CANONICAL GROUP QUANTIZATION AND BOUNDARY CONDITIONS

Dissertation

zur Erlangung des Grades »Doktor der Naturwissenschaften«
am Fachbereich *Physik, Mathematik und Informatik*
der Johannes Gutenberg-Universität in Mainz

Florian Jung
geboren in Saarbrücken

Mainz, April 2012

Florian Jung
Theoretische Hochenergie-Physik (THEP)
Institut für Physik
Staudingerweg 7
55128 Mainz (Germany)

Typeset in Adobe® Minion® Pro using pdfLaTeX. Figures drawn in metapost and Inkscape, diagrams created using TikZ/pgf.

version 1.0 ¶ 2 4 6 8 10 9 7 5 3

D77 Dissertation Universität Mainz

1. Gutachter —
2. Gutachter —
3. Gutachter —

Datum der mündlichen Prüfung 16. Juli 2012

There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable.

There is another which states that this has already happened.

Douglas Adams, The Restaurant at the End of the Universe

Abstract

In the present thesis, we study quantization of classical systems with non-trivial phase spaces using the group-theoretical quantization technique proposed by Isham. Our main goal is a better understanding of global and topological aspects of quantum theory. In practice, the group-theoretical approach enables direct quantization of systems subject to constraints and boundary conditions in a natural and physically transparent manner – cases for which the canonical quantization method of Dirac fails.

First, we provide a clarification of the quantization formalism. In contrast to prior treatments, we introduce a sharp distinction between the two group structures that are involved and explain their physical meaning. The benefit is a consistent and conceptually much clearer construction of the Canonical Group. In particular, we shed light upon the 'pathological' case for which the Canonical Group must be defined via a central Lie algebra extension and emphasise the role of the central extension in general.

In addition, we study direct quantization of a particle restricted to a half-line with 'hard wall' boundary condition. Despite the apparent simplicity of this example, we show that a naive quantization attempt based on the cotangent bundle over the half-line as classical phase space leads to an incomplete quantum theory; the reflection which is a characteristic aspect of the 'hard wall' is not reproduced. Instead, we propose a different phase space that realises the necessary boundary condition as a topological feature and demonstrate that quantization yields a suitable quantum theory for the half-line model. The insights gained in the present special case improve our understanding of the relation between classical and quantum theory and illustrate how contact interactions may be incorporated.

Zusammenfassung

In der vorliegenden Dissertation beschäftigen wir uns mit der Quantisierung von klassischen Systemen mit nicht-trivialen Phasenräumen mittels der gruppentheoretischen Quantisierungsmethode, welche von Isham vorgeschlagen wurde. Unser Hauptziel ist ein besseres Verständnis globaler und topologischer Aspekte der Quantentheorie. In der Praxis erlaubt der gruppentheoretische Zugang die direkte Quantisierung von Systemen mit Zwangs- und Randbedingungen in natürlicher und physikalisch transparenter Weise – Fälle, in denen die kanonische Quantisierungsmethode von Dirac versagt.

Als Erstes liefern wir eine Präzisierung des Quantisierungsformalismus. Im Gegensatz zu vorherigen Arbeiten führen wir eine strikte Unterscheidung zwischen den beiden beteiligten Gruppenstrukturen ein und erläutern deren physikalische Bedeutung. Das Ergebnis ist eine konsistente und konzeptionell deutlich klarere Konstruktion der Kanonischen Gruppe. Insbesondere diskutieren wir den »pathologischen« Fall, in dem die Kanonische Gruppe mittels zentraler Liealgebra-Erweiterung definiert werden muss, und betonen die Rolle der zentralen Erweiterung im Allgemeinen.

Außerdem betrachten wir die direkte Quantisierung eines Teilchens, welches durch eine unendlich hohe Potentialwand auf eine Halbgerade eingeschränkt wird. Trotz der scheinbaren Einfachheit dieses Beispiels zeigen wir, dass ein naiver Quantisierungsversuch basierend auf dem Kotangentialbündel über der Halbgeraden als Phasenraum zu einer unvollständigen Quantentheorie führt; die Reflexion, ein wesentlicher Aspekt der harten Potentialwand, wird nicht reproduziert. Stattdessen schlagen wir einen alternativen Phasenraum vor, der die notwendige Randbedingung als topologische Eigenschaft realisiert und demonstrieren, dass durch Quantisierung eine geeignete Quantentheorie für die Halbgeraden entsteht. Die aus diesem Spezialfall gewonnenen Erkenntnisse verbessern das Verständnis der Beziehung zwischen Klassischer und Quantenmechanik und zeigen, wie Kontaktwechselwirkungen berücksichtigt werden können.

