$$

then we denote the Grassmann algebra as $\mathcal{G}_L$, with L finite. An arbitrary element $q \in \mathcal{G}_L$ can then be written as

$$q = q_{\mathbf{b}} + \sum_{1 \leq \mu_1 < \mu_2 < \cdots < \mu_k \leq L} q_{\mu_1,\ldots,\mu_k} \xi^{\mu_1} \xi^{\mu_2} \cdots \xi^{\mu_k},$$

where $\mu_1, \ldots, \mu_k \in \mathbb{N}$. The sum is finite, because higher order terms vanish due to the nilpotency of the generators: $\xi^i \xi^i = 0$. The variables $q_{\mathbf{b}}$ and $q_{\mu_1,\ldots,\mu_k}$ are elements of the field K. We call $q_{\mathbf{b}}$ the *body* of q and the remaining terms $q - q_{\mathbf{b}}$ constitute the *soul*. Numbers with nonzero body will turn out to be very important, because they are precisely the invertible elements of the Grassmann algebra.

We now return to the issue of determining the even and odd parts of $\mathcal{G}_L$. Even numbers $x \in \mathcal{G}_{L,0}$ can be generated by monomials of even degree: [29]

$$x = x_{\mathbf{b}} + x_{ij} \xi^i \xi^j + x_{ijkl} \xi^i \xi^j \xi^k \xi^l + \ldots, \tag{4.3}$$

where we implicitly sum over repeated indices. Odd numbers $\theta, \bar{\theta} \in \mathcal{G}_{L,1}$ can be written in terms of monomials of odd degree:

$$\begin{aligned} \theta &= \theta_i \xi^i + \theta_{ijk} \xi^i \xi^j \xi^k + \ldots, \\ \bar{\theta} &= \bar{\theta}_i \xi^i + \bar{\theta}_{ijk} \xi^i \xi^j \xi^k + \ldots. \end{aligned}$$

Written in this way, it is easy to see that the even numbers x commute with all other numbers and the odd numbers $\theta, \bar{\theta}$ anticommute amongst themselves, as required by the graded commutativty property (4.1). In supersymmetric field theories, the coordinates $x = x^\mu$ are real and so the coefficients $x_i, x_{ij}, \ldots$ are real. Meanwhile, the coordinates $\theta = \theta^\alpha, \bar{\theta} = \bar{\theta}^{\dot{\alpha}}$ are Weyl spinors and so the coefficients $\theta_i, \theta_{ijk}, \ldots$ are complex. In section 4.2, we show how to deal with tuples such as $x^\mu = (x^1, x^2, x^3, x^4)$, where each of the x^i are of the form (4.3). We now prove the following results about Grassmann numbers with nonzero body.

Lemma 4.1. *A number $x \in \mathcal{G}_L$ has nonzero body if and only if $x^n \neq 0$ for all $n \in \mathbb{N}$.*

Proof. Decomposing x into body and soul, we see that

$$x^n = (x_{\mathbf{b}} + \tilde{x})^n = x_{\mathbf{b}}^n + \ldots.$$

It is clear that $x^n \neq 0$ if $x_{\mathbf{b}} \neq 0$, because x^n has body $x_{\mathbf{b}}^n \neq 0$. The converse is immediate. □

Lemma 4.2. *Numbers $x \in \mathcal{G}_L$ have a multiplicative inverse $x^{-1} \in \mathcal{G}_L$ if and only if they have nonzero body. Moreover, the inverse is unique if it exists. Furthermore, Grassmann numbers are either invertible or nilpotent.*

Proof. First, suppose that x has nonzero body. Decomposing x into body and soul, we write

$$x = x_{\mathbf{b}} + \tilde{x}.$$

Now, using an identity akin to the geometric series for regular numbers, we observe that the number $1 - \alpha\tilde{x}$ (for $0 \neq \alpha \in K$) has an inverse given by

$$1 = (1 - \alpha\tilde{x})(1 + \alpha\tilde{x} + \alpha^2\tilde{x}^2 + \dots).$$

This sum converges for all $\tilde{x}$, because $\tilde{x}^n = 0$ for some finite n by Lemma 4.1. We now multiply the above with $x_{\mathbf{b}}$ and set $\alpha = -1/x_{\mathbf{b}}$ to find

$$x_{\mathbf{b}} = (x_{\mathbf{b}} + \tilde{x}) \left(1 - \frac{\tilde{x}}{x_{\mathbf{b}}} + \frac{\tilde{x}^2}{x_{\mathbf{b}}^2} - \dots \right).$$

Hence, the inverse is given by

$$x^{-1} = \frac{1}{x_{\mathbf{b}}} - \frac{\tilde{x}}{x_{\mathbf{b}}^2} + \frac{\tilde{x}^2}{x_{\mathbf{b}}^3} - \dots$$

This proves existence in case x has nonzero body. On the other hand, suppose that x has an inverse $y = x^{-1}$. If $x_{\mathbf{b}} = 0$, then the equation $1 = (x_{\mathbf{b}} + \tilde{x})(y_{\mathbf{b}} + \tilde{y})$ has no solutions, because the left-hand side has body 1 and the right-hand side has body 0. Hence, we must have $x_{\mathbf{b}} \neq 0$. For uniqueness, suppose there exist two inverses y and z. Then,

$$1 = (yx)z = y(xz) = z = y.$$

By associativity of $\mathcal{G}_L$, it follows that x^{-1} is unique. The final statement now follows from Lemma 4.1. □

Corollary 4.3. *Let $x, y \in \mathcal{G}_L$ be nonzero. If $xy = 0$, then the body of both x and y vanishes.*

Proof. A zero divisor cannot be a unit. The result now follow from Lemma 4.2. □

We can define a module over $\mathcal{G}_L$, allowing us to handle systems of linear equations involving Grassmann variables. There are general theorems concerning the existence of solutions of systems of linear equations in such modules. See for instance [39, p. 15]. For our application in chapter 6, we shall encounter systems involving only even Grassmann numbers. This is fortunate, because the even Grassmann numbers $\mathcal{G}_{L,0}$ form a commutative subring of $\mathcal{G}_L$. This makes life much easier when working with systems of linear equations. For instance, determinants of matrices over commutative rings are well-defined [11, III, §8]. We now prove the result below, which will be used in section 6.3.

Lemma 4.4. *Let R be a unital, commutative ring. Let $M \in R^{(n+1)\times(n+1)}$ be the block matrix*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

with $A \in R^{n\times n}$, $B \in R^{n\times 1}$, $C \in R^{1\times n}$, and $D \in R^{1\times 1}$. Then, $\det(M) = \det(AD - BC)$. *Furthermore, if A is invertible, then* $\det(M) = \det(A)\det(D - CA^{-1}B)$.

The first statement is proved using the techniques of [71], where an analogous result is proved for the case where A, B, C, D are all square matrices of the same order. The second statement follows immediately from a simple matrix identity.

Proof. Denote by 1 the multiplicative identity in R. We have the following identity

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} D & 0 \\ -C & 1 \end{pmatrix} = \begin{pmatrix} AD - BC & B \\ 0 & D \end{pmatrix}, \tag{4.4}$$

where we have used the fact that $CD = DC$, because $D \in R$ and R is commutative. Observing that $\det(D) = D$, it follows that

$$\det(M)\det(D) = \det(AD - BC)\det(D).$$

If $D \neq 0$ and D is not a zero divisor, then $\det(M) = \det(AD - BC)$ as desired. However, this is not required. To see this, let $R[x]$ be the polynomial ring in indeterminate x. Then, $R[x]$ is also commutative. Let $D_x = x + D$ and

$$M_x = \begin{pmatrix} A & B \\ C & D_x \end{pmatrix}.$$

Once again, we can use (4.4) with D_x instead of D and M_x instead of M, to find that

$$[\det(M_x) - \det(AD_x - BC)]\det(D_x) = 0.$$

In this equation, $\det(D_x) = \det(x + D) = x + D$ is not a zero divisor. It follows that $\det(M_x) - \det(AD_x - BC) = 0$. Evaluating at $x = 0$ gives the desired result.

For the second part, we assume that A is invertible, which allows us to write M as

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & 1 \end{pmatrix} \begin{pmatrix} I_{n\times n} & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix}.$$

Now, the result follows from the first part. ∎

4.2 Superspace and superfunctions

Remark 2. We now specify the field, setting $K = \mathbb{C}$. Furthermore, the body $q_{\mathbf{b}}$ of each number $q \in \mathcal{G}_L$ is assumed to be real.

It is possible to construct a Hilbert space over a Grassmann algebra [64, 53], but we will follow a different route based on the notion of superfunctions [61, 29]. In general, a superfunction is a function $F\colon \mathcal{G}_L \to \mathcal{G}'_L$ between Grassmann algebras. In supersymmetric field theories, superfields may be considered as superfunctions $(x, \theta, \bar{\theta}) \mapsto F(x, \theta, \bar{\theta})$. Superfields are defined by their expansion in powers of the anticommuting coordinates $\theta, \bar{\theta}$:

$$F(x, \theta, \bar{\theta}) = \sum_{i=0,j=0}^{K} f_{i,j}(x)\theta^i\bar{\theta}^j, \tag{4.5}$$

where the sum terminates at some finite power K. The coefficient functions $f_{i,j}(x)$ depend only on the even coordinates x and are interpreted as the component fields of the supermultiplet F (see section 5.2). Henceforth, we assume that superfunctions are always of the form (4.5) and we use the words *superfunctions* and *superfields* interchangeably.

Before we can discuss smooth superfunctions, we first need to find a topology for $\mathcal{G}_L$. We shall use the topology induced by the norm [63, 29]:

$$\|q\|_p = \left(|q_{\mathbf{b}}|^p + \sum_{1 \le \mu_1 < \mu_2 < \cdots < \mu_k \le L} |q_{\mu_1,\mu_2,\ldots,\mu_k}|^p \right)^{1/p}.$$

This norm satisfies $\|\mathcal{I}\| = 1$ and $\|qq'\| \le \|q\|\|q'\|$. Together with this norm, the Grassmann algebra becomes a Banach algebra [29]. We introduce the following definition.

Definition 4.3. Let $\mathcal{G}_L = \mathcal{G}_{L,0} \oplus \mathcal{G}_{L,1}$ be a *Grassmann-Banach algebra*, which is a Grassmann algebra that is also a Banach algebra. The (m, n)-dimensional space $\mathcal{G}_L^{m,n} = (\mathcal{G}_{L,0})^m \times (\mathcal{G}_{L,1})^n$ is called *superspace* and generalises $\mathbb{R}^m$.

Superfields in supersymmetric field theories depend on tuples of even and odd variables. For example, in the case of $\mathcal{N} = 1$ supersymmetry with spacetime dimension $d = 4$, the superfields depend on $x^\mu = (x^1, x^2, x^3, x^4)$, $\theta_\alpha = (\theta_1, \theta_2)$, and $\bar{\theta}_{\dot{\alpha}} = (\bar{\theta}_1, \bar{\theta}_2)$. Hence, considering Definition 4.3, they are best viewed as maps $\mathcal{G}_L^{4,4} \to \mathcal{G}_L$. We therefore need to adapt the expansion in (4.5) to the general case, where θ and $\bar{\theta}$ are tuples of odd numbers: $\theta = (\theta_1, \ldots, \theta_r)$ and $\bar{\theta} = (\bar{\theta}_1, \ldots, \bar{\theta}_k)$. The expansion should then be written as

$$F(x, \theta, \bar{\theta}) = \sum_{(i),(j)} f_{(i),(j)}(x)\ \theta^{(i)}\bar{\theta}^{(j)}, \tag{4.6}$$

where the sum is over all the monomials involving $\theta_1, \ldots, \theta_r, \bar{\theta}_1, \ldots, \bar{\theta}_k$ up to degree $(r+k)K$, analogous to (4.5). Of course, the coefficient functions are easily adapted to the case where $x = (x_1, \ldots, x_m)$ is a tuple of even numbers.

We now define the idea of smooth or G^∞ superfunctions as superfunctions with smooth coefficient functions $f_{(i),(j)} \colon \mathbb{R}^m \to \mathcal{G}_L$, extended to functions $\mathcal{G}_L^{m,0} \to \mathcal{G}_L$ through a process called z-continuation. The basic idea is simple. Assume that the coefficient functions are infinitely differentiable when restricted to soulless variables $x = x_{\mathbf{b}} \in \mathbb{R}^m$. Then, we expand the coefficient functions as a power series in the soul of x. Such a function is then called $G^\infty\left(\mathcal{G}_L^{m,0}, \mathcal{G}_L\right)$. By extension, a superfunction $F \colon \mathcal{G}_L^{m,n} \to \mathcal{G}_L$ with such coefficient functions is then $G^\infty\left(\mathcal{G}_L^{m,n}, \mathcal{G}_L\right)$. Let's make this more precise.

Definition 4.4 (Z-continuation)**.** We let $\epsilon \colon \mathcal{G}_L \to \mathbb{R}$ be the body projection map, such that $\epsilon(q) = q_{\mathbf{b}} \in \mathbb{R}$. Given an open set $U \subset \mathcal{G}_L^{m,0}$, we denote by V the open set $\epsilon(U) \subset \mathbb{R}^m$. Then a function $f \in C^\infty(V, \mathcal{G}_L)$ can be extended to a function $z(f) \in G^\infty(U, \mathcal{G}_L)$ through a power series

$$z(f)(x_1, \ldots, x_m) = \sum_{i_1 = \cdots = i_m = 0}^{L} \frac{1}{i_1! \cdots i_m!} \left[\frac{\partial^{i_1}}{\partial x_1^{i_1}} \cdots \frac{\partial^{i_m}}{\partial x_m^{i_m}}\right] f\left(\epsilon(x_1), \ldots, \epsilon(x_m)\right) \Delta_1^{i_1} \cdots \Delta_m^{i_m},$$

where $\Delta_i = x_i - \epsilon(x_i)$ for $i = 1, \ldots, m$.

With all this in place, G^∞ superfunctions can be differentiated. Because superfunctions only depend on the even coordinates x through the smooth component functions $f_{(i),(j)}(x)$, differentiation with respect to the even coordinates is done by simply differentiating the component functions. Differentiation with respect to the odd coordinates $\theta = (\theta_1, \ldots, \theta_r)$ and $\bar{\theta} = (\bar{\theta}_1, \ldots, \bar{\theta}_k)$ is defined by the rules

$$\frac{\partial \theta_j}{\partial \theta_i} = \delta_i^j, \qquad \frac{\partial \bar{\theta}_j}{\partial \theta_i} = 0, \qquad \frac{\partial \bar{\theta}_j}{\partial \bar{\theta}_i} = \delta_i^j, \qquad \frac{\partial \theta_j}{\partial \bar{\theta}_i} = 0,$$

in addition to the Leibniz rules

$$\frac{\partial (FG)}{\partial \theta_i} = \left(\frac{\partial F}{\partial \theta_i}\right) G + F \left(\frac{\partial G}{\partial \theta_i}\right), \qquad \frac{\partial (FG)}{\partial \bar{\theta}_i} = \left(\frac{\partial F}{\partial \bar{\theta}_i}\right) G + F \left(\frac{\partial G}{\partial \bar{\theta}_i}\right),$$

where $F, G \in G^\infty\left(\mathcal{G}_L^{m,n}, \mathcal{G}_L\right)$. These rules completely define differentiation for G^∞ functions of the form (4.6).

We now extend the concept of superspace to that of a supermanifold. Roughly speaking, a smooth (m,n)-dimensional supermanifold $\mathcal{M}$ is topological space that is locally homeomorphic to "flat superspace" $\mathcal{G}_L^{m,n}$, in the same way that an ordinary smooth real m-dimensional manifold is locally homeomorphic to $\mathbb{R}^m$. Formally, we have the following definition.

Definition 4.5 (Supermanifold)**.** A *supermanifold* $\mathcal{M}$ is a paracompact topological Hausdorff space together with an atlas of charts $\{(X_\alpha, \phi_\alpha) \mid \alpha \in I\}$ for some indexing set I, where the coordinate maps $\phi_\alpha \colon X_\alpha \to \tilde{X}_\alpha \subset \mathcal{G}_L^{m,n}$ are local homeomorphisms into open subsets $\tilde{X}_\alpha$ of flat superspace $\mathcal{G}_L^{m,n}$.

The theory of supermanifolds is incredibly rich, but we stop here. Our ultimate goal is to formalise the idea that superfields are linear representations of the superPoincaré group. This is a challenge which we will pick up in the next section.

4.3 Representations of supersymmetry

In supersymmetric field theories, we are interested in representations of the superPoincaré group. See section 5.1 for a detailed discussion of supersymmetry and the physical aspects of the superPoincaré algebra. Before discussing this, let us briefly recall the relation between ordinary Lie algebras and Lie groups. We follow [12, 23] by defining a Lie group as follows.

Definition 4.6. A *Lie group* is a smooth manifold G with a smooth map $\mu \colon G \times G \to G$, under which G forms a group.

Definition 4.7. A *Lie algebra* is vector space $\mathfrak{g}$ over a field F together with an operation $[\,\cdot\,,\,\cdot\,] \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, satisfying

1. (Bilinearity): $[ax + by, z] = a[x, z] + b[y, z]$ and $[z, ax + by] = a[z, x] + b[z, y]$, for $a, b \in F$ and $x, y, z \in \mathfrak{g}$,
2. (Alternativity): $[x, x] = 0$, for $x \in \mathfrak{g}$,
3. (Jacobi identity): $[x[y, z]] + [z, [x, y]] + [y, [z, x]] = 0$, for $x, y, z \in \mathfrak{g}$.

The relation between Lie algebras and Lie groups is given by the following theorem, which we do not prove.

Theorem 4.5. *Each Lie group G is associated with a Lie algebra $\mathfrak{g} = T_eG$ formed by the tangent space at the identity $e \in G$. Moreover, the exponential map* $\exp \colon \mathfrak{g} \to G$ *maps a vector $v \in \mathfrak{g}$ to a point $V \in G$ defined by $V = \exp(v) = \gamma_e(1)$, where $\gamma_e \colon \mathbb{R} \to G$ is the integral curve of the vector field X_v with initial condition e.*

Now, we generalise the notion of a Lie algebra to that of a Lie superalgebra.

Definition 4.8. A *Lie superalgebra* $V = V_0 \oplus V_1$ is a Lie algebra that is also a commutative superalgebra (see Definition 4.1). The bilinear multiplication operation is now the graded Lie bracket $\{\!\![\,\cdot, \cdot\,]\!\!\}$, satisfying

1. (Supercommutativity): $\{\!\!\{ v_1, v_2 \}\!\!\} = -(-1)^{|v_1||v_2|} \{\!\!\{ v_2, v_1 \}\!\!\}$,
2. (Super Jacobi identity): $(-1)^{|v_1||v_3|} \{\!\!\{ v_1, \{\!\!\{ v_2, v_3 \}\!\!\}\}\!\!\} +$ cyclic permutations $= 0$,

where $|v| = 0$ if $v \in V_0$ and $|v| = 1$ if $v \in V_1$.