Contents

Introduction and Overview

First quantization is a mystery,
but second quantization is a functor.

Edward Nelson, in [RS.II, section X.7]

Today, almost nine decades after the birth of the 'new' quantum theory, pioneered by Heisenberg, Born and Jordan [Hei25; BJ25; BHJ26], and developed independently by Schrödinger [Sch26], non-relativistic quantum mechanics stands as a well-established and experimentally well-tested theory. Nevertheless, quantum theory still hides some intricate secrets not just of philosophical value. Specifically, there is the question how to pass from classical to quantum theory via a procedure known as *quantization*:

> How can we construct a quantum theory if a classical system is given?

In fact, the very concept of quantization seems pointless if we consider quantum theory to be the more fundamental theory and classical mechanics to be only approximately correct. A direct construction of a quantum theory based on first principles, however, is in most cases – or at least in the more interesting ones – infeasible, and history has shown that the similarity between classical and quantum physics is large enough to make an investigation of quantization methods a worthwhile endeavour. Quantization in its modern sense is therefore often understood as construction of a quantum theory with *help* of a classical reference, not necessarily as a strict mapping. The *relationship* between classical and quantum theory shifts into the foreground. In the present thesis, our aim is to study this relation in case of systems with non-trivial classical phase spaces using the group-theoretical quantization scheme proposed by Christopher Isham in [Ish83].

Historically, the question of quantization in its most basic form was already present in Heisenberg's 1925 paper [Hei25, p. 881]:

> 'Gegeben sei eine an Stelle der klassischen Größe $x(t)$ tretende quantentheoretische Größe; welche quantentheoretische Größe tritt dann an Stelle von $x(t)^2$?'

('If instead of a classical quantity $x(t)$ we have a quantum-theoretical quantity, what quantum-theoretical quantity will appear in place of $x(t)^2$?' [Wae67, section 12]). One of Heisenberg's main achievements in this groundbreaking work was the insight that quantum observables, in contrast to classical observables, have to be represented by *non-commutative* quantities. Born and Jordan immediately realised that Heisenberg's 'quantum-theoretical quantities' can be understood as (infinite-dimensional) matrices [BJ25]. Using the more powerful and economical language of matrices, they were able to derive the infinitesimal 'quantum condition' in its modern form:

$$\boldsymbol{pq} - \boldsymbol{qp} = h/(2\pi i)\,\mathbf{1}\,.$$

The famous 'Dreimännerarbeit' ('three men's paper') [BHJ26] – a collaboration of Born, Heisenberg and Jordan – later provided a logically consistent exposition of the *Matrix Mechanics* formalism, including applications.

Some of these early papers on quantum theory are translated in [Wae67], accompanied by a historical commentary. Our remarks on the development of quantum mechanics also follow this work.

Elaborating on Heisenberg's work, Dirac first mentioned the relation between Born and Jordan's 'quantum condition' and the classical Poisson bracket in [Dir25, section 4]:

$$\hat{q}_r\hat{p}_s - \hat{p}_s\hat{q}_r = \frac{ih}{2\pi}\,\delta_{rs} = \frac{ih}{2\pi}\,\{q_r, p_s\}\ .$$

The generalisation of this correspondence lies at the heart of the *canonical quantization* programme. Dirac proposed this recipe as means to obtain the 'quantum conditions' in the general case in his seminal book [Dir58], first published in 1930. After introducing the 'mathematical scheme connecting states and observables in quantum mechanics' (the bra-ket formalism in abstract Hilbert space with quantum observables realised as self-adjoint operators), he suggested [Dir58, p. 87]:

> 'The problem of finding quantum conditions now reduces to the problem of determining P.B.s [Poisson brackets] in quantum mechanics. The strong analogy between the quantum P.B. [quantum commutators] defined by (7) and the classical P.B. defined by (1) leads us to make the assumption that the quantum P.B.s, or at any rate the simpler ones of them, have the same values as the corresponding