By analogy with Definition 4.6, we define a Lie supergroup as follows [74].

Definition 4.9. A *Lie supergroup* is a supermanifold G with a smooth map $\mu\colon G \times G \to G$, under which G forms a group.

We are particularly interested in the $\mathcal{N} = 1$ superPoincaré algebra, which is generated by the 10 even generators of the Poincaré algebra: 4 generators of translations $\mathcal{P}_\mu$, 6 generators of Lorentz transformations $\mathcal{M}_{\mu\nu}$, in addition to 2 odd generators $\mathcal{Q}^\alpha$ and $\bar{\mathcal{Q}}^{\dot\alpha}$. For the full commutation and anticommutation relations, refer to equations (5.2) and (5.3) and the surrounding discussion in chapter 5. Here, we use calligraphic letters for the Lie algebra generators, to distinguish them from the differential operators $P^\mu, Q^\alpha, \bar{Q}^{\dot\alpha}, M_{\mu\nu}$.

A result analogous to that of Theorem 4.5 holds for Lie superalgebras and Lie supergroups. For instance, refer to chapter VI of [74] for an explicit construction. Here, we simply follow [81] and use the exponential map to define a group element of the identity component of the $\mathcal{N} = 1$ Lie superPoincaré group by

$$G(x^\mu, \theta^\alpha, \bar\theta^{\dot\alpha}, \omega^{\mu\nu}) = \exp\left(ix^\mu \mathcal{P}_\mu + i\theta^\alpha \mathcal{Q}_\alpha - i\bar\theta^{\dot\alpha}\bar{\mathcal{Q}}_{\dot\alpha} + \tfrac{1}{2} i\omega^{\mu\nu}\mathcal{M}_{\mu\nu}\right).$$

The minus sign is because $\bar\theta\bar{\mathcal{Q}} \equiv \bar\theta_{\dot\alpha}\bar{\mathcal{Q}}^{\dot\alpha} = -\bar\theta^{\dot\alpha}\bar{\mathcal{Q}}_{\dot\alpha}$. We shall see in chapter 5 that these group elements satisfy the property[1]:

$$G\begin{pmatrix} y^\mu \\ \xi^\alpha \\ \bar\xi^{\dot\alpha} \\ \omega^{\rho\sigma} \end{pmatrix}^{\mathbf{T}} G\begin{pmatrix} x^\mu \\ \theta^\alpha \\ \bar\theta^{\dot\alpha} \\ 0 \end{pmatrix}^{\mathbf{T}} = G\begin{pmatrix} y^\mu + x^\mu + i\xi^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha} - i\theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha} + \frac14 x_\nu\left(\omega^{\nu\mu} - \omega^{\mu\nu}\right) \\ \xi^\alpha + \theta^\alpha - \frac14\omega^{\rho\sigma}\xi^\beta\left(\sigma_{\rho\sigma}\right)_\beta{}^\alpha \\ \bar\xi^{\dot\alpha} + \bar\theta^{\dot\alpha} - \frac14\omega^{\rho\sigma}\bar\xi_{\dot\beta}\left(\bar\sigma_{\mu\nu}\right)^{\dot\beta}{}_{\dot\alpha} \\ 0 \end{pmatrix}^{\mathbf{T}} \tag{4.7}$$

We note that the subspace of elements of the form $G(x, \theta, \bar\theta, 0)$ is invariant under the group action $G \times G \to G\colon (g, h) \mapsto gh$. Equation (4.7) is only a second-order approximation. However, we usually assume that $\omega^{\rho\sigma} = 0$, in which case the expression is exact. Moreover, we will only be dealing with infinitesimal transformations. Regardless, the statement that the subspace of elements of the form $G(x, \theta, \bar\theta, 0)$ is invariant is also exact.

Equation (4.7) allows us to define a consistent group action on superspace $\mathcal{G}_L^{4,4}$ by

$$G\left(y^\mu, \xi^\alpha, \bar\xi^{\dot\alpha}, \omega^{\rho\sigma}\right) \cdot \begin{pmatrix} x^\mu \\ \theta^\alpha \\ \bar\theta^{\dot\alpha} \end{pmatrix}^{\mathbf{T}} = \begin{pmatrix} x^\mu + y^\mu + i\xi^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha} - i\theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha} + \frac14 x_\nu\left(\omega^{\nu\mu} - \omega^{\mu\nu}\right) \\ \xi^\alpha + \theta^\alpha - \frac14\omega^{\rho\sigma}\xi^\beta\left(\sigma_{\rho\sigma}\right)_\beta{}^\alpha \\ \bar\xi^{\dot\alpha} + \bar\theta^{\dot\alpha} - \frac14\omega^{\rho\sigma}\bar\xi_{\dot\beta}\left(\bar\sigma_{\mu\nu}\right)^{\dot\beta}{}_{\dot\alpha} \end{pmatrix}^{\mathbf{T}}.$$

[1]Here we use the notation $G\left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)^{\mathbf{T}}$ as shorthand for $G(a, b)$.

This is compatible with the group structure, by which we mean that the action of $G'G$ is the same (to first order) as the action of G followed by the action of G'. As a result, the map $\rho\colon G\times\mathcal{G}_L^{4,4}\to\mathcal{G}_L^{4,4}$ is a non-linear realisation. Often, we are only interested in supersymmetry transformations, so we set $y^\mu=\omega^{\rho\sigma}=0$, which yields the much simpler action

$$G\left(0,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},0\right)\cdot\begin{pmatrix}x^\mu\\ \theta^\alpha\\ \bar{\theta}^{\dot{\alpha}}\end{pmatrix}^{\mathbf{T}}=\begin{pmatrix}x^\mu+i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}-i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\\ \xi^\alpha+\theta^\alpha\\ \bar{\xi}^{\dot{\alpha}}+\bar{\theta}^{\dot{\alpha}}\end{pmatrix}^{\mathbf{T}}. \tag{4.8}$$

We are now in a position to discuss linear representations on superspace and on the space of superfunctions. Given a group action $G\times V\to V$ on a vector space V, there exists an induced group action $G\times V^*\to V^*$ on the dual space V^*, defined by $(T,f)\mapsto f\circ T^{-1}$ [43]. Similarly, we can extend a G^∞ action $G\times\mathcal{G}_L\to\mathcal{G}_L$ on the Grassmann algebra $\mathcal{G}_L$ to a related action $G\times G^\infty(\mathcal{G}_L)\to G^\infty(\mathcal{G}_L)$ on the space of smooth superfunctions. The same can be done for superspace $\mathcal{G}_L^{m,n}$ in general.

Let's use this to define an action of the superPoincaré group on the space of smooth superfields. We already have a smooth group action on $\mathcal{G}_L^{4,4}$, so we immediately obtain the related action $\rho^*\colon G\times G^\infty\left(\mathcal{G}_L^{4,4}\right)\to G^\infty\left(\mathcal{G}_L^{4,4}\right)$ on the space of smooth superfunctions:

$$G\left(y^\mu,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},\omega^{\rho\sigma}\right)\cdot F\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right)=F\left(G\left(y^\mu,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},\omega^{\rho\sigma}\right)\cdot\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right)\right)$$

where we chose to act with G rather than G^{-1} in order to be consistent with the physics literature. Written out in full, and neglecting the second-order terms $\omega\xi\sigma$ and $\omega\bar{\xi}\bar{\sigma}$, the action becomes

$$\begin{aligned}G\left(y^\mu,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},\omega^{\rho\sigma}\right)F\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right)=F\big(&x^\mu+y^\mu+i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}-i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}+\tfrac{1}{4}x_\nu\left(\omega^{\nu\mu}-\omega^{\mu\nu}\right),\\ &\xi^\alpha+\theta^\alpha,\bar{\xi}^{\dot{\alpha}}+\bar{\theta}^{\dot{\alpha}}\big).\end{aligned} \tag{4.9}$$

By taking the differential of this non-linear realisation (a Lie group homomorphism) at the identity, we obtain a linear representation of the superPoincaré algebra. We write the action (4.9) in terms of the linear differential operators

$$\begin{aligned}Q_\alpha&=-i\frac{\partial}{\partial\theta^\alpha}-\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\frac{\partial}{\partial x^\mu},\\ \bar{Q}_{\dot{\alpha}}&=i\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}}+\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\frac{\partial}{\partial x^\mu},\\ P_\mu&=-i\frac{\partial}{\partial x^\mu}\\ M_{\mu\nu}&=-ix_\mu\frac{\partial}{\partial x^\nu}+ix_\nu\frac{\partial}{\partial x^\mu}.\end{aligned}$$

The names $Q_\alpha,\bar{Q}_{\dot{\alpha}},P_\mu,M_{\mu\nu}$ are justified, because the operators close the algebra (5.3). For example, it can easily be checked that $\{Q_\alpha,\bar{Q}_{\dot{\alpha}}\}=-2i\sigma^\mu_{\alpha\dot{\alpha}}\partial_\mu=2\sigma^\mu_{\alpha\dot{\alpha}}P_\mu$. The group action for infinitesimal $y^\mu, \xi^\alpha, \bar{\xi}^{\dot{\alpha}}, \omega^{\mu\nu}$ then becomes

$$G\left(y^\mu, \xi^\alpha, \bar{\xi}^{\dot{\alpha}}, \omega^{\mu\nu}\right) F\left(x^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}\right) = \left(1 + iy^\mu P_\mu + i\xi^\alpha Q_\alpha - i\bar{\xi}^{\dot{\alpha}}\bar{Q}_{\dot{\alpha}} + \tfrac{1}{2}i\omega^{\mu\nu}M_{\mu\nu}\right) F\left(x^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}\right).$$

Expanding this expression, one recovers (4.9). We remarked before that the group action on superspace was a non-linear realisation $\rho\colon G \times \mathcal{G}_L^{4,4} \to \mathcal{G}_L^{4,4}$. Now, in fact, we have found a linear representation $\rho^*\colon G \times G^\infty\left(\mathcal{G}_L^{4,4}\right) \to G^\infty\left(\mathcal{G}_L^{4,4}\right)$. We refer to this representation – and by abuse of notation, also the function space $G^\infty\left(\mathcal{G}_L^{4,4}\right)$ and its elements – as the linear representation of supersymmetry.

This concludes our discussion of the linear representation of supersymmetry, which is all we need for chapter 5. However, in chapter 6, we will be dealing with the Volkov-Akulov realisation of supersymmetry [77]. Without going into too much detail here, it suffices to say that the Volkov-Akulov realisation describes a single goldstino spinor $\lambda_\alpha(x)$, which transforms non-linearly under a supersymmetry transformation, according to:

$$\lambda_\alpha(x) \to \lambda_\alpha(x) + \frac{i}{\sqrt{2}f}\left(\lambda^\alpha \sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - \xi^\alpha \sigma^\mu_{\alpha\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}}\right)\partial_\mu \lambda_\alpha + \sqrt{2}f\xi_\alpha. \tag{4.10}$$

This transformation law can be derived by introducing a correspondence $\lambda_\alpha(x) \leftrightarrow \sqrt{2}f\theta_\alpha$. See also section 6.1. Can we formulate the Volkov-Akulov field as a non-linear realisation, using the machinery of this chapter? The answer is yes. We define a spurious superfunction

$$F\left(x^\mu, \lambda^\alpha, \bar{\lambda}^\alpha\right) = \lambda^\alpha \lambda_\alpha\left(x^\mu\right).$$

We call this superfunction "spurious", because $\lambda(x)$ plays both the role of a component function and that of an odd coordinate. Under a supersymmetry transformation $G(\xi, \bar{\xi}) \equiv G(0, \xi, \bar{\xi}, 0)$, this function transforms as

$$G\left(\xi^\alpha, \bar{\xi}^{\dot{\alpha}}\right) F\left(x^\mu, \lambda^\alpha, \bar{\lambda}^{\dot{\alpha}}\right) = \left(1 + i\xi^\alpha Q_\alpha - i\bar{\xi}^{\dot{\alpha}}\bar{Q}_{\dot{\alpha}}\right) F\left(y^\mu, \lambda^\alpha, \bar{\lambda}^{\dot{\alpha}}\right). \tag{4.11}$$

Substituting our correspondence $\lambda_\alpha(x) = \sqrt{2}f\theta_\alpha$ into the expressions for the differential operators $Q_\alpha, \bar{Q}_{\dot{\alpha}}$, we find

$$\begin{aligned} G\left(\xi^\alpha, \bar{\xi}^{\dot{\alpha}}\right)\lambda^\alpha\lambda_\alpha &= \left(1 + \xi^\alpha \frac{\partial}{\partial\theta^\alpha} + \frac{1}{\sqrt{2}f} i\lambda^\beta \sigma^\mu_{\beta\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\frac{\partial}{\partial x^\mu} - \frac{1}{\sqrt{2}f} i\xi^\beta \sigma^\mu_{\beta\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}}\frac{\partial}{\partial x^\mu}\right)\lambda^\alpha\lambda_\alpha\left(x^\mu\right) \\ &= \lambda^\alpha\left[\lambda_\alpha + 2\sqrt{2}f\xi_\alpha + \frac{2i}{\sqrt{2}f}\left(\lambda^\beta\sigma^\mu_{\beta\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - \xi^\beta\sigma^\mu_{\beta\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}}\right)\partial_\mu\lambda_\alpha\right]. \end{aligned}$$

Now, splitting off the odd coordinate λ^α, we obtain the transformed coordinate function λ'_α, which matches the transformation law (4.10) up to a factor 2 in the variation. This factor 2 is a result of properly treating the odd coordinate $\lambda^\alpha = \lambda^\alpha(x)$ as a function that should be differentiated. The factor 2 is absent in the conventional derivation of the Volkov-Akulov transformation law (e.g. [81, p. 88]), where the interpretation of the analogy $\lambda_\alpha(x) \leftrightarrow \sqrt{2} f \theta_\alpha$ is unclear. If we absorb the factor 2 into the correspondence $\lambda_\alpha(x) = 2\sqrt{2} f \theta_\alpha$, then the terms proportional to $\partial_\mu \lambda$ in the resulting transformation law match those in (4.10). However, as a result, the supersymmetry breaking term in the transformation law becomes $4\sqrt{2} f \xi_\alpha$, which is 4 times larger than its conventional value. In what follows, we will continue to use (4.10) without the factor 2, so that our results in chapter 6 can be compared with the literature.

A final comment is in order. Although the expression in (4.11) suggests that the Volkov-Akulov field is a linear representation, this is actually false: the inhomogeneous term $2\sqrt{2} f \xi_\alpha$ ensures that the origin is not invariant. This happens because of the double role that λ_α plays in the spurious superfunction.

5

Supersymmetry

5.1 Introduction

Supersymmetry is a proposed symmetry of Nature relating fermions and bosons. If supersymmetry is actually realised, then we expect some operators $(Q, \bar{Q})$, mapping fermions into bosons and vice versa, by raising or lowering spin quantum numbers by units of $\frac{1}{2}$. In its most basic conception, there is just one such pair. However, if one considers multiple pairs of operators $(Q, \bar{Q})$, one is dealing with so-called extended supersymmetry. The number of pairs is denoted $\mathcal{N}$. For instance, in case $\mathcal{N} = 2$ and given a one-particle state $|s\rangle$ with spin s, one might expect a multiplet of states satisfying something like

$$\begin{aligned} Q^1|s\rangle &= Q^2|s\rangle = 0, \\ \bar{Q}^1|s\rangle &= |s + \tfrac{1}{2}\rangle_1, \\ \bar{Q}^2|s\rangle &= |s + \tfrac{1}{2}\rangle_2, \\ \bar{Q}^1\bar{Q}^2|s\rangle &= |s + 1\rangle. \end{aligned} \tag{5.1}$$

If we set $s = 0$, then this *supermultiplet* contains one spin-0 state, two spin-$\frac{1}{2}$ states, and one spin-1 state. Supermultiplets are important for supersymmetric field theories, where the Lagrangian is supposed to be invariant under supersymmetry transformations. This is best accomplished by working directly with supermultiplets using the *superfield formalism* of section 5.2. The Standard Model can be extended with supersymmetry by associating new superparticles with each Standard Model particle. One such extension is the Minimal Supersymmetric Standard Model. In general, the supermultiplets are made up of a Standard Model particle together with its superpartners, which differ in spin, but otherwise share the same mass and quantum numbers. This is a consequence of commutators such as

$$[Q^I, P_\mu] = [\bar{Q}^I, P_\mu] = 0.$$

This means, for instance, that we should observe a "selectron" with the same mass and electric charge as an electron, but with spin 0. We know from observations that no such superparticles exist. Hence, if supersymmetry exists, it must be in a broken form. This explains our interest in supersymmetry breaking. The relations (5.1) also imply that the Q's should transform as spin-$\frac{1}{2}$ particles themselves, which means that supersymmetry is a spacetime symmetry, unlike the internal symmetries considered in chapter 2. We could say a lot more at this point, but considering the lack of evidence for supersymmetry, we should first take a moment to justify our interest in this topic.

Back in 2013, amidst the excitement about the discovery of the Higgs boson, theorists were starting to turn their backs on supersymmetry [82]. The results of Run 1 (2009-2013) of the Large Hadron Collider [30, 66] had been inconclusive. On the one hand, no statistically significant evidence for physics beyond the Standard Model had been found. On the other hand, the results did not rule out the Minimal Supersymmetric Standard Model, which would have been the case if the Higgs had been found to be heavier [5]. Three years later, the first results of Run 2 (2015-2017) are in [4, 68]. The story is much the same, although lower bounds on the masses of some of the lightest predicted superpartners have increased. Interestingly, there was also a slight hint of a new particle at around 750 GeV [18, 3]. Whether future results bear out this observation remains to be seen, but supersymmetric explanations have already been offered [56]. The question remains: without any direct evidence, why are we still interested in supersymmetry? There are three main reasons.

The first reason is that supersymmetry is the only allowed nontrivial combination of spacetime and internal symmetries for realistic[1] quantum field theories [81]. This is a consequence of the no-go theorem of Coleman and Mandula [20], which states that in a realistic quantum field theory, the Lie algebra of symmetries of the S-matrix must be generated by translations P_μ, Lorentz transformation $M_{\mu\nu}$ and internal symmetry generators B_a, without mixing between the internal and spacetime symmetries. This implies that the Lie group must be a direct product of the Poincaré group and the compact Lie group G of internal symmetries: $\mathrm{ISO}(3,1) \times G$. However, if one generalises the notion of Lie algebras to *graded Lie algebras*, a wider set of generators is allowed. In this light, the Haag-Łopuszański-Sohnius theorem [37] is an important generalisation of the Coleman-Mandula theorem. It tells us that the most general graded Lie algebra of symmetries of a realistic quantum field theory is generated by $P_\mu, M_{\mu\nu}, B_a$ in addition to the supergenerators $Q^I, \bar{Q}^I$. The Poincaré generators together with the supergenerators make up the superPoincaré algebra, which is an example of the superalgebras considered in chapter 4.

The preceding argument only shows that supersymmetry is possible. There is also reason to believe that supersymmetry is actually detectable at the Large Hadron Collider [8]. This

[1] A "realistic" theory is a theory that assumes locality, four spacetime dimensions, a finite number of particles of a given mass, and a nonzero energy gap between the vacuum and the one-particle states.

has to do with the hierarchy problem in particle physics: the large discrepancy between the Higgs mass (125 GeV) and the Planck mass (1.22×10^{19} GeV). In the Standard Model, the coupling of fermions to the Higgs field causes huge corrections to the Higgs mass above its tree-level value. These corrections are only limited by the UV cut-off scale Λ_{SM}. To stabilise the Higgs mass at 125 GeV, these corrections must somehow be cancelled. Supersymmetry provides a natural way to do this, by associating scalar superpartners with the same Higgs-coupling to each of the Standard Model fermions. This solves the hierarchy problem, because the scalar induced corrections to the Higgs mass are opposite in sign to the fermion induced corrections. For this to work, the supersymmetry breaking scale should be of order $\Lambda_{\mathrm{SM}} \simeq 1$ TeV, which is a feasible scale for detection at the Large Hadron Collider. A third reason to study supersymmetry is that some of the new superparticles are candidates for dark matter. In the case of the Minimal Supersymmetric Standard Model, this is the neutralino [8].

Let us now return to our discussion of what supersymmetry entails. We already mentioned that the superPoincaré algebra is an example of a Lie superalgebra. We formally defined Lie superalgebras in section 4.3. The ordinary Poincaré algebra, together with the Lie algebra of internal symmetries, is defined by the relations

$$\begin{aligned}
[P_\mu, P_\nu] &= 0, \\
[M_{\mu\nu}, M_{\rho\sigma}] &= -i\eta_{\mu\rho}M_{\nu\sigma} - i\eta_{\nu\sigma}M_{\mu\rho} + i\eta_{\mu\sigma}M_{\nu\rho} + i\eta_{\nu\rho}M_{\mu\sigma}, \\
[M_{\mu\nu}, P_\rho] &= -i\eta_{\rho\mu}P_\nu + i\eta_{\rho\nu}P_\mu, \\
[B_a, B_b] &= if_{ab}^c B_c, \\
[P_\mu, B_a] &= 0, \\
[M_{\mu\nu}, B_a] &= 0,
\end{aligned} \tag{5.2}$$

where f_{ab}^c are the structure constants of the internal symmetry algebra. The last two relations are required in order to be consistent with the Coleman-Mandula theorem. Now, the superPoincaré algebra is defined by extending the preceding relations with:

$$\begin{aligned}
[P_\mu, Q_\alpha^I] &= 0, \\
[P_\mu, \bar{Q}_{\dot\alpha}^I] &= 0, \\
[M_{\mu\nu}, Q_\alpha^I] &= i\,(\sigma_{\mu\nu})_\alpha{}^\beta\, Q_\beta^I, \\
[M_{\mu\nu}, \bar{Q}^{\dot\alpha, I}] &= i\,(\bar\sigma_{\mu\nu})^{\dot\alpha}{}_{\dot\beta}\, \bar{Q}^{\dot\beta, I}, \\
\{Q_\alpha^I, \bar{Q}_{\dot\beta}^J\} &= 2\sigma^\mu_{\alpha\dot\beta} P_\mu \delta^{IJ}, \\
\{Q_\alpha^I, Q_\beta^J\} &= \epsilon_{\alpha\beta} Z^{IJ}, \quad \text{with} \quad Z^{IJ} = -Z^{JI}, \\
\{\bar{Q}_{\dot\alpha}^I, \bar{Q}_{\dot\beta}^J\} &= \bar\epsilon_{\dot\alpha\dot\beta}(Z^{IJ})^*,
\end{aligned} \tag{5.3}$$

where the Z^{IJ} are the *central charges*. For $\mathcal{N} = 1$ supersymmetry, these relations simplify considerably. In particular, there are no central charges: $Z^{IJ} = 0$. This leaves us with five interesting relations. We already mentioned that the first two commutators $[P_\mu, Q] = [P_\mu, \bar{Q}] = 0$ imply that states in a supermultiplet share the same mass. The third and fourth relations can be used to show that the supercharges change the spin quantum numbers of states. Indeed, noting that $M_{12} = J_3$ corresponds to the z-component of angular momentum, we see that[2]

$$\begin{aligned} J_3 Q_\alpha |s\rangle &= Q_\alpha J_3 |s\rangle + [M_{12}, Q_\alpha]|s\rangle \\ &= Q_\alpha J_3 |s\rangle + i\left(\sigma_{\mu\nu}\right)_\alpha{}^\beta Q_\beta |s\rangle \\ &= \left(s \pm \tfrac{1}{2}\right) Q_\alpha |s\rangle, \end{aligned}$$

and similarly for $\bar{Q}$. This shows that linear combinations of Q and $\bar{Q}$ can be used to raise and lower spin by units of $\frac{1}{2}$. Finally, the fifth relation implies that the commutator of two supersymmetry transformations corresponds to a translation. This has important implications for supergravity, which we do not consider here. In the next section, we shall introduce representations of the $\mathcal{N} = 1$ superPoincaré algebra.

5.2 Field representations

It is possible to construct state representations of the superPoincaré algebra, such as the multiplet in (5.1). However, we are ultimately interested in constructing field representations. The easiest way to accomplish this is with the superfield formalism, which we now introduce. Most of this section is based on [8, 81]. A first step is to express the $\mathcal{N} = 1$ superPoincaré algebra in terms of commutation relations only. This can be done with the help of anticommuting variables $\theta^\alpha, \bar{\theta}^{\dot{\alpha}}$ transforming as Weyl spinors and satisfying

$$\{\theta^\alpha, \theta^\beta\} = \{\theta^\alpha, \bar{\theta}^{\dot{\alpha}}\} = \cdots = [P_\mu, \theta^\alpha] = [P_\mu, \bar{\theta}^{\dot{\alpha}}] = 0.$$

This effectively places us in the $(4,4)$-dimensional Grassmann-Banach algebra of chapter 4 with coordinates $\left(x^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}\right)$. As an example, the anticommutator (5.3) becomes

$$[\theta Q, \bar{\theta}\bar{Q}] = 2\theta^\alpha \sigma^\mu_{\alpha\dot{\alpha}} \bar{\theta}^{\dot{\alpha}} P_\mu.$$

We can now write an element of the corresponding superPoincaré group as

$$G\left(x^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}, \omega^{\mu\nu}\right) = \exp\left(i x^\mu P_\mu + i\theta^\alpha Q_\alpha - i\bar{\theta}^{\dot{\alpha}} Q_{\dot{\alpha}} + \tfrac{1}{2} i \omega^{\mu\nu} M_{\mu\nu}\right).$$

In chapter 4, we made use of the group property (4.7), which we shall now derive with the help of the Baker-Campbell-Hausdorff formula [14]:

$$\exp(A)\exp(B) = \exp\left(A + B + \tfrac{1}{2}[A,B] + \tfrac{1}{12}[A,[A,B]] + \tfrac{1}{12}[B,[B,A]] + O\left(A^2B^2\right)\right).$$

[2]Here, we used the fact that $i\left(\sigma_{\mu\nu}\right)_\alpha{}^\beta X_\beta = \frac{i}{4}\left(\sigma^\mu_{\alpha\dot{\gamma}}\bar{\sigma}^{\nu,\dot{\gamma}\beta} - \sigma^\nu_{\alpha\dot{\gamma}}\bar{\sigma}^{\mu,\dot{\gamma}\beta}\right) X_\beta = \pm\frac{1}{2} X_\alpha$.

We find that the product of $G(y,\xi,\bar{\xi},\omega)$ and $G(x,\theta,\bar{\theta},0)$ up to second order is

$$\begin{aligned}&\exp\left(iy^\mu P_\mu + i\xi^\alpha Q_\alpha + i\bar{\xi}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}} + \tfrac{1}{2}i\omega^{\rho\sigma}M_{\rho\sigma}\right)\exp\left(ix^\mu P_\mu + i\theta^\alpha Q_\alpha + i\bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}\right)\\ =&\exp\Big(i\left(y^\mu + x^\mu\right)P_\mu + i\left(\xi^\alpha+\theta^\alpha\right)Q_\alpha + \ i\left(\bar{\xi}_{\dot{\alpha}} + \bar{\theta}_{\dot{\alpha}}\right)\bar{Q}^{\dot{\alpha}} - \tfrac{1}{2}\left[\bar{\xi}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}},\theta^\alpha Q_\alpha\right]\\ &-\tfrac{1}{2}\left[\xi^\alpha Q_\alpha, \bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}\right] - \tfrac{1}{4}\left[\omega^{\rho\sigma}M_{\rho\sigma}, x^\mu P_\mu\right] - \tfrac{1}{4}\left[\omega^{\rho\sigma}M_{\rho\sigma},\theta^\alpha Q_\alpha\right] - \tfrac{1}{4}\left[\omega^{\rho\sigma}M_{\rho\sigma},\bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}\right]\Big),\end{aligned}$$

where the other commutators $[P,Q]=[\theta Q,\xi Q]=[\bar{\theta}\bar{Q},\bar{\xi}\bar{Q}]=0$ vanish. Hence, if $\omega^{\rho\sigma}=0$, then the equality is exact to all orders. The presence of commutators involving M means that there are terms involving Q and $\bar{Q}$ at all orders. This follows from the commutators

$$\begin{aligned}\left[\xi^\alpha Q_\alpha, \bar{\xi}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}\right] &= 2\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}P_\mu,\\ \left[\omega^{\rho\sigma}M_{\rho\sigma}, x^\mu P_\mu\right] &= ix_\nu\left(\omega^{\mu\nu}-\omega^{\nu\mu}\right)P_\mu,\\ \left[\omega^{\rho\sigma}M_{\rho\sigma},\theta^\alpha Q_\alpha\right] &= i\omega^{\rho\sigma}\xi^\beta\left(\sigma_{\rho\sigma}\right)_\beta{}^\alpha Q_\alpha,\\ \left[\omega^{\rho\sigma}M_{\rho\sigma},\bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}\right] &= i\omega^{\rho\sigma}\bar{\xi}_{\dot{\beta}}\left(\bar{\sigma}_{\mu\nu}\right)^{\dot{\beta}}{}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}},\end{aligned}$$

which are based on the commutators and anticommutators in (5.2) and (5.3). These can be used to evaluate our exponential:

$$\begin{aligned}&\exp\left(iy^\mu P_\mu + i\xi^\alpha Q_\alpha + i\bar{\xi}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}} + \tfrac{1}{2}i\omega^{\rho\sigma}M_{\rho\sigma}\right)\exp\left(ix^\mu P_\mu + i\theta^\alpha Q_\alpha + i\bar{\theta}_{\dot{\alpha}}\bar{Q}^{\dot{\alpha}}\right)\\ =&\exp\Big(i\left(y^\mu + x^\mu + i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}} - i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} + \tfrac{1}{4}x_\nu\left(\omega^{\nu\mu}-\omega^{\mu\nu}\right)\right)P_\mu\\ &+ i\left(\xi^\alpha + \theta^\alpha - \tfrac{1}{4}\omega^{\rho\sigma}\xi^\beta\left(\sigma_{\rho\sigma}\right)_\beta{}^\alpha\right)Q_\alpha + i\left(\bar{\xi}_{\dot{\alpha}} + \bar{\theta}_{\dot{\alpha}} - \tfrac{1}{4}\omega^{\rho\sigma}\bar{\xi}_{\dot{\beta}}\left(\bar{\sigma}_{\mu\nu}\right)^{\dot{\beta}}{}_{\dot{\alpha}}\right)\bar{Q}^{\dot{\alpha}}\Big),\end{aligned}$$

from which we read off the desired result:

$$G\begin{pmatrix} y^\mu \\ \xi^\alpha \\ \bar{\xi}^{\dot{\alpha}} \\ \omega^{\rho\sigma}\end{pmatrix}^{\mathbf{T}} G\begin{pmatrix} x^\mu \\ \theta^\alpha \\ \bar{\theta}^{\dot{\alpha}} \\ 0\end{pmatrix}^{\mathbf{T}} = G\begin{pmatrix} y^\mu + x^\mu + i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}} - i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} + \tfrac{1}{4}x_\nu\left(\omega^{\nu\mu}-\omega^{\mu\nu}\right) \\ \xi^\alpha + \theta^\alpha - \tfrac{1}{4}\omega^{\rho\sigma}\xi^\beta\left(\sigma_{\rho\sigma}\right)_\beta{}^\alpha \\ \bar{\xi}^{\dot{\alpha}} + \bar{\theta}^{\dot{\alpha}} - \tfrac{1}{4}\omega^{\rho\sigma}\bar{\xi}_{\dot{\beta}}\left(\bar{\sigma}_{\mu\nu}\right)^{\dot{\beta}}{}_{\dot{\alpha}} \\ 0\end{pmatrix}^{\mathbf{T}}.$$

Often, we are only interested in supersymmetry transformations, so we set $y^\mu=\omega^{\rho\sigma}=0$, which yields the much simpler property:

$$G\left(0,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},0\right)G\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}},0\right) = G\left(x^\mu + i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}} - i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}, \xi^\alpha+\theta^\alpha, \bar{\xi}^{\dot{\alpha}} + \bar{\theta}^{\dot{\alpha}}, 0\right).$$

In chapter 4, we used these group properties to define linear representations of the super-Poincaré group. These are the superfields.

Superfields

Superfields are linear representations of the $\mathcal{N}=1$ superPoincaré group, but we shall first consider them simply as superfunctions. We saw in chapter 4 that superfunctions can be expressed as a series expansion in $\theta, \bar{\theta}$. We thus write

$$\begin{aligned} Y(x,\theta,\bar{\theta}) = f(x) + \theta\psi(x) + \bar{\theta}\bar{\chi}(x) + \theta\theta m(x) + \bar{\theta}\bar{\theta}n(x) + \theta\sigma^\mu\bar{\theta}v_\mu(x) \\ + \theta\theta\bar{\theta}\bar{\lambda}(x) + \bar{\theta}\bar{\theta}\theta\rho(x) + \theta\theta\bar{\theta}\bar{\theta}d(x). \end{aligned} \tag{5.4}$$

We stress once more that the coefficient functions, now interpreted as component fields, only depend on the even coordinates x^μ. The odd coordinates $\theta^\alpha, \bar{\theta}^{\dot{\alpha}}$ have no physical meaning. The Lorentz properties of the superfield $Y(x,\theta,\bar{\theta})$ are those of the lowest component field $f(x)$. The properties of the other components then follow naturally. For example, if $f(x)$ is a scalar, then $\psi(x), \bar{\chi}(x), \bar{\lambda}(x), \rho(x)$ are Weyl spinors, $m(x), n(x), d(x)$ are scalars and $v_\mu(x)$ is a vector.

In chapter 4, we postulated an action of the superPoincaré group on $(4,4)$-dimensional superspace $\mathcal{G}_L^{4,4}$ and a corresponding action on the space of smooth superfields $G^\infty\left(\mathcal{G}_L^{4,4}\right)$. The latter action, for infinitesimal $\xi^\alpha, \bar{\xi}^{\dot{\alpha}}$,

$$G\left(0,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},0\right)\cdot Y\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right) = Y\left(x^\mu + i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}, \theta^\alpha+\xi^\alpha, \bar{\theta}^{\dot{\alpha}}+\bar{\xi}^{\dot{\alpha}}\right),$$

is a linear representation. In terms of the differential operators

$$\begin{aligned} Q_\alpha &= -i\frac{\partial}{\partial\theta^\alpha} - \sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\frac{\partial}{\partial x^\mu}, \\ \bar{Q}_{\dot{\alpha}} &= i\frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}} + \theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\frac{\partial}{\partial x^\mu}, \end{aligned}$$

the action becomes

$$G\left(0,\xi^\alpha,\bar{\xi}^{\dot{\alpha}},0\right)\cdot Y\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right) = Y\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right) + i\left(\xi^\alpha Q_\alpha - \bar{\xi}^{\dot{\alpha}}\bar{Q}_{\dot{\alpha}}\right)Y\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right).$$

Chiral superfields

Superfields are useful, but the most general representation (5.4) is highly reducible. A way out of this is to postulate some supersymmetric constraint which eliminates a number of the component fields. A common choice is

$$D_\alpha\Phi\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right) = 0, \qquad \text{or} \qquad \bar{D}_{\dot{\alpha}}\Phi\left(x^\mu,\theta^\alpha,\bar{\theta}^{\dot{\alpha}}\right) = 0, \tag{5.5}$$

where $D_\alpha, \bar{D}_{\dot{\alpha}}$ are an alternative choice of differential operators to $Q_\alpha, \bar{Q}_{\dot{\alpha}}$:

$$\begin{aligned} D_\alpha &= \frac{\partial}{\partial\theta^\alpha} + i\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}}\frac{\partial}{\partial x^\mu}, \\ \bar{D}_{\dot{\alpha}} &= \frac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}} + i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\frac{\partial}{\partial x^\mu}. \end{aligned}$$

These operators anticommute with $Q_\alpha, \bar{Q}_\alpha$, which implies that if (5.5) holds for Φ, then it also holds for $\Phi' = \Phi + i(\xi Q + \bar{\xi}\bar{Q})\Phi$. Superfields which satisfy the condition $\bar{D}_{\dot{\alpha}}\Phi = 0$ are called *chiral superfields*. Those that satisfy $D_\alpha \Phi = 0$ are *anti-chiral*. In particular, if Φ is chiral, then $\bar{\Phi}$ is anti-chiral. To find the general component expression of a chiral superfield $\Phi\left(x^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}\right)$, it is prudent to define new coordinates

$$
\begin{aligned}
y^\mu &= x^\mu + i\theta^\alpha \sigma^\mu_{\alpha\dot{\alpha}} \bar{\theta}^{\dot{\alpha}}, \\
\bar{y}^\mu &= x^\mu - i\theta^\alpha \sigma^\mu_{\alpha\dot{\alpha}} \bar{\theta}^{\dot{\alpha}}.
\end{aligned}
$$

These coordinates are useful, because

$$
\bar{D}_{\dot{\alpha}} \bar{y}^\mu = \left(\frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}} + i\theta^\alpha \sigma^\nu_{\alpha\dot{\alpha}} \frac{\partial}{\partial x^\nu} \right) \left(x^\mu - i\theta^\beta \sigma^\mu_{\beta\dot{\beta}} \bar{\theta}^{\dot{\beta}} \right) = i\theta^\alpha \sigma^\nu_{\alpha\dot{\alpha}} \delta^\mu_\nu - i\theta^\beta \sigma^\mu_{\beta\dot{\beta}} \delta^{\dot{\alpha}}_{\dot{\beta}} = 0.
$$

Similarly, $D_\alpha y^\mu = D_\alpha \bar{\theta}^{\dot{\beta}} = \bar{D}_{\dot{\alpha}} \theta^\alpha = 0$. It follows that the condition $\bar{D}_{\dot{\alpha}} \Phi\left(y^\mu, \bar{y}^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}\right) = 0$ kills all terms that depend explicitly on y^μ and $\bar{\theta}^{\dot{\alpha}}$. As a result, Φ is only explicitly a function of $\bar{y}^\mu$ and θ^α:

$$
\begin{aligned}
\Phi\left(\bar{y}^\mu, \theta^\alpha\right) &= \phi\left(\bar{y}^\mu\right) + \sqrt{2}\theta^\alpha \psi_\alpha\left(\bar{y}^\mu\right) + \theta\theta F\left(\bar{y}^\mu\right) \\
&= \phi\left(x^\mu\right) + \sqrt{2}\theta^\alpha \psi_\alpha\left(x^\mu\right) + \theta\theta F\left(x^\mu\right) - i\theta\sigma^\mu\bar{\theta}\partial_\mu \phi\left(x^\mu\right) \\
&\quad - i\sqrt{2}\theta\theta\partial_\mu \psi_\alpha\left(x^\mu\right) \sigma^\mu_{\alpha\dot{\alpha}} \bar{\theta}^{\dot{\alpha}} - \tfrac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\Box\phi\left(x^\mu\right),
\end{aligned}
$$

where the $\sqrt{2}$ is standard. It is now easy to derive the analogous expression for anti-chiral superfields. The expansion above shows that a chiral superfield has the same degrees of freedom[3] as a chiral supermultiplet: a complex scalar $\phi(x)$, a 2-component complex spinor $\psi(x)$, and another complex "auxiliary" scalar field $F(x)$. The reason that $F(x)$ is called the auxiliary field is that it is a non-propagating field, as we will see. The equation of motion of F is algebraic and can be substituted into the Lagrangian to eliminate F from the theory.

Let us now determine the effect of a supersymmetry transformation on $\Phi = (\phi, \psi, F)$. In terms of the alternative coordinates $\left(\bar{y}^\mu, \theta^\alpha, \bar{\theta}^{\dot{\alpha}}\right)$, the differential operators $Q_\alpha, \bar{Q}_{\dot{\alpha}}$ read

$$
\begin{aligned}
Q_\alpha &= -i\frac{\partial}{\partial \theta^\alpha}, \\
\bar{Q}_{\dot{\alpha}} &= i\frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}} + 2\theta^\alpha \sigma^\mu_{\alpha\dot{\alpha}} \frac{\partial}{\partial \bar{y}^\mu}.
\end{aligned}
$$

By substituting these into the variation equation $\delta\Phi = i(\xi Q + \bar{\xi}\bar{Q})\Phi$, one finds

$$
\delta\Phi = \sqrt{2}\xi\psi + \sqrt{2}\theta\left(\sqrt{2}\xi F + \sqrt{2}i\sigma^\mu\bar{\xi}\frac{\partial\phi}{\partial\bar{y}^\mu}\right) - \theta\theta\left(i\sqrt{2}\frac{\partial\psi^\alpha}{\partial\bar{y}^\mu}\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\right),
$$

where we used the identity $(\theta\sigma^\mu\bar{\phi})(\theta\psi) = -\frac{1}{2}(\psi\sigma^\mu\bar{\phi})(\theta\theta)$ for arbitrary Weyl spinors $\theta, \psi, \bar{\phi}$.

[3]In general, supermultiplets should contain the same number of bosonic and fermionic degrees of freedom. The chiral superfield clearly satisfies this condition.

Hence, component-wise the variations are

$$\begin{aligned}\delta\phi &= \sqrt{2}\xi^\alpha\psi_\alpha,\\ \delta\psi_\alpha &= \sqrt{2}i\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha}\partial_\mu\phi + \sqrt{2}\xi_\alpha F,\\ \delta F &= -i\sqrt{2}\partial_\mu\psi^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha}.\end{aligned}$$

In practice, many other types of superfields are used besides chiral superfields. The *vector superfields* constitute an important class. As the name suggests, vector superfields carry a vector component v_μ. Vector superfields are defined by the reality condition $V = \bar V$. Hence, if $\Phi = (\phi, \psi, F)$ is a chiral superfield, then $\Phi + \bar\Phi$ is a vector superfield. It turns out that under a transformation $V \to V + \Phi + \bar\Phi$, the vector component transforms similarly to a gauge transformation $v_\mu \to v_\mu - \partial_\mu\,(\mathrm{Im}\ \phi)$. This is why vector superfields are suitable for supersymmetric gauge theories. Another type of superfield is the *Ferrara-Zumino multiplet* $\mathcal{J}_\mu$. One of its component fields is the supercurrent $S_{\alpha\mu}$ associated with unbroken supersymmetry. We will meet this superfield later in chapter 6.

Supersymmetric actions

The general action for n chiral superfields $\Phi^i = (\phi^i, \psi^i, F^i)$ is

$$S = \int \mathrm{d}^4x\,\mathcal{L} = \int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\overline\theta\ K(\Phi^i,\overline\Phi_i) + \int \mathrm{d}^4x\,\mathrm{d}^2\theta\ W(\Phi^i) + \int \mathrm{d}^4x\,\mathrm{d}^2\overline\theta\ \overline{W}(\overline\Phi_i), \tag{5.6}$$

where $K(\Phi^i, \bar\Phi_i)$ is a real function called the *Kähler potential* and the holomorphic function $W(\Phi^i)$ is called the *superpotential*. It is easy to define Grassmann integration as a formal operation, similar to the way Grassmann differentiation was defined in chapter 4. However, it is sufficient to note that the integral $\int \mathrm{d}^2\theta\mathrm{d}^2\bar\theta$ simply picks out the $\theta^2\bar\theta^2$-component of the integrand, whereas the effect of $\int \mathrm{d}^2\theta$ and $\int \mathrm{d}^2\bar\theta$ is to pick out the θ^2- and $\bar\theta^2$-components, respectively. We will only encounter these three cases.

The typical example is the Wess-Zumino model of $n = 1$ chiral superfield: [10]

$$K = \Phi\bar\Phi, \qquad W = \tfrac{1}{2}m\Phi^2 + \tfrac{1}{3}\lambda\Phi^3, \tag{5.7}$$

with m and λ two real parameters. Substituting this into the Lagrangian, we find

$$\mathcal{L} = \int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\overline\theta\ \Phi\bar\Phi + \int \mathrm{d}^4x\,\mathrm{d}^2\theta\ \left(\tfrac{1}{2}m\Phi^2 + \tfrac{1}{3}\lambda\Phi^3\right) + \int \mathrm{d}^4x\,\mathrm{d}^2\overline\theta\ \left(\tfrac{1}{2}m\bar\Phi^2 + \tfrac{1}{3}\lambda\bar\Phi^3\right).$$

The easiest way to evaluate the component form of the Lagrangian is two write Φ in terms of the alternative coordinates $(\bar y^\mu, \theta^\alpha)$. This allows one to easily work out terms like $\Phi\bar\Phi$ and Φ^2. One should then expand the expression in a power series around $y^\mu = x^\mu$ to obtain the full expression in $(x^\mu, \theta^\alpha, \bar\theta^\alpha)$ coordinates. Of course, we only have to consider the $\theta^2\bar\theta^2$-

component of $K(\Phi, \bar{\Phi})$ and the θ^2-component of $W(\Phi)$. Doing this for the Kähler potential, we find

$$\mathcal{L}_{\mathrm{kin}} = \int \mathrm{d}^4x\, \mathrm{d}^2\theta\, \mathrm{d}^2\overline{\theta}\; K(\Phi, \overline{\Phi}) = \partial_\mu\bar{\phi}\, \partial^\mu\phi + \tfrac{1}{2}i\left(\partial_\mu\psi\sigma^\mu\bar{\psi} - \psi\sigma^\mu\partial_\mu\bar{\psi}\right) + \bar{F}F,$$

up to a total derivative term which we ignore. These are precisely the kinetic terms for a complex scalar field ϕ, a spinor ψ, and a non-propagating field F. For the superpotential, we do the same and find

$$\begin{aligned}\mathcal{L}_{\mathrm{int}} &= \int \mathrm{d}^4x\, \mathrm{d}^2\theta\; W(\Phi^i) + \int \mathrm{d}^4x\, \mathrm{d}^2\overline{\theta}\; \overline{W}(\overline{\Phi}_i) \\ &= \left(m\phi + \lambda\phi^2\right)F - \left(\tfrac{1}{2}m + \lambda\phi\right)\psi\psi + \left(m\bar{\phi} + \lambda\bar{\phi}^2\right)\bar{F} - \left(\tfrac{1}{2}m + \lambda\bar{\phi}\right)\bar{\psi}\bar{\psi}.\end{aligned}$$

The total Lagrangian $\mathcal{L} = \mathcal{L}_{\mathrm{kin}} + \mathcal{L}_{\mathrm{int}}$ is supersymmetric invariant (up to a total derivative). This is not easy to see in component form, but any integral of a superfield in superspace is invariant under supersymmetry transformations. This follows from the fact that the variation of a superfunction $f(x, \theta, \overline{\theta})$ is

$$\delta f(x, \theta, \overline{\theta}) = \xi^\alpha\partial_\alpha f + \overline{\xi}_{\dot{\alpha}}\partial^{\dot{\alpha}} f + \partial_\mu\left(-i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}} f + i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\overline{\xi}^{\dot{\alpha}} f\right),$$

which becomes a total derivative when integrated over superspace, because the $\theta^2\bar{\theta}^2$-components of the first two terms vanish. This is the power of the superfield formalism. Let us now eliminate the auxiliary field. The equation of motion for F is

$$0 = \frac{\partial\mathcal{L}}{\partial F} = \bar{F} + m\phi + \lambda\phi^2.$$

We eliminate F by substituting this back into the Lagrangian:

$$\begin{aligned}\mathcal{L} = \partial_\mu\phi\, \partial^\mu\bar{\phi} &+ \tfrac{1}{2}i\left(\partial_\mu\psi\sigma^\mu\bar{\psi} - \psi\sigma^\mu\partial_\mu\bar{\psi}\right) - \left(\tfrac{1}{2}m + \lambda\phi\right)\psi\psi \\ &- \left(\tfrac{1}{2}m + \lambda\bar{\phi}\right)\bar{\psi}\bar{\psi} - \left|m\phi + \lambda\phi^2\right|^2.\end{aligned}$$

By starting with the simple superfield Lagrangian (5.7) and working out the component expression, we have obtained a theory of a massive scalar ϕ and a massive spinor ψ with ϕ^4 and Yukawa interactions.

Let us now turn to the more general case, still with $n = 1$ chiral superfield, but with unspecified Kähler and superpotentials. The most general Kähler potential is

$$K(\Phi, \bar{\Phi}) = \sum_{m=1,n=1}^{\infty} c_{mn}\bar{\Phi}^m\Phi^n, \qquad c_{mn} = c^*_{mn}.$$

Because the mass dimension of a Lagrangian should be 4, and the mass dimension of $\int \mathrm{d}^2\theta \mathrm{d}^2\bar{\theta}$ is 2, we see that the mass dimension of K must be 2. Moreover, because the dimension of Φ

is 1, the coefficients c_{mn} for $m+n>2$ have negative mass dimensions. The coefficients are therefore usually suppressed by the cut-off scale Λ of the theory:

$$c_{mn} \simeq \Lambda^{2-(m+n)}.$$

If the theory is to be renormalisable, then all such coefficients with $m+n>2$ should vanish and we end up with the canonical Kähler potential $K=\Phi\bar{\Phi}$. Meanwhile, the most general superpotential can be written as

$$W(\Phi) = \sum_{n=1}^{\infty} a_n \Phi^n.$$

There is an easy way to evaluate the integral of the superpotential, by expanding it in a power series around $\Phi = \phi$:

$$W(\Phi) = W(\phi) + \sqrt{2}\theta \frac{\partial W}{\partial \phi}\psi + \theta\theta \left(\frac{\partial W}{\partial \phi} F - \frac{1}{2}\frac{\partial^2 W}{\partial \phi^2}\psi\psi \right),$$

where the partial derivatives are evaluated at $W(\Phi) = W(\phi)$. The Grassmann integral only picks up the θ^2-term. Hence,

$$\mathcal{L}_{\text{int}} = \int \mathrm{d}^2\theta W(\Phi) + \int \mathrm{d}^2\overline{\theta W}(\bar{\Phi}) = \frac{\partial W}{\partial \phi} F - \frac{1}{2}\frac{\partial^2 W}{\partial \phi^2}\psi\psi + \text{h.c.},$$

where h.c. refers to the Hermitian conjugate of the preceding terms. If we now assume that the Kähler potential has the canonical form $K(\Phi, \bar{\Phi}) = \Phi\bar{\Phi}$, but leave W unspecified, then we obtain the Lagrangian

$$\begin{aligned} \mathcal{L} &= \int \mathrm{d}^4x \, \mathrm{d}^2\theta \, \mathrm{d}^2\bar{\theta} \ \Phi\bar{\Phi} + \int \mathrm{d}^2\theta W(\Phi) + \int \mathrm{d}^2\overline{\theta W}(\bar{\Phi}) \\ &= \partial_\mu \bar{\phi} \, \partial^\mu \phi + \tfrac{1}{2} i \left(\partial_\mu \psi \sigma^\mu \bar{\psi} - \psi \sigma^\mu \partial_\mu \bar{\psi} \right) + \bar{F}F + \frac{\partial W}{\partial \phi} F - \frac{1}{2}\frac{\partial^2 W}{\partial \phi^2}\psi\psi + \text{h.c.} \end{aligned}$$

We follow the same procedure as for the Wess-Zumino model and eliminate the auxiliary field F. The equation of motion for F is

$$\bar{F} = -\frac{\partial W}{\partial \phi}.$$

Substituting this back into the Lagrangian, we then find

$$\mathcal{L} = \partial_\mu \bar{\phi} \, \partial^\mu \phi + \tfrac{1}{2} i \left(\partial_\mu \psi \sigma^\mu \bar{\psi} - \psi \sigma^\mu \partial_\mu \bar{\psi} \right) - \left| \frac{\partial W}{\partial \phi} \right|^2 - \frac{1}{2}\frac{\partial^2 W}{\partial \phi^2}\psi\psi - \frac{1}{2}\frac{\partial^2 \overline{W}}{\partial \bar{\phi}^2}\bar{\psi}\bar{\psi}. \tag{5.8}$$

Clearly, the scalar potential of this theory is

$$V(\phi, \bar{\phi}) = \left| \frac{\partial W}{\partial \phi} \right|^2 = \bar{F}F.$$

Generalising this to the case for n chiral superfields, we have to consider the full action (5.6). If the Kähler potential is in its canonical "diagonalised" form

$$K\left(\Phi^i, \bar{\Phi}_i\right) = \Phi^i \bar{\Phi}_i,$$

then the scalar potential will turn out to be

$$V\left(\phi^i, \bar{\phi}_i\right) = \sum_{i=1}^{n} \left|\frac{\partial W}{\partial \phi^i}\right|^2 = \bar{F}_i F^i, \qquad \bar{F}_i = -\frac{\partial W}{\partial \bar{\phi}_i}, \quad F^i = -\frac{\partial W}{\partial \phi^i}.$$

We could say a lot more about supersymmetric Lagrangians, but we should now turn to a discussion of supersymmetry breaking.

5.3 Spontaneous supersymmetry breaking

In section 5.1, we noted that supersymmetry breaking is necessary in order to account for the fact that superparticles do not have the same mass as their Standard Model partners. This can either be modelled through spontaneous or explicit symmetry breaking. In the latter case, the theory is usually an effective theory of a theory with spontaneous supersymmetry breaking [8]. Hence, we focus on the former scenario. Spontaneous supersymmetry breaking does not differ in principle from ordinary spontaneous symmetry breaking. Analogous to what we saw in chapter 2, spontaneous supersymemtry breaking occurs when the stable vacuum of the theory is not invariant under superPoincaré transformations. What sets supersymmetry breaking apart is the role of the scalar potential, as we shall see.

Things simplify considerably, when we demand that Poincaré symmetry remains unbroken. This implies that only the scalar fields $\phi^i, \bar{\phi}_i, F^i, \bar{F}_i$ can have a nonzero vacuum expectation value. Otherwise, some field would obtain a preferred direction, which contradicts Lorentz invariance. Furthermore, kinetic terms in the Lagrangian become irrelevant, because the vacuum expectation values must be constant. This means that we can focus entirely on the scalar potential $V(\phi^i, \bar{\phi}_i)$. The equation of motion of the scalar fields, evaluated on the vacuum, becomes

$$\left\langle \frac{\partial V}{\partial \phi^i} \right\rangle = \left\langle \frac{\partial V}{\partial \bar{\phi}_i} \right\rangle = 0. \tag{5.9}$$

For models with canonical Kähler potentials, we saw in the previous section that the scalar potential is given by

$$V = \bar{F}_i F^i, \qquad \bar{F}_i = -\frac{\partial W}{\partial \bar{\phi}_i}, \qquad F^i = -\frac{\partial W}{\partial \phi^i}. \tag{5.10}$$

In general, even for non-canonical Kähler potentials, the scalar potential can be brought into this form; at least when evaluated on the vacuum. Combining (5.9) and (5.10), we find

$$\left\langle \bar{F}_i \frac{\partial F^i}{\partial \phi^j} \right\rangle = -\left\langle \bar{F}_i \frac{\partial W}{\partial \phi^i \partial \phi^j} \right\rangle = 0. \tag{5.11}$$

The second derivative W_{ij} of the superpotential is proportional to the mass^2 coefficient of the fermions $\psi^i\psi^j$. See for example (5.8). Hence, the fermion mass matrix is

$$\mathcal{M}_{ij} = -\frac{1}{2}\frac{\partial W}{\partial \phi^i \partial \phi^j}.$$

Now, suppose that $\langle \bar{F}_i \rangle \neq 0$ for some i. Then, (5.11) implies that $\mathcal{M}_{ij}$ must have an eigenvector with zero eigenvalue:

$$\lambda_\alpha \propto \langle \bar{F}_i \rangle \psi^i_\alpha, \qquad \mathcal{M}_{ij}\lambda_\alpha = 0.$$

This direction represents a massless spin-$\frac{1}{2}$ particle, which is analogous to the Goldstone bosons of the bosonic theories of chapter 2. In this case, the massless particle is called a *goldstino.*

5.4 Nilpotency constraints

Nilpotency constraints are instrumental to our study of non-linear realisations of supersymmetry in chapter 6. To prepare for this, let us first recall some facts about the simplest nilpotency condition $\Phi^2 = 0$. We let $\Phi = (\phi, \psi_\alpha, F)$ denote a chiral multiplet consisting of a complex scalar ϕ, a Weyl spinor ψ_α, and an auxiliary field F. In terms of components, the condition $\Phi^2 = 0$ reads

$$\Phi^2 = \phi^2 + 2\sqrt{2}\theta^\alpha (\phi\psi_\alpha) + \theta\theta (2\phi F - \psi\psi) = 0, \tag{5.12}$$

where we made use of the Fierz identity $(\theta\phi)(\theta\psi) = -\frac{1}{2}(\phi\psi)(\theta\theta)$ for arbitrary spinors ϕ, ψ, θ. Of course, the trivial solution is $\phi = F = \psi = 0$, which we henceforth ignore. The constraint is generally solved by setting

$$\phi = \frac{\psi\psi}{2F}, \tag{5.13}$$

provided that $F \neq 0$. Clearly, the last component of (5.12) then vanishes. Furthermore, because $\phi^2 \propto \psi^4 = 0$ and $\phi\psi \propto \psi^3 = 0$, it follows that the first two components vanish as well. The field F in the denominator can be problematic, but may be dealt with by means of non-Gaussian integration [7, 46]. Simply put, one may use the equation of motion of F to set

$$F = f + O\left(\psi^2\right) \quad \Longrightarrow \quad \frac{1}{F} = \frac{1}{f} - O\left(\psi^2\right). \tag{5.14}$$

Nevertheless, on this view, the validity of (5.13) critically depends on the equation of motion of F. It is not true that computing ϕ^2 on the basis of (5.13) is valid on the assumption $F \neq 0$ alone. Indeed, suppose that $F \propto \psi\psi$. In that case, $F^2 = 0$ and it is certainly not the case that $\Phi^2 = 0$. In what follows, we assume that (5.12) and (5.13) are equivalent, as in [15, 47, 46, 32]. This follows from the requirement that $F^n \neq 0$ for all integers n. Fortunately, this is a supersymmetric invariant condition, as we now show.

Lemma 5.1. *The condition $F^n \neq 0$ for $n \in \mathbb{Z}$ on the auxiliary field F is SUSY-invariant.*

Proof. The auxiliary field F transforms under a supersymmetry transformation as

$$F \to F + \delta F,$$

with $\delta F = -i\sqrt{2}\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha}\partial_\mu\psi$. The quantity δF satisfies $(\delta F)^3 = 0$ due to the nilpotency of $\bar\xi$. The desired conclusion follows if we can show that $(F+\delta F)^n \neq 0$ for $n \in \mathbb{Z}$. The case $n = 0$ is immediate. For $n > 0$, assume first that the converse is true, namely that $(F+\delta F)^n = 0$. Then, we find that

$$0 = (\delta F)^2(F+\delta F)^n = (\delta F)^2 F^n,$$

where all other terms vanish due to the nilpotency of δF. Because F^n is not nilpotent, it has an inverse by Lemma 4.2. Right-multiplication with $(F^n)^{-1}$ yields $0 = (\delta F)^2$, which is a contradiction. It follows that $(F+\delta F)^n$ is nonzero for $n \geq 0$. Clearly then, the same holds for $1/(F+\delta F)^n$. □

Next, we should investigate whether the nilpotency constraint is supersymmetric invariant. This is a requirement for the application of the Non-Linearisation Lemma of section 3.4. To see that it is invariant, consider an arbitrary supersymmetry transformation on Φ. In components, this reads

$$\begin{aligned}
\delta\phi &= \sqrt{2}\xi^\alpha\psi_\alpha, \\
\delta\psi_\alpha &= \sqrt{2}i\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha}\partial_\mu\phi + \sqrt{2}\xi_\alpha F, \\
\delta F &= -i\sqrt{2}\partial_\mu\psi\sigma^\mu\bar\xi.
\end{aligned}$$

To work out the effect on $2\phi F - \psi\psi$, we proceed as follows:

$$\begin{aligned}
\delta(2\phi F - \psi\psi) &= 2\delta(\phi)F + 2\phi\delta F - 2\psi\delta(\psi) \\
&= 2\sqrt{2}\xi\psi F - 2\sqrt{2}i\phi\left(\partial_\mu\psi\right)\sigma^\mu\bar\xi - 2\sqrt{2}i\psi\sigma^\mu\bar\xi\left(\partial_\mu\phi\right) - 2\sqrt{2}\psi\xi F \\
&= -2\sqrt{2}i\partial_\mu\left(\phi\psi\right)\sigma^\mu\bar\xi,
\end{aligned} \tag{5.15}$$

where we used that $\delta\left(\psi^\alpha\psi_\alpha\right) = 2\psi^\alpha\delta\psi_\alpha$. Finally, because $\phi\psi = 0$, we see that $2\phi F - \psi\psi$ is supersymmetric invariant. It follows that the constraint $\Phi^2 = 0$ is also invariant.

Now, we turn to higher order nilpotency constraints. We derive the following expression for Φ^n in terms of components:

$$\Phi^n = \phi^n + \sqrt{2}\theta\left(n\psi\phi^{n-1}\right) + \theta\theta\left(n\phi^{n-1}F - a(n)\phi^{n-2}\psi\psi\right), \tag{5.16}$$

$$a(n) = \frac{n(n-1)}{2}. \tag{5.17}$$

To see that (5.17) is correct, consider only the terms proportional to $\theta\theta$ in the expansion

$$\begin{aligned} \Phi\Phi^n\big|_{\theta\theta} &= \left(\phi + \sqrt{2}\theta\psi + \theta\theta F\right) \times \left(\phi^n + \sqrt{2}\theta\left(n\psi\phi^{n-1}\right) + \theta\theta\left(n\phi^{n-1}F - a(n)\phi^{n-2}\psi\psi\right)\right)\Big|_{\theta\theta} \\ &= \theta\theta F\phi^n + \phi\theta\theta\left(n\phi^{n-1}F - a(n)\phi^{n-2}\psi\psi\right) + 2\theta\psi\theta\left(n\psi\phi^{n-1}\right) \\ &= \theta\theta\left[(n+1)\phi^n F - (a(n)+n)\,\phi^{n-1}\psi\psi\right], \end{aligned}$$

where we again used the identity $(\theta\phi)(\theta\psi) = -\frac{1}{2}(\phi\psi)(\theta\theta)$. From (5.12), we read off that $a(2) = 1$. Equation (5.17) then follows from the fact that $a(n+1) = a(n) + n$. We now consider whether the $\theta\theta$-component of the condition $\Phi^n = 0$ implies the that the other components vanish as well, since this is the case for $\Phi^2 = 0$. We may write the last component suggestively as

$$\theta\theta\phi^{n-2}\left[n\phi F - a(n)\psi\psi\right] = 0. \tag{5.18}$$

Unfortunately, the condition $\gamma\phi F - \psi\psi = 0$ (for fixed γ) is only supersymmetric invariant for $\gamma = 2$. This can be seen by exchanging 2 with γ in the derivation of (5.15):

$$\begin{aligned} \delta(\gamma\phi F - \psi\psi) &= \gamma\delta(\phi)F + \gamma\phi\delta(F) - 2\delta(\psi)\psi \\ &= (\gamma-2)\sqrt{2}\xi\psi F + (\gamma-2)\sqrt{2}i\partial_\mu\left(\phi\psi\right)\sigma^\mu\bar{\xi}. \end{aligned}$$

Nevertheless, the full condition $\Phi^n = 0$ is still supersymmetric invariant for $n \geq 3$. This is a consequence of the factor ϕ^{n-2} in front of the square brakets in (5.18). A more complicated version of the proof for $\Phi^2 = 0$ also applies to $\Phi^3 = 0$.

Lemma 5.2. *The condition $\Phi^3 = 0$ is supersymmetric invariant.*

Proof. We want to show that the constraint

$$\Phi^3 = \phi^3 + \sqrt{2}\theta^\alpha\left(3\psi_\alpha\phi^2\right) + \theta\theta\left(3\phi^2 F - 3\phi\psi\psi\right) = 0$$

is invariant under the transformations

$$\begin{aligned} \delta\phi &= \sqrt{2}\xi^\alpha\psi_\alpha, \\ \delta\psi_\alpha &= \sqrt{2}i\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\partial_\mu\phi + \sqrt{2}\xi_\alpha F, \\ \delta F &= -i\sqrt{2}\partial_\mu\psi\sigma^\mu\bar{\xi}. \end{aligned}$$

First, consider the lowest order component $\phi^3 = 0$. The variation is

$$\delta\left(\phi^3\right) = 3\phi^2\delta\phi = 3\phi^2\sqrt{2}\xi^\alpha\psi_\alpha = 0,$$

because $\phi^2\psi_\alpha = 0$ by the θ-component of $\Phi^3 = 0$. Next, we consider the variation of the θ-component of $\Phi^3 = 0$:

$$\begin{aligned}\delta\left(3\theta^\alpha\psi_\alpha\phi^2\right) &= 3\theta^\alpha\left[\delta\left(\psi_\alpha\right)\phi^2 + 2\psi_\alpha\phi\delta\phi\right]\\ &= 3\sqrt{2}\left[i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\left(\partial_\mu\phi\right)\phi^2 + \theta^\alpha\xi_\alpha F\phi^2 + 2\theta^\alpha\psi_\alpha\phi\xi^\beta\psi_\beta\right]\\ &= 3\sqrt{2}\left[i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\tfrac{1}{3}\partial_\mu\left(\phi^3\right) + \theta^\alpha\xi_\alpha F\phi^2 - \theta^\alpha\xi_\alpha\phi\psi^\beta\psi_\beta\right],\end{aligned}$$

where we used the fact that $\partial_\mu\left(\phi^3\right) = 3\phi^2\partial_\mu\phi$ and the identity $(\theta\psi)(\xi\psi) = -\frac{1}{2}(\theta\xi)(\psi\psi)$. Next, we use the lowest order component $\phi^3 = 0$ and the $\theta\theta$-component $3\phi^2 F = 3\phi\psi\psi$ to conclude that

$$\delta\left(3\theta^\alpha\psi_\alpha\phi^2\right) = 0.$$

Finally, we turn to the variation of the $\theta\theta$-component of $\Phi^3 = 0$. We see that

$$\begin{aligned}\delta\left(3\phi^2 F - 3\phi\psi^\alpha\psi_\alpha\right) &= 3\left[2\phi(\delta\phi)F + \phi^2\delta F - (\delta\phi)\psi\psi - 2\phi\psi^\alpha\delta\psi_\alpha\right]\\ &= 3\sqrt{2}\left[2\phi\xi^\alpha\psi_\alpha F - i\phi^2\left(\partial_\mu\psi^\alpha\right)\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - \xi^\alpha\psi_\alpha\psi^\beta\psi_\beta\right.\\ &\qquad\left. -2\phi\psi^\alpha\left(i\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\partial_\mu\phi + \xi_\alpha F\right)\right].\end{aligned}$$

The first and the last term cancel, because $\xi^\alpha\psi_\alpha = \psi^\alpha\xi_\alpha$. The third term vanishes due to the nilpotency of ψ. That leaves the second and fourth terms, which we write as

$$\delta\left(3\phi^2 F - 3\phi\psi\psi\right) = -3i\sqrt{2}\partial_\mu\left(\phi^2\psi^\alpha\right)\sigma^\mu_{\alpha\dot{\alpha}}\xi^{\dot{\alpha}} = 0,$$

which follows because $\partial_\mu\left(\phi^2\psi^\alpha\right) = \phi^2\partial_\mu\psi^\alpha + 2\phi\psi^\alpha\partial_\mu\phi$ and because the θ-component of $\Phi^3 = 0$ is $\phi^2\psi = 0$. We have seen that the variations of all three components vanish. This concludes the proof that $\Phi^3 = 0$ is supersymmetric invariant. □

A similar computation can be applied to higher order cases. However, we have a much simpler proof for the case $n > 4$. For this simpler proof, we first need the following.

Lemma 5.3. *Let $\Phi = (\phi, \psi, F)$ be a chiral superfield. Assume that F has nonzero body. Then, we have the following implications:*

$$\begin{aligned}\phi \propto \psi\psi \quad&\Longrightarrow\quad \Phi^3 = 0 \quad\Longrightarrow\quad \phi(\phi F - \psi\psi) = 0,\\ \phi \propto \psi \quad&\Longrightarrow\quad \Phi^4 = 0 \quad\Longrightarrow\quad \phi^2(6\phi F - 4\psi\psi) = 0,\\ \phi \propto \chi \quad&\Longrightarrow\quad \Phi^5 = 0 \quad\Longrightarrow\quad \phi^3(5\phi F - 10\psi\psi) = 0,\end{aligned}$$

for some spinor χ. Moreover, for $k > 5$, we also have $\phi \propto \chi \Longrightarrow \Phi^k = 0$.

Proof. This can be seen immediately by substituting the relevant expressions into (5.16). $\square$

Now, it is easy to prove the invariance of the constraints starting with $\Phi^5 = 0$.

Lemma 5.4. *The condition $\Phi^n = 0$ is supersymmetric invariant for $n > 4$.*

Proof. Let $n > 4$. Suppose that $\Phi^n = 0$. By Lemma 4.3 and Lemma 5.3, this is equivalent to $\phi \propto \chi$ for some spinor χ. Then, an arbitrary supersymmetry transformation will act on ϕ as

$$\phi \to \phi' = \phi + \sqrt{2}\xi^\alpha \psi_\alpha.$$

Since $\phi' \propto \chi'$, where χ' is some new spinor, we see that $(\Phi')^n = 0$. Hence, $\Phi^n = 0$ is supersymmetric invariant for all $n > 4$. $\square$

6

Non-linear realisations of supersymmetry

6.1 The Volkov-Akulov field

We have seen that when a chiral superfield $\Phi = (\phi, \psi, F)$ breaks supersymmetry, the auxiliary field F attains a nonzero vacuum expectation value: $\langle F \rangle \neq 0$. This implies that the spinor ψ transforms non-linearly on the vacuum. To see this, consider the transformation law for a spinor in a chiral superfield:

$$\delta\psi_\alpha = i\sqrt{2}\left(\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\right)\partial_\mu\phi + \sqrt{2}\xi_\alpha F. \tag{6.1}$$

Even though the term $\propto \langle \partial_\mu\phi \rangle$ drops out, the transformation law still has an inhomogeneous term $\propto \langle F \rangle$, analogous to the shift symmetry of pions. Compare (6.1) with the non-linear transformation laws for the U(1) toy model and Standard Model pion fields (2.4) and (2.15).

The first model of a non-linearly transforming goldstino was that of Volkov and Akulov [77]. The model includes a single Weyl spinor $\lambda_\alpha(x)$ with a transformation law like (6.1). The specific transformation law for $\lambda_\alpha(x)$ can be derived by first copying the infinitesimal coordinate transformations from (4.8):

$$\begin{aligned}
x^\mu &\to (x')^\mu = x^\mu + i\theta^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - i\xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\theta}^{\dot{\alpha}},\\
\theta^\alpha &\to (\theta')^\alpha = \theta^\alpha + \xi^\alpha,\\
\bar{\theta}^{\dot{\alpha}} &\to \left(\bar{\theta}'\right)^{\dot{\alpha}} = \bar{\theta}^{\dot{\alpha}} + \bar{\xi}^{\dot{\alpha}}.
\end{aligned}$$

We then introduce the correspondence $\lambda_\alpha(x) \leftrightarrow \sqrt{2}f\theta_\alpha$, which yields [15]

$$\lambda_\alpha(x) \to \lambda_\alpha\left(x'\right) + \sqrt{2}f\xi_\alpha = \lambda_\alpha(x) + \frac{i}{\sqrt{2}f}\left(\lambda^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - \xi^\alpha\sigma^\mu_{\alpha\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}}\right)\partial_\mu\lambda_\alpha + \sqrt{2}f\xi_\alpha,$$

which indeed satisfies the superPoincaré algebra relation $\{Q, \bar{Q}\}\lambda \propto P_\mu \lambda$. To construct a supersymmetric invariant action, we define the tensor [81]

$$A_\mu{}^\nu = \delta_\mu^\nu - \frac{i}{2f^2}\left[(\partial_\mu \lambda^\alpha)\,\sigma^\mu_{\alpha\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}} - \lambda^\alpha \sigma^\mu_{\alpha\dot{\alpha}}\partial_\mu \bar{\lambda}^{\dot{\alpha}}\right],$$

which transforms as

$$\delta A_\mu{}^\nu = \frac{i}{\sqrt{2}f}\left[\xi^\alpha \sigma^\rho_{\alpha\dot{\alpha}}\partial_\mu \bar{\lambda}^{\dot{\alpha}} - (\partial_\mu \lambda^\alpha)\,\sigma^\rho_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}}\right] A_\rho{}^\nu - \frac{i}{\sqrt{2}f}\left[\lambda^\alpha \sigma^\rho_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - \xi^\alpha \sigma^\rho_{\alpha\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}}\right]\partial_\rho A_\mu{}^\nu.$$

Next, we take the determinant of A, which itself transforms according to

$$\begin{aligned}\delta\,(\det A) &= \det A \;\mathrm{Tr}\;\delta\left(AA^{-1}\right)\\ &= -\frac{i}{\sqrt{2}f}\partial_\mu\left[\left(\lambda^\alpha \sigma^\mu_{\alpha\dot{\alpha}}\bar{\xi}^{\dot{\alpha}} - \xi^\alpha \sigma^\mu_{\alpha\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}}\right)\det A\right],\end{aligned}$$

which is a total derivative. Using this, we define the supersymmetric invariant action

$$\begin{aligned}\mathcal{L}_{\mathrm{VA}} &= -f^2 \det A\\ &= -f^2 + i\left[(\partial_\mu \lambda^\alpha)\,\sigma^\mu_{\alpha\dot{\alpha}}\bar{\lambda}^{\dot{\alpha}} - \lambda^\alpha \sigma^\mu_{\alpha\dot{\alpha}}\partial_\mu \bar{\lambda}^{\dot{\alpha}} + O\left(\lambda^2\,(\partial_\mu \lambda)^2\right)\right].\end{aligned}$$

To find the higher order terms, we make use of the following expression [22]

$$\det A = \exp(\mathrm{Tr}\;\log\;A) = \sum_{n=0}^{\infty}\left(\sum_{m=1}^{\infty}\frac{1}{m}\mathrm{Tr}\;B^m\right)^n,$$

where we introduced the tensor $B_\mu{}^\nu = A_\mu{}^\nu - \delta_\mu^\nu$. When written out in full, the Lagrangian then looks like

$$\begin{aligned}\mathcal{L}_{\mathrm{VA}} = -f^2\Bigg(1 + B_\mu{}^\nu + \tfrac{1}{2}\left(B_\mu{}^\mu B_\nu{}^\nu - B_\mu{}^\nu B_\nu{}^\mu\right) + \frac{1}{3!}\sum_p (-1)^p B_\mu{}^\mu B_\nu{}^\nu B_\rho{}^\rho\\ + \frac{1}{4!}\sum_p (-1)^p B_\mu{}^\mu B_\nu{}^\nu B_\rho{}^\rho B_\sigma{}^\sigma\Bigg),\end{aligned} \tag{6.2}$$

where the sums $\sum_p$ are over all possible permutations of B's upper indices.

6.2 Literature overview

We now give an overview of some of the techniques that have been used in the literature in order to obtain the non-linear Volkov-Akulov realisation, starting from a linear chiral multiplet. There are also other ways of going about this, which make use of component fields or non-linear superfields. For an overview, we refer to [22].

Roček [60] was the first to realise that it was possible to describe the Volkov-Akulov action in terms of a chiral multiplet $\Phi = (\phi, \psi, F)$ by imposing the constraint

$$\Phi^2 = 0.$$

We already met this *nilpotency constraint* in section 5.4. Under the assumption that F is not nilpotent ($F^n \neq 0$ for $n \in \mathbb{N}$), the constraint is solved by setting

$$\phi = \frac{\psi\psi}{2F}.$$

In other words, the effect of imposing the nilpotency constraint is that the scalar field ϕ is integrated out in favour of the goldstino ψ. In this sense, the constraint $\Phi^2 = 0$ is analogous to the pion model constraints $\phi\phi^* = v^2$ (2.5) and $\Sigma\Sigma^\dagger = \frac{1}{2}v^2$ (2.16), where the massive scalar fields $\sigma(x)$ are integrated out, leaving the Goldstone bosons $\pi(x)$. In addition to the nilpotency constraint, Roček also imposed the constraint $4f\Phi = -\frac{1}{2}\Phi\bar{D}^2\bar{\Phi}$ which eliminates the auxiliary field F. We will not do this here.

To see how the constraint $\Phi^2 = 0$ allows one to express the Volkov-Akulov Lagrangian in terms of a chiral multiplet, we write down the most general action for a single chiral multiplet Φ satisfying $\Phi^2 = 0$:

$$\mathcal{L} = \int \mathrm{d}^2\theta \mathrm{d}^2\overline{\theta}\ \Phi\bar{\Phi} + \int \mathrm{d}^2\theta\ f\Phi + \int \mathrm{d}^2\overline{\theta}\ f\bar{\Phi}, \tag{6.3}$$

where f is a real constant. Substituting the solution $\phi = \psi\psi/2F$ into the expression for a chiral superfield, we see that

$$\Phi = \frac{\psi\psi}{2F} + \sqrt{2}\theta\psi + \theta^2 F.$$

Substituting this back into (6.3), we find

$$\mathcal{L} = \tfrac{1}{2}i\left(\partial_\mu\psi\sigma^\mu\bar{\psi} - \psi\sigma^\mu\partial_\mu\bar{\psi}\right) + \bar{F}F + \partial_\mu\left(\frac{\bar{\psi}^2}{2\bar{F}}\right)\partial^\mu\left(\frac{\psi^2}{2F}\right) + fF + f\bar{F}.$$

The equations of motion for F and $\bar{F}$ are [47]

$$\bar{F} + f - \frac{\bar{\psi}^2}{2\bar{F}^2}\partial_\mu\partial^\mu\left(\frac{\psi^2}{2F}\right) = 0,$$
$$F + f - \frac{\psi^2}{2F^2}\partial_\mu\partial^\mu\left(\frac{\bar{\psi}^2}{2\bar{F}}\right) = 0,$$

with solutions

$$F = -f\left(1 + \frac{\bar{\psi}^2}{4f^4}\partial_\mu\partial^\mu\psi^2 - \frac{3}{16f^8}\psi^2\bar{\psi}^2\partial_\mu\partial^\mu\psi^2\partial_\mu\partial^\mu\bar{\psi}^2\right),$$
$$\bar{F} = -f\left(1 + \frac{\psi^2}{4f^4}\partial_\mu\partial^\mu\bar{\psi}^2 - \frac{3}{16f^8}\psi^2\bar{\psi}^2\partial_\mu\partial^\mu\psi^2\partial_\mu\partial^\mu\bar{\psi}^2\right).$$

This can be verified using (5.14) and substituting the expressions back into the equations of motion. This is easier than it looks, because the terms with coefficient $3/16f^8$ do not enter into the derivative terms of the equations of motion, due to the nilpotency of ψ. Finally, the solutions can be substituted back into the Langrangian, resulting in

$$\mathcal{L} = -f^2 + \tfrac{1}{2}i\left(\partial_\mu\psi\sigma^\mu\bar{\psi} - \psi\sigma^\mu\partial_\mu\bar{\psi}\right) + \frac{1}{4f^2}\bar{\psi}^2\partial_\mu\bar{\psi}^2\partial^\mu\psi^2 - \frac{1}{16f^6}\psi^2\bar{\psi}^2\partial_\mu\partial^\mu\psi^2\partial_\mu\partial^\mu\bar{\psi}^2, \quad (6.4)$$

which is equivalent to the Volkov-Akulov action (6.2), after a suitable field redefinition [49].

Next, we consider the work of Casalbuoni et al. [15]. They used a different approach to arrive at the condition $\Phi^2 = 0$. First, start with the general supersymmetric action for a single chiral multiplet:

$$S = \int \mathrm{d}^4x\,\mathcal{L} = \int \mathrm{d}^4x\,\mathrm{d}^2\theta\,\mathrm{d}^2\overline{\theta}\;K(\Phi,\overline{\Phi}) + \int \mathrm{d}^4x\,\mathrm{d}^2\theta\;W(\Phi) + \int \mathrm{d}^4x\,\mathrm{d}^2\overline{\theta}\;\overline{W}(\overline{\Phi}),$$

where $K(\Phi,\overline{\Phi})$ is the Kähler potential and $W(\Phi)$ the superpotential. In terms of component fields, this reads

$$\begin{aligned} S = \int \mathrm{d}^4x\,\Big[&K_1^1\left(\partial_\mu\phi\partial^\mu\overline{\phi} + \tfrac{1}{2}i\bar{\psi}\bar{\sigma}^\mu\partial_\mu\psi - \tfrac{1}{2}i\partial_\mu\bar{\psi}\bar{\sigma}^\mu\psi + F\bar{F}\right) - \tfrac{1}{2}K_{11}^1\left(\bar{F}\psi\psi - i\bar{\psi}\bar{\sigma}\cdot\partial(\phi\psi)\right) \\ &-\tfrac{1}{2}K_1^{11}\left(F\bar{\psi}\bar{\psi} + i\bar{\psi}\bar{\sigma}\cdot\partial(\bar{\phi}\psi)\right) + \tfrac{1}{4}K_{11}^{11}\psi\psi\bar{\psi}\bar{\psi} - \tfrac{1}{2}W_{11}\psi\psi - \tfrac{1}{2}W^{11}\bar{\psi}\bar{\psi} + W_1F + W^1\bar{F}\Big], \end{aligned} \quad (6.5)$$

where lower indices K_1 and upper indices K^1 refer to derivatives with respect to ϕ and $\bar{\phi}$, respectively. In normal coordinates, the derivatives $\langle K_{11}^1\rangle$ and $\langle K_1^{11}\rangle$ vanish when evaluated on the vacuum. Spontaneous supersymmetry breaking implies that $\langle F\rangle \neq 0$. In the limit of infinite curvature $\langle K_{11}^{11}\rangle \to \infty$, we then obtain a non-linear realisation associated with spontaneous supersymmetry breaking. This turns out to be equivalent to imposing the constraint $\Phi^2 = 0$. To see this, we expand the Kähler potential $K(\phi,\bar{\phi})$ around the vacuum and write

$$K_1^1 = \phi\bar{\phi}\langle K_{11}^{11}\rangle, \qquad K_1^{11} = \phi\langle K_{11}^{11}\rangle, \qquad K_{11}^1 = \bar{\phi}\langle K_{11}^{11}\rangle,$$

so that the divergent terms in the Lagrangian can be written as

$$\begin{aligned} S|_{\mathrm{div}} = \langle K_{11}^{11}\rangle \int \mathrm{d}^4x\,\Big[&\phi\bar{\phi}\left(\partial_\mu\phi\partial^\mu\overline{\phi} + \tfrac{1}{2}i\bar{\psi}\bar{\sigma}^\mu\partial_\mu\psi - \tfrac{1}{2}i\partial_\mu\bar{\psi}\bar{\sigma}^\mu\psi + F\bar{F}\right) \\ &-\tfrac{1}{2}\phi\left(\bar{F}\psi\psi - i\bar{\psi}\bar{\sigma}\cdot\partial(\phi\psi)\right) - \tfrac{1}{2}\bar{\phi}\left(F\bar{\psi}\bar{\psi} + i\bar{\psi}\bar{\sigma}\cdot\partial(\bar{\phi}\psi)\right) + \psi\psi\bar{\psi}\bar{\psi}\Big]. \end{aligned}$$

These divergent terms have to made stationary, which requires that the equality

$$-\tfrac{1}{2}\bar{\phi}\partial_\mu\partial^\mu\phi^2 + i\partial_\mu\left(\phi\bar{\psi}\bar{\sigma}^\mu\psi\right) - i\phi\left(\partial_\mu\bar{\psi}\right)\bar{\sigma}^\mu\psi + \bar{F}\left(\phi F - \tfrac{1}{2}\psi\psi\right) = 0,$$

is satisfied. The last term suggests that this equation is equivalent to $\phi = \psi\psi/2F$, which can be proved by iteration. Of course, we already saw that the condition $\phi = \psi\psi/2F$ was equivalent to $\Phi^2 = 0$. Casalbuoni et al. then go on to show that the action (6.5) together with the condition $\Phi^2 = 0$ results in a non-linear action comparable to the Volkov-Akulov model. This is essentially analogous to the procedure we used above to derive (6.4). In conclusion, the work of Casalbuoni et al. gives us another way of looking at the constraint $\Phi^2 = 0$, namely as a way of imposing the limit of infinite Kähler curvature. A similar approach is used by Kallosh, Karlsson, and Murli [46] for global supersymmetry, which they subsequently extend to local supersymmetry.

Next, we consider a recent paper by Seiberg and Komargodski [47], which spawned renewed interest in this topic. They argued that, in the infrared limit of low energies, the goldstino should reside in what they call the *superconformal symmetry breaking multiplet* X. This is a chiral multiplet that is associated with the Ferrara-Zumino multiplet through

$$\overline{D}^{\dot{\alpha}} \mathcal{J}_{\alpha\dot{\alpha}} = D_\alpha X.$$

See section 6.3 for details. They also argued that $X^2 = 0$ in the infrared limit, and investigated ways to couple this constrained superfield X to the MSSM. Later, Antoniadis and Ghilencea [2] confirmed that the goldstino lives in the superfield X in the limit of low energies and infinite cut-off scale Λ, but they also showed that the constraint $X^2 = 0$ is not necessarily true in models with multiple sources of supersymmetry breaking. For the case of two chiral multiplets breaking supersymmetry, they derived conditions on the Kähler and superpotential, which guarantee that the constraint $X^2 = 0$ is satisfied. Additionally, Ghilencea [32] offered specific counter-examples for which only higher order nilpotency constraints $X^k = 0$, with $k \geq 3$, are satisfied. In the next section, we will extend the analysis of Antoniadis and Ghilencea to the case of n chiral multiplets, but first let's make the following observation. We have seen in section 5.4 that the constraint $X^3 = 0$ can be solved by setting

$$\phi \propto \psi\psi.$$

In contrast to the constraint $X^2 = 0$, it is not necessary that $F^n \neq 0$ for all $n \in \mathbb{Z}$. In other words, $X^3 = 0$ could also be imposed if there were no supersymmetry breaking. Regardless, the approach in the next section, which is based on the Non-Linearisation Lemma of section 3.4, is to consider a theory which already features supersymmetry breaking, and to impose the constraint only in order to integrate out the massive scalar partner of the goldstino.

6.3 Conditions for nilpotency

We start by introducing the standard non-linear sigma model with n chiral superfields $\Phi^i = (\phi^i, \psi^i, F^i)$ [8] and then review the part of the analysis of [2, 32] pertinent to our discussion.

The general action for n chiral superfields is

$$S = \int \mathrm{d}^4x\, \mathcal{L} = \int \mathrm{d}^4x\, \mathrm{d}^2\theta\, \mathrm{d}^2\overline{\theta}\; K(\Phi^i, \overline{\Phi}_i) + \int \mathrm{d}^4x\, \mathrm{d}^2\theta\; W(\Phi^i) + \int \mathrm{d}^4x\, \mathrm{d}^2\overline{\theta}\; \overline{W}(\overline{\Phi}_i), \tag{6.6}$$

where $K(\Phi^i, \bar{\Phi}_i)$ is the Kähler potential and $W(\Phi^i)$ the superpotential. We let

$$K_i(\phi_k) = \frac{\partial K(\phi_k)}{\partial \phi^i}, \qquad K^i(\phi_k) = \frac{\partial K(\phi_k)}{\partial \overline{\phi}_i}, \qquad W_{ij}(\phi_k) = \frac{\partial^2 W(\phi_k)}{\partial \phi^i \partial \phi^j}, \qquad W^i = \overline{W_i}, \quad \text{etc.},$$

where ϕ^k refers to the n-vector of scalar fields, such that $K(\Phi^k)$ and $W(\Phi^k)$ are evaluated at $\Phi^k = \phi^k$. Dimensional considerations imply that third and higher derivatives of K are suppressed by inverse powers of the UV cut-off scale Λ. The full component expression is

$$\begin{aligned} \mathcal{L} = K_i^j \left[\partial_\mu \phi^i \partial^\mu \overline{\phi}_j + \tfrac{1}{2} i \psi^i \sigma^\mu D_\mu \overline{\psi}_j - \tfrac{1}{2} i D_\mu \psi^i \sigma^\mu \overline{\psi}_j + F^i \overline{F}_j\right] + \tfrac{1}{4} K_{ij}^{kl} \psi^i \psi^j \overline{\psi}_k \overline{\psi}_l \\ + \left(W_k - \tfrac{1}{2} K_k^{ij} \overline{\psi}_i \overline{\psi}_j\right) F^k + \left(W^k - \tfrac{1}{2} K_{ij}^k \psi^i \psi^j\right) \overline{F}_k - \tfrac{1}{2} W_{ij} \psi^i \psi^j - \tfrac{1}{2} W^{ij} \overline{\psi}_i \overline{\psi}_j. \end{aligned}$$

Here, the covariant derivative D_μ is given by

$$\begin{aligned} D_\mu \psi^i &= \partial_\mu \psi^i - \Gamma^i_{jk} (\partial_\mu \phi^i) \psi^k, \\ D_\mu \overline{\psi}_i &= \partial_\mu \overline{\psi}_i - \Gamma_i^{jk} (\partial_\mu \overline{\phi}_j) \overline{\psi}_k, \end{aligned}$$

where the connection is given in terms of (derivatives of) the Kähler metric K_i^j as

$$\begin{aligned} \Gamma^i_{jk} &= (K^{-1})^i_l K^l_{jk}, \\ \Gamma_i^{jk} &= (K^{-1})^l_i K_l^{jk}. \end{aligned}$$

The equations of motion for F^m and $\overline{F}_m$ are

$$\begin{aligned} \overline{F}_m &= -(K^{-1})^i_m W_i + \tfrac{1}{2} \Gamma^{lj}_m \overline{\psi}_l \overline{\psi}_j, \\ F^m &= -(K^{-1})^i_m W + i + \tfrac{1}{2} \Gamma^{lj}_m \overline{\psi}_l \overline{\psi}_j. \end{aligned}$$

These can be used to find the on-shell form of the Lagrangian:

$$\begin{aligned} \mathcal{L} = K_i^j \left(\partial_\mu \phi^i \partial^\mu \overline{\phi}_j + \tfrac{1}{2} i \psi^i \sigma^\mu D_\mu \overline{\psi}_j - \tfrac{1}{2} i D_\mu \psi^i \sigma^\mu \overline{\psi}_j\right) - (K^{-1})^i_k W^k W_i \\ - \tfrac{1}{2} \left(W_{ij} - \Gamma^m_{ij} W_m\right) \psi^i \psi^j - \tfrac{1}{2} \left(W^{ij} - \Gamma^{ij}_m W^m\right) \overline{\psi}_i \overline{\psi}_j + \tfrac{1}{4} R^{kl}_{ij} \psi^i \psi^j \overline{\psi}_k \overline{\psi}_l, \end{aligned}$$

where the Riemann curvature tensor is $R^{kl}_{ij} = K^{kl}_{ij} - K^n_{ij} \Gamma^{kl}_n$.

We also introduce the Ferrara-Zumino multiplet $\mathcal{J}_\mu$ [28, 47], which carries the supercurrent among its components. The FZ multiplet satisfies the constraint

$$\overline{D}^{\dot{\alpha}} \mathcal{J}_{\alpha\dot{\alpha}} = D_\alpha X,$$

where $J_{\alpha\dot{\alpha}} = -2\sigma^{\mu}_{\alpha\dot{\alpha}} J_\mu$ and where $X = (\phi_X, \psi_X, F_X)$ is a chiral superfield. It turns out that for the general action (6.6), X can be written as follows: [17, 47]

$$X = 4W - \tfrac{1}{3}\overline{D}^2 K - \tfrac{1}{2}\overline{D}^2 \overline{Y}(\overline{\Phi}^i),$$

where the last term is an "improvement term" containing an arbitrary holomorphic function $Y(\Phi^i)$. Ignoring the improvement term, the components of X can be written as [2]

$$\begin{aligned}
\phi_X &= 4W(\phi^i) + \frac{4}{3}\left(K^j \overline{F}_j - \tfrac{1}{2} K^{ij} \overline{\psi}_i \overline{\psi}_j\right), \\
\psi_X &= \psi^k \frac{\partial \phi_X}{\partial \phi^k} - \frac{4}{3} i\sigma^\mu \left(K^j \partial_\mu \overline{\psi}_j + K^{ij} \overline{\psi}_j \partial_\mu \overline{\phi}_i\right), \\
F_X &= F^i \frac{\partial \phi_X}{\partial \phi^i} - \tfrac{1}{2}\psi^i \psi^j \frac{\partial^2 \phi_X}{\partial \phi^i \partial \phi^j} + O(\partial_\mu),
\end{aligned} \tag{6.7}$$

where the last component is given up to terms vanishing in the limit of zero momentum (see the discussion surrounding (6.9)). These expressions are easily verified using $\overline{D}_{\dot{\alpha}} = \overline{\partial}_{\dot{\alpha}} + i\theta^\beta \sigma^\mu_{\beta\dot{\alpha}} \partial_\mu$ and $\overline{D}_{\dot{\alpha}} \Phi^i = 0$. We now expand X around the vacuum and evaluate all quantities $\langle W \rangle$, $\langle K^i \rangle$, $\langle F^i \rangle$, etc. at the vacuum. Field fluctuations are written as

$$\delta\phi^i = \phi^i - \langle \phi^i \rangle, \qquad \delta\psi^i = \psi^i - \langle \psi^i \rangle = \psi^i, \qquad \text{and} \qquad \delta F^i = F^i - \langle F^i \rangle.$$

From the equations of motion of ϕ^i and F^i, we know that $\langle \overline{F}_j \rangle = -\langle W_j \rangle$. Furthermore, in normal coordinates $\langle K^i_j \rangle = \delta^i_j$. Using this, we find to first order in $\delta\phi^i, \psi^i, \delta F^i$,

$$\begin{aligned}
\phi_X &= 4\langle W \rangle + \frac{8}{3}\langle W_j \rangle \delta\phi^j + O\left(\Lambda^{-1}\right), \\
\psi_X &= \frac{8}{3}\psi^k \langle W_k \rangle + O\left(\Lambda^{-1}\right), \\
F_X &= -\frac{8}{3}\langle W_i \rangle \langle W^i \rangle - 4\langle W^i \rangle \langle W_{ij} \rangle \delta\phi^j - \frac{4}{3}\langle W^i \rangle \delta\overline{F}_i + \frac{8}{3}\langle W_i \rangle \delta F^k + O\left(\Lambda^{-1}\right).
\end{aligned} \tag{6.8}$$

Following [47, 2], we ignore the constant term $4\langle W \rangle$ and consider $\tilde{X} \equiv X - 4\langle W \rangle$. If one chooses a field redefinition so that only one supermultiplet contains the supersymmetry breaking auxiliary field, then it can be shown [2] that the goldstino resides in this superfield $\tilde{X}$ in the limit of a large, but finite sgoldstino mass and infinite cut-off scale Λ. Because the mass of the sgoldstino is of the order of f/Λ, where f is the supersymmetry breaking scale, we restrict ourselves to the régime where the energy E satisfies

$$E \ll m_{\text{sgoldstino}} \cong \frac{f}{\Lambda}, \qquad f \ll \Lambda, \qquad \Lambda \to \infty. \tag{6.9}$$

At these energies, the momentum of the sgoldstino is negligible, justifying (6.7). We now seek conditions on the Kähler and superpotential that guarantee that $\tilde{X}^3 = 0$. To do this, we determine the equation of motion for the scalar field $\overline{\phi}_l$, for each chiral superfield Φ^l, in the limit of zero momentum. We find immediately

$$0 = \frac{\partial \mathcal{L}}{\partial \overline{\phi}_l} = W^{kl}(K^{-1})^i_k W_i + W^k (K^{-1})^{il}_k W_i + \tfrac{1}{2}\left(W^{ijl} - \partial^l(\Gamma^{ij}_m W^m)\right)\overline{\psi}_i\overline{\psi}_j - \tfrac{1}{2}\Gamma^m_{ij} W_m \psi^i\psi^j.$$

Once more, we expand this around the vacuum, using the simplifying assumptions $\langle W_{ij}\rangle = 0$, $\langle W^{ijl}\rangle\langle W_l\rangle = 0$, and $\langle W^{ijlm}\rangle = 0$. Then, for $l = 1, 2, \ldots, n$, we have

$$-\langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle\delta\phi^j + \tfrac{1}{2}\langle W^{ijl}\rangle\overline{\psi}_i\overline{\psi}_j - \tfrac{1}{2}\langle K^{lm}_{ij}\rangle\langle W_m\rangle\psi^i\psi^j + O\left(\Lambda^{-3}\right) = 0. \tag{6.10}$$

This system of linear equations may be solved in closed form for $\delta\phi^i$. To see whether $\tilde{X}^n = 0$, one could then substitute $\delta\phi^i$ into the component expressions (6.8). This is the method used by [2, 32] for $n = 2$. However, we will take a more general approach, which allows us to consider simultaneously the conditions for $\tilde{X}^2 = 0$ and $\tilde{X}^3 = 0$ for any n. To do this, we will need to use the fact that $\phi_{\tilde{X}} \propto \psi_{\tilde{X}}\psi_{\tilde{X}}$ implies that $\tilde{X}^3 = 0$, whereas $2F_{\tilde{X}}\phi_{\tilde{X}} = \psi_{\tilde{X}}\psi_{\tilde{X}}$ implies that $\tilde{X}^2 = 0$. This can be confirmed by substituting these expressions into (5.16). We combine both cases into the condition

$$\alpha\phi_{\tilde{X}} = \psi_{\tilde{X}}\psi_{\tilde{X}}, \tag{6.11}$$

where α is arbitrary (for $\tilde{X}^3 = 0$) or equals $2F_{\tilde{X}}$ (for $\tilde{X}^2 = 0$). In the latter case, we see from (6.8), ignoring second order terms, that we must set

$$\alpha = 2F_{\tilde{X}} = -\frac{16}{3}\langle W_i\rangle\langle W^i\rangle.$$

Using (6.8) again, we rewrite (6.11) as

$$\alpha(8/3)\langle W_j\rangle\delta\phi^j = \langle W_a\rangle\langle W_b\rangle\psi^a\psi^b + O\left(\Lambda^{-1}\right). \tag{6.12}$$

Equation (6.12) is one additional linear equation in $\delta\phi^i$. Combined with the equations of motion (6.10), this results in an $(n+1)\times n$ linear system, which we write as

$$\begin{pmatrix} -\langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle \\ \alpha(8/3)\langle W_j\rangle \end{pmatrix}\delta\phi^j + \begin{pmatrix} \frac{1}{2}\langle W^{abl}\rangle\overline{\psi}_a\overline{\psi}_b - \frac{1}{2}\langle K^{lm}_{ab}\rangle\langle W_m\rangle\psi^a\psi^b \\ -\langle W_a\rangle\langle W_b\rangle\psi^a\psi^b \end{pmatrix} = 0, \tag{6.13}$$

where the top n rows are indexed by l and the columns in the matrix on the left by j. Note that all coefficients are even elements of a Grassmann algebra $\mathcal{G}_L$. Here, we use the fact that products of odd elements, such as $\psi^a\psi^b$, are also even. Even elements commute, so it follows that we are dealing with a system of linear equations over a commutative ring. A first condition for system (6.13) to be consistent is that $\det\left(\langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle\right)$ is invertible, which is necessary and sufficient for the equations of motion in (6.10) to have a (unique) solution for any value of $\psi^a\psi^b$. This is an application of Proposition 7.13 of [11, III, §8]. Assuming that this is satisfied, we turn to the $(n+1)\times(n+1)$ augmented matrix

$$M = \left(\begin{array}{c|c} A & B \\ C & D \end{array}\right) \equiv \left(\begin{array}{c|c} -\langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle & \frac{1}{2}\langle W^{abl}\rangle\overline{\psi}_a\overline{\psi}_b - \frac{1}{2}\langle K^{lm}_{ab}\rangle\langle W_m\rangle\psi^a\psi^b \\ \alpha(8/3)\langle W_j\rangle & -\langle W_a\rangle\langle W_b\rangle\psi^a\psi^b \end{array}\right). \quad (6.14)$$

We note that $\delta\phi^j = A^{-1}B$ is the unique solution of the top n rows. Hence, the system is consistent if and only if $CA^{-1}B = D$. Of course, the inverse of A may not be readily available, so we look for an equivalent characterisation. Using Lemma 4.4, we can write the determinant of the augmented matrix as

$$\det(M) = \det(AD - BC) = \det(A)\det\left(D - CA^{-1}B\right) = \det(A)\left(D - CA^{-1}B\right).$$

Since $\det(A)$ is a unit and therefore not a zero divisor, it follows that $CA^{-1}B = D$ is equivalent to $\det(AD - BC) = 0$. We want this to be true for any choice of $\psi^a\psi^b = \psi^b\psi^a = \left(\overline{\psi}_a\overline{\psi}_b\right)^\dagger$. To emphasise this, we replace $\psi^a\psi^b$ with an arbitrary symmetric tensor V^{ab}. Substituting the necessary expressions from (6.14), we derive our main result:

Nilpotency test At energies far below the mass of the sgoldstino and in the limit $\Lambda \to \infty$, the goldstino superfield satisfies $\tilde{X}^3 = 0$ if for every symmetric tensor V^{ab}, there exists a nonzero scalar α, such that

$$\det\left(\frac{16}{3}\langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle\langle W_a\rangle\langle W_b\rangle V^{ab} - \alpha\left[\langle W^{abl}\rangle\overline{V}_{ab} - \langle K^{lm}_{ab}\rangle\langle W_m\rangle V^{ab}\right]\langle W_j\rangle\right) = 0. \quad (6.15)$$

Moreover, it satisfies $\tilde{X}^2 = 0$ if and only if this is true for $\alpha = -(16/3)\langle W_i\rangle\langle W^i\rangle$.

Remark 3. Our nilpotency test gives a necessary and sufficient condition for the second-order nilpotency constraint $\tilde{X}^2 = 0$, whereas the condition for $\tilde{X}^3 = 0$ is sufficient, but not necessary. The reason is that $2F_{\tilde{X}}\phi_{\tilde{X}} = \psi_{\tilde{X}}\psi_{\tilde{X}}$ is equivalent to $\tilde{X}^2 = 0$, whereas $\phi_{\tilde{X}} \propto \psi_{\tilde{X}}\psi_{\tilde{X}}$ is not necessary for $\tilde{X}^3 = 0$.

It might seem convenient to strip off V^{ab} and impose the conditions

$$\langle W^{abl}\rangle\langle W_j\rangle = 0,$$
$$2\langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle\langle W_a\rangle\langle W_b\rangle = -\alpha\langle K^{lm}_{ab}\rangle\langle W_m\rangle\langle W_j\rangle.$$

However, this results in trivial equations of motion unless $\alpha = -2\langle W_a\rangle\langle W^a\rangle$. This can be seen by contracting both sides of the second condition with $\langle W^a\rangle$, which yields $0 = \langle K^{il}_{kj}\rangle\langle W^k\rangle\langle W_i\rangle\langle W_j\rangle$. Substituting this back into the second condition then yields $\langle K^{lm}_{ab}\rangle\langle W_m\rangle = 0$, which makes (6.10) trivial. We thank Pelle Werkman for pointing this out.

Although condition (6.15) looks unwieldy, it is quite easy to apply in practice. We will show that geometrically, the condition can be interpreted in terms of sectional curvatures of planes spanned by the superpotential directions $\langle W_i \rangle$ and other vectors. We will also see that a number of these curvatures must vanish if the nilpotency constraint $\tilde{X}^2 = 0$ is satisfied. But first, we will illustrate the condition with a few examples.

Example 6.1. First, we consider a model with only 1 chiral superfield Φ_1 breaking supersymmetry. The Kähler and superpotentials are

$$K = \overline{\Phi}_1 \Phi_1 - \epsilon_1 \left(\overline{\Phi}_1 \Phi_1\right)^2 - \epsilon_2 \left(\Phi_1^3 \overline{\Phi}_1 + \Phi_1 \overline{\Phi}_1^3\right), \qquad W = f\Phi_1,$$

where ϵ_1, ϵ_2 are coefficients of order $1/\Lambda^2$ and where f is the supersymmetry breaking scale. This model was already considered by Komargodski and Seiberg [47], so we expect that $\tilde{X}^2 = 0$ in the limit of low energies. To confirm this with condition (6.15), we compute the following quantities:

$$\langle W_1 \rangle = \langle W^1 \rangle = f, \qquad \langle W^{111} \rangle = 0, \qquad \langle K_{11}^{11} \rangle = -4\epsilon_1.$$

Additionally, we compute $\alpha = -(16/3)\langle W_i \rangle \langle W^i \rangle = -(16/3)f^2$. Substituting these values in (6.15), we see that for any V^{11}, both terms exactly cancel:

$$-\frac{64}{3}\epsilon_1 f^4 V^{11} + \frac{64}{3}\epsilon_1 f^4 V^{11} = 0,$$

which means that $\tilde{X}^2 = 0$ in the infrared, as expected. □

Example 6.2. Next, we consider a model [1] with two chiral superfields Φ_1, Φ_2. The Kähler potential is given by

$$K = \overline{\Phi}_1 \Phi_1 + \overline{\Phi}_2 \Phi_2 - \epsilon_1 \left(\overline{\Phi}_1 \Phi_1\right)^2 - \epsilon_2 \left(\overline{\Phi}_2 \Phi_2\right)^2 - \epsilon_3 \left(\overline{\Phi}_1 \Phi_1\right)\left(\overline{\Phi}_2 \Phi_2\right) - \epsilon_4 \left[\overline{\Phi}_1^2 \Phi_2^2 + \text{h.c.}\right],$$

where the coefficients $\epsilon_1, \ldots, \epsilon_4$ are $O\left(1/\Lambda^2\right)$. The superpotential is given by

$$W = f\Phi_1,$$

where f is the supersymmetry breaking scale. Ghilencea [32] used this model as a counterexample to $\tilde{X}^2 = 0$, which was found to be false unless $\epsilon_4 = 0$. Let us apply condition (6.15) to confirm this. We compute the following quantities:

$$\langle W_1 \rangle = \langle W^1 \rangle = f, \quad \langle K_{11}^{11} \rangle = -4\epsilon_1, \quad \langle K_{12}^{12} \rangle = \langle K_{21}^{12} \rangle = -\epsilon_3, \quad \langle K_{22}^{11} \rangle = -4\epsilon_4,$$
$$\langle W_2 \rangle = \langle W^2 \rangle = \langle W^{111} \rangle = \cdots = \langle W^{222} \rangle = \langle K_{11}^{12} \rangle = \langle K_{22}^{21} \rangle = 0.$$

Substituting these values into (6.15), we see that $\tilde{X}^3 = 0$ in the infrared if for every symmetric tensor V^{ab}, there exists a nonzero scalar α, such that

$$\left(-\frac{64}{3}f^4\epsilon_1 V^{11} - 4f^2\alpha\left[\epsilon_1 V^{11} + \epsilon_4 V^{22}\right]\right)\left(-\frac{16}{3}f^4\epsilon_3\right) = 0.$$

If $\epsilon_1 V^{11} + \epsilon_4 V^{22}$ is a unit, then this is trivial. However, this is usually not the case, so the nilpotency test is inconclusive with regards to $\tilde{X}^3 = 0$. We can say with certainty that $\tilde{X}^2 \neq 0$ in the infrared, because α cannot, in general, be chosen to be equal to $-(16/3)\langle W_i\rangle\langle W^i\rangle = -(16/3)f^2$, except when $\epsilon_4 = 0$. $\square$

Noting that $\langle K^{il}_{kj}\rangle = R^{il}_{kj}$ is the curvature tensor in normal coordinates[1], it is possible to obtain a geometric interpretation of (6.15). First, we use the fact that $\det(AD - BC) = 0$ implies that $AD - BC$ has a nontrivial kernel [59, 11, III, §8.7.14]. That is, there exists a nonzero vector U^j, such that

$$(AD - BC)^l_j U^j = 0. \tag{6.16}$$

Substituting $\langle K^{il}_{kj}\rangle = R^{il}_{kj}$ and the relevant expressions from (6.14) into (6.16), we find

$$\forall V^{ab}, \exists U^j : \quad R^l{}_{kij}\langle W^k\rangle\langle W^i\rangle\langle W_a\rangle\langle W_b\rangle V^{ab}U^j = \tfrac{4}{3}\alpha\left(\langle W^{abl}\rangle\overline{V}_{ab} - R^l{}_{amb}\langle W\rangle^m V^{ab}\right)\langle W_j\rangle U^j.$$

We now transition into a more convenient geometric notation, where $\langle\cdot,\cdot\rangle$ denotes an inner product and $\langle R(A,B)C, D\rangle = R^a{}_{bcd}A^c B^d C^b D_a$. We also write the symmetric tensor V^{ab} in terms of its Takagi's decomposition[2]

$$V^{ab} = \sum_{i=1}^{n} \lambda_i V_i^a V_i^b.$$

From here on out, U, V_i, and W all denote vectors and i is no longer a tensor index. We then have the geometric condition

$$\forall V_i, \exists U : R(W,U)W\sum_{i=1}^{n}\lambda_i\langle W, V_i\rangle^2 = \tfrac{4}{3}\alpha\langle W, U\rangle\sum_{i=1}^{n}\left(W^{abl}\bar{\lambda}_i\left(V_i^a V_i^b\right)^\dagger - \lambda_i R(W,V_i)V_i\right). \tag{6.17}$$

Disregarding the W^{abl}-term, this expression has an interesting interpretation: the left-hand side is a product of the curvature in the (W,U)-plane and the lengths of the V_i along W, whereas the right-hand side is a product of the length of U along W and a sum of the curvatures in the (W,V_i)-planes. This condition is not easy to satisfy unless the curvatures

[1] In moving from K^{lm}_{ab} to the Riemann tensor $R^l{}_b{}^m{}_a$, the correct index placement follows from (see (6.10) and above): $K^{lm}_{ab} = \partial^l\Gamma^m{}_{ab} = R^m{}_a{}^l{}_b = R^l{}_b{}^m{}_a$.

[2] Every square complex symmetric matrix A has a decomposition $A = VDV^{\mathbf{T}}$, where V is unitary and D is diagonal with real nonnegative entries, which are the square roots of the eigenvalues of $AA^\dagger$ [42].

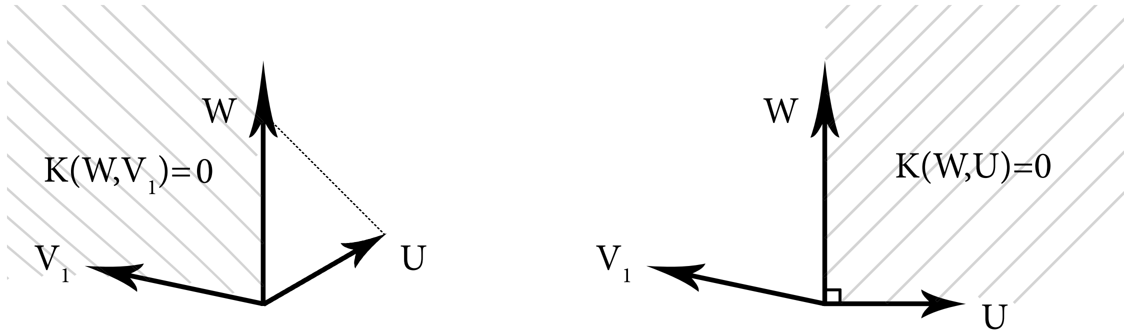


Figure 6.1: The vector W is given, whereas V_1 is arbitrary and U must be chosen so that the geometric condition (6.17) is satisfied. If $\langle W, U\rangle \neq 0$ (left), then the curvature of any plane spanned by W and V_1 must vanish. On the other hand, if $\langle W, U\rangle = 0$ (right), then the curvature in the (W, U)-plane must vanish.

vanish. In fact, it turns out that this is required. To see this, we take an inner product of both sides with W by contracting with W_l:

$$\langle R(W,U)W, W\rangle \sum_{i=1}^{n} \lambda_i \langle W, V_i\rangle^2 = -\tfrac{4}{3}\alpha\langle W, U\rangle \sum_{i=1}^{n} \lambda_i \langle R(W, V_i)V_i, W\rangle,$$

where the W^{abl}-term drops out by virtue of the earlier assumption that $W^{abl}W_l = 0$. The left-hand side also vanishes, due to the antisymmetry $R^l{}_{kij} = -R^k{}_{lij}$ in the Riemann tensor. Hence, we can write the right-hand side as

$$0 = -\tfrac{4}{3}\alpha\langle W, U\rangle \sum_{i=1}^{n} \lambda_i K(W, V_i) \left[\langle W, W\rangle\langle V_i, V_i\rangle - \langle W, V_i\rangle^2\right],$$

where $K(W, V_i)$ is the sectional curvature of the (W, V_i)-plane. Of course for $n > 1$, it is easy to pick U to be orthogonal to W, so that the above equation is satisfied. However, then (6.17) is inconsistent unless

$$R(W,U)W = 0.$$

After all, we assumed that $D = -\sum_{i=1}^{n} \lambda_i \langle W, V_i\rangle^2 \neq 0$. On the other hand, if $\langle W, U\rangle \neq 0$, then we must have $K(W, V) = 0$ for each vector V, because we are free to set $\lambda_2 = \cdots = \lambda_n = 0$ and choose V_1 arbitrarily[3]. In either case, a number of sectional curvatures at the vacuum must vanish. See figure 6.1. Of course, the vanishing of these curvatures is only a necessary condition, but condition (6.17) is both necessary and sufficient in order to guarantee that $\tilde{X}^3 = 0$ (or $\tilde{X}^2 = 0$ if $\alpha = \frac{16}{3}\langle W, W\rangle$), for all $\psi^a\psi^b$.

[3]Note that $\langle W, W\rangle\langle V_1, V_1\rangle - \langle W, V_1\rangle^2 = 0$ if and only if W and V_1 are parallel. However, in this case, also $K(W, V_1) = 0$.

One might wonder what we can say about other constraints on $\tilde{X}$. First, let's consider the condition $\tilde{X}^4 = 0$. By Lemma 5.3, the condition is true whenever $\phi_{\tilde{X}} \propto \psi_{\tilde{X}}\chi$ for an arbitrary spinor χ. It is easy to fit this into our approach, by relaxing the symmetry condition on V^{ab}. Unfortunately, we can no longer use Takagi's decomposition, but we can simply use the singular value decomposition[4] of $V^{ab} = \sum_{i=1}^{n} \sigma_i P_i^a \left(R_i^b\right)^*$. This then yields the geometric condition that for each V^{ab}, there must exist a vector U, such that

$$R(W,U)W \sum_{i=1}^{n} \sigma_i \langle W, P_i\rangle\langle W, R_i^*\rangle = \tfrac{4}{3}\alpha\langle W, U\rangle \sum_{i=1}^{n} \left(\bar{\sigma}_i W^{abl} \left(R_i^a\right)^{\mathbf{T}} \left(P_i^b\right)^{\dagger} - \sigma_i R(W, P_i) R_i^*\right).$$

Any higher order nilpotency conditions, starting with $\tilde{X}^5 = 0$, are trivially satisfied if ϕ_X is nilpotent. In a way, we have now said all there is to say, because the most general condition on a chiral superfield is exactly

$$Z(\tilde{X}) = Z(\phi_{\tilde{X}}) + \sqrt{2}\theta\psi_{\tilde{X}} \left.\frac{\partial Z}{\partial \phi_{\tilde{X}}}\right|_{\tilde{X}=\phi_{\tilde{X}}} + \theta\theta \left[\left.\frac{\partial Z}{\partial \phi_{\tilde{X}}}\right|_{\tilde{X}=\phi_{\tilde{X}}} F_{\tilde{X}} - \tfrac{4}{3} \left.\frac{\partial^2 Z}{\partial \phi_{\tilde{X}}^2}\right|_{\tilde{X}=\phi_{\tilde{X}}} \psi_{\tilde{X}}\psi_{\tilde{X}} \right] = 0.$$

Expanding this as a power series in $\phi_{\tilde{X}}$, we find

$$\begin{aligned} Z(\tilde{X}) = {}& [\beta_0 + \beta_1\phi_{\tilde{X}} + \ldots] + \sqrt{2}\theta\psi_{\tilde{X}} [\beta_1 + \phi_{\tilde{X}}\beta_2 + \ldots] \\ &+ \theta\theta \left[(\beta_1 + \phi_{\tilde{X}}\beta_2 + \ldots) F_{\tilde{X}} - \tfrac{4}{3} (\beta_2 + \phi_{\tilde{X}}\beta_3 + \ldots) \psi_{\tilde{X}}\psi_{\tilde{X}} \right]. \end{aligned}$$

The $\theta\theta$-component gives us the only interesting condition:

$$(\beta_1 + \beta_2\phi_{\tilde{X}}) F_{\tilde{X}} - \tfrac{1}{2} (\beta_2 + \beta_3\phi_{\tilde{X}}) \psi_{\tilde{X}}\psi_{\tilde{X}} = 0,$$

which can be treated in the same way as the condition $2\phi_{\tilde{X}}F_{\tilde{X}} = \psi_{\tilde{X}}\psi_{\tilde{X}}$ which gave us the $\tilde{X}^2 = 0$ condition.

[4] Here, A^* denotes complex conjugation.

7

Conclusion

We covered a number of related topics in this thesis. First of all, we discussed the role of non-linear realisations in spontaneous symmetry breaking. In chapter 3, we obtained a generalisation of the Linearisation Lemma of Bochner [9] and Coleman, Wess, and Zumino [21], which tells us that a non-linear realisation becomes linear when restricted to the isotropy group of the vacuum. Although we were able to relax a few of the assumptions, the lemma still does not apply to the case of supersymmetry, because the Poincaré group is not compact. Regardless, similar results do hold for the superPoincaré group [81, p. 246], which suggests that a proof for the general non-compact case might be possible. Such a proof would probably not use group integration. An alternative resolution might be through the use of Weyl's unitarian trick. This is discussed at the end of section 3.3.

In section 3.4, we also proved the Non-Linearisation Lemma, which states that imposing a constraint to eliminate a field yields a consistent non-linear realisation if and only if the constraint is invariant under the transformation group. This result is essential to the approach of chapter 6, where we used nilpotency constraints of the form $\tilde{X}^n = 0$ to eliminate the scalar partner of the goldstino. In order to be able to successfully apply the Non-Linearisation Lemma in this case, we proved in section 5.4 that the nilpotency constraints are supersymmetric invariant.

Another way of looking at the use of the nilpotency constraints is as a way of writing the Volkov-Akulov action in terms of a chiral superfield. In chapter 6, we reviewed this approach and we expanded on the analysis of Antoniadis and Ghilencea [2] by deriving conditions for the validity of $\tilde{X}^2 = 0$ and $\tilde{X}^3 = 0$ in the case of n chiral superfields. We also considered more general constraints on $\tilde{X}$. A few observations are in order. First of all, the models considered in section 6.3 only contain chiral superfields. Of course, if we wish to apply this approach to more realistic models, such as in [47], then we should be able to incorporate other types of superfields. As a start, for supersymmetric gauge theories it would be necessary to include vector superfields. In that case, it might no longer be sufficient to consider only F-term breaking, as we assumed in section 5.3. Secondly, and on a related note, it would be interesting to see what conditions like (6.15) imply for the matter couplings derived by Komargodski and Seiberg [47]. Finally, it might be prudent to study possible extensions to local supersymmetry. This is useful if one wishes to use the nilpotency constraints in cosmological applications, as is done, for example, in [27].

Acknowledgements

I would like to thank Pelle Werkman and Diederik Roest for many fruitful discussions. I would also like to extend my gratitude to Arthemy Kiselev for his detailed comments on the mathematical aspects.

Bibliography

[1] I. Antoniadis, E. Dudas, and D. Ghilencea. Goldstino and sgoldstino in microscopic models and the constrained superfields formalism. *Nuclear Physics B*, 857(1):65–84, 2012.

[2] I. Antoniadis and D. Ghilencea. Low-scale SUSY breaking and the (s)goldstino physics. *Nuclear Physics B*, 870(1):278–291, 2013.

[3] ATLAS collaboration and others. Search for resonances decaying to photon pairs in 3.2 fb- 1 of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. ATLAS-CONF-2015-081, 2015.

[4] ATLAS Collaboration and others. Search for supersymmetry at $\sqrt{s} = 13$ TeV in final states with jets and two same-sign leptons or three leptons with the ATLAS detector. *arXiv:1602.09058*, 2016.

[5] P. Bechtle, T. Plehn, and C. Sander. The Status of Supersymmetry after the LHC Run 1. *arXiv:1506.03091*, 2015.

[6] M. Berger, P. Pansu, J.-P. Berry, and X. Saint-Raymond. *Problems in geometry*. Springer Science & Business Media, 2013.

[7] E. A. Bergshoeff, D. Z. Freedman, R. Kallosh, and A. Van Proeyen. Pure de Sitter supergravity. *Physical Review D*, 92(8):085040, 2015.

[8] M. Bertolini. Lectures on Supersymmetry. *Lecture notes given at SISSA*, 2015.

[9] S. Bochner. Compact groups of differentiable transformation. *Annals of Mathematics*, pages 372–381, 1945.

[10] R. Boels and F. Brümmer. Introduction to Supersymmetry and Supergravity. 2012.

[11] N. Bourbaki. *Algebra I: chapters 1-3*. Springer Science & Business Media, 1998.

[12] R. L. Bryant. An introduction to Lie groups and symplectic geometry. Geometry and quantum field theory. *IAS/Park City Math. Ser*, 1, 1991.

[13] C. G. Callan Jr, S. Coleman, J. Wess, and B. Zumino. Structure of phenomenological Lagrangians. II. *Physical Review*, 177(5):2247, 1969.

[14] J. E. Campbell. On a law of combination of operators (second paper). *Proceedings of the London Mathematical Society*, 1(1):14–32, 1897.

[15] R. Casalbuoni, S. De Curtis, D. Dominici, F. Feruglio, and R. Gatto. Non-Linear realization of supersymmetry algebra from supersymmetric constraint. *Physics Letters B*, 220(4):569–575,

1989.

[16] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Bleick. *Analysis, Manifolds, and Physics*. Elsevier, 1982. Page 73.

[17] T. Clark and S. Love. The supercurrent in supersymmetric field theories. *International Journal of Modern Physics A*, 11(16):2807–2821, 1996.

[18] CMS collaboration and others. Search for new physics in high mass diphoton events in proton-proton collisions at 13 TeV. *CMS PAS EXO-15-004 (Dec. 2015)*, 2015.

[19] D. L. Cohn. *Measure theory*, volume 165. Springer, 1980.

[20] S. Coleman and J. Mandula. All possible symmetries of the S matrix. *Physical Review*, 159(5):1251, 1967.

[21] S. Coleman, J. Wess, and B. Zumino. Structure of phenomenological lagrangians. I. *Physical Review*, 177(5):2239, 1969.

[22] A. Coser. Non-linear Realizations of Supersymmetry. Master's thesis, Universita degli studi di Padova, 2009.

[23] M. Dafermos. Part III Differential Geometry Lecture Notes. Retrieved from `https://www.dpmms.cam.ac.uk/~md384/neessnmeiwseis.pdf` on 20-12-2015.

[24] B. de Wit and J. Smith. *Field theory in particle physics*, volume 1. Elsevier, 2012.

[25] E. Fabri and L. Picasso. Quantum field theory and approximate symmetries. *Physical Review Letters*, 16(10):408, 1966.

[26] S. Ferrara, R. Kallosh, and A. Linde. Cosmology with nilpotent superfields. *Journal of High Energy Physics*, 2014(10):1, 2014.

[27] S. Ferrara, R. Kallosh, and J. Thaler. Cosmology with orthogonal nilpotent superfields. *arXiv:1512.00545*, 2015.

[28] S. Ferrara and B. Zumino. Transformation properties of the supercurrent. *Nuclear Physics B*, 87(2):207–220, 1975.

[29] D. H. Franco and C. M. Polito. Supersymmetric field-theoretic models on a supermanifold. *Journal of Mathematical Physics*, 45(4):1447–1473, 2004.

[30] G. Aad et al. (Atlas collaboration). Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using $\sqrt{s} = 8$ TeV proton-proton collision data. *Journal of high energy physics*, 2014(9):1–52, 2014.

[31] J. Gallier. *Basics of affine geometry*. Springer, 2011.

[32] D. Ghilencea. Comments on the nilpotent constraint of the goldstino superfield. *arXiv:1512.07484*, 2015.

[33] J. Gleason. Existence and Uniqueness of Haar Measure. 2010. Retrieved from `http://www.math.uchicago.edu/~may/VIGRE/VIGRE2010/REUPapers/Gleason.pdf`.

[34] H. Glöckner. Implicit functions from topological vector spaces to Banach spaces. *Israel Journal*

of Mathematics, 155(1):205–252, 2006.

[35] J. Goldstone. Field theories with «Superconductor» solutions. *Il Nuovo Cimento (1955-1965)*, 19(1):154–164, 1961.

[36] G. Guralnik, C. Hagen, and T. Kibble. Broken symmetries and the Goldstone theorem. 1968.

[37] R. Haag, J. T. Łopuszański, and M. Sohnius. All possible generators of supersymmetries of the S-matrix. *Nuclear Physics B*, 88(2):257–274, 1975.

[38] A. Haar. Der Massbegriff in der Theorie der kontinuierlichen Gruppen. *Annals of mathematics*, pages 147–169, 1933.

[39] M. Hazewinkel. *Handbook of algebra*, volume 3. Elsevier, 2000.

[40] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80. Academic press, 1979.

[41] Hermes. Is this true: Every open set A contains a neighborhood whose closure is a subset of A. Mathematics Stack Exchange. Retrieved from `http://math.stackexchange.com/q/1404582`.

[42] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.

[43] A. Huckleberry and T. Wurzbacher. *Infinite dimensional Kähler manifolds*, volume 31. Birkhäuser, 2012.

[44] B. Iyer, N. Mukunda, and C. Vishveshwara. *Gravitation, gauge theories and the early universe*, volume 29. Springer Science & Business Media, 2012.

[45] X. Ji. Chiral Dynamics. *Lecture notes*, 2015. Retrieved from `http://www.physics.umd.edu/courses/Phys741/xji/chapter5.pdf`.

[46] R. Kallosh, A. Karlsson, and D. Murli. From linear to nonlinear supersymmetry via functional integration. *Physical Review D*, 93(2):025012, 2016.

[47] Z. Komargodski and N. Seiberg. From linear SUSY to constrained superfields. *Journal of High Energy Physics*, 2009(09):066, 2009.

[48] R. Kupferman. Basic Notions in Functional Analysis - Topological Vector Spaces. *Lecture notes*, 2014. Retrieved from `http://www.ma.huji.ac.il/~razk/iWeb/My_Site/Teaching_files/TVS.pdf`.

[49] S. M. Kuzenko and S. J. Tyler. Relating the Komargodski–Seiberg and Akulov–Volkov actions: Exact nonlinear field redefinition. *Physics Letters B*, 698(4):319–322, 2011.

[50] S. Lang. *Differential manifolds*, volume 212. Springer, 1972.

[51] I. Low and A. V. Manohar. Spontaneously broken spacetime symmetries and Goldstone's theorem. *Physical review letters*, 88(10):101602, 2002.

[52] J. Mickelsson and J. Niederle. On non-linear realizations of the group SU(2). *Communications in Mathematical Physics*, 16(3):191–206, 1970.

[53] S. Nagamachi and Y. Kobayashi. Hilbert superspace. *Journal of mathematical physics*, 33(12):4274–4282, 1992.

[54] J. Nelson. The Six-fold Nature of Snow. 2011. Retrieved from `http://www.storyofsnow.com/blog1.php/2011/03/15/how-the-crystal-got-its-six`.

[55] E. Noether. Invariante variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse*, 1918:235–257, 1918.

[56] C. Petersson and R. Torre. The 750 GeV diphoton excess from the goldstino superpartner. *arXiv:1512.05333*, 2015.

[57] L. S. Pontryagin. *Topological groups*. Gordon and Breach, 1966.

[58] I. Porteous. *Topological Geometry*. Cambridge University Press, 1981. Chapters 18-19.

[59] R. B. Richter and W. P. Wardlaw. Diagonalization over commutative rings. *The American Mathematical Monthly*, 97(3):223–227, 1990.

[60] M. Roček. Linearizing the Volkov-Akulov model. *Physical Review Letters*, 41(7):451, 1978.

[61] A. Rogers. A global theory of supermanifolds. *Journal of Mathematical Physics*, 21(6):1352–1365, 1980.

[62] W. Rudin. *Functional analysis*. McGraw-Hill, 1991.

[63] O. Rudolph. Super hilbert spaces. *Communications in Mathematical Physics*, 214(2):449–467, 2000.

[64] W. Rühl and B. Yunn. Superfields as representations. *Fortschritte der Physik*, 23(7):431–449, 1975.

[65] L. H. Ryder. *Quantum field theory*. Cambridge university press, 1996.

[66] S. Chatrchyan et al. (CMS collaboration and others). Search for new physics in the multi-jet and missing transverse momentum final state in proton-proton collisions at $\sqrt{s}$= 8 TeV. *arXiv:1402.4770*, 2014.

[67] E. B. Saff and A. D. Snider. *Fundamentals of complex analysis for mathematics, science, and engineering*. Prentice-Hall, 1976.

[68] A. Santra, A. Askew, M. Hildreth, A. Kyriakis, D. Mason, A. Reinsvold, I. Topsis-Giotis, and M. Weinberg. A Search for Supersymmetry at CMS with Two Photons and Missing Transverse Energy at $\sqrt{s} = 13$ TeV. *Bulletin of the American Physical Society*, 2016.

[69] A. Schellekens. Beyond the standard model. *Lecture Notes, june*, 2007.

[70] M. D. Schwartz. *Quantum field theory and the standard model*. Cambridge University Press, 2013.

[71] J. R. Silvester. Determinants of block matrices. *The Mathematical Gazette*, 84(501):460–467, 2000.

[72] E. Snapper and R. J. Troyer. *Metric affine geometry*. Elsevier, 2014.

[73] D. Tong. Quantum field theory. *Lecture Notes*, 2007.

[74] G. M. Tuynman. *Supermanifolds and supergroups: basic theory*, volume 570. Springer Science & Business Media, 2004.

[75] V. S. Varadarajan. *Supersymmetry for mathematicians: an introduction.* American Mathematical Soc., 2004.

[76] R. Vinroot. Topological groups. *Lecture notes*, 2014. Retrieved from `http://www.math.wm.edu/~vinroot/PadicGroups/topgroups.pdf`.

[77] D. Volkov and V. Akulov. Is the neutrino a goldstone particle? *Physics Letters B*, 46:109–110, 1973.

[78] H. Watanabe and T. Brauner. Spontaneous breaking of continuous translational invariance. *Physical Review D*, 85(8):085010, 2012.

[79] A. Weil, A. Weil, A. Weil, F. Mathematician, and A. Weil. *L'intégration dans les groupes topologiques et ses applications*, volume 1145. Hermann Paris, 1965.

[80] E. Weinberg. Quantum Field Theory III. *Lecture notes*, 2011. Retrieved from `http://phys.columbia.edu/~cyr/notes/QFT_3/lecture18.pdf`.

[81] J. Wess and J. Bagger. *Supersymmetry and supergravity.* Princeton university press, 1992.

[82] N. Wolchover. As supersymmetry fails tests, physicists seek new ideas. *Quanta Magazine*, 2012.

[83] H. Yukawa. On the interaction of elementary particles. I. *Nippon Sugaku-Buturigakkwai Kizi Dai 3 Ki*, 17(0):48–57, 1935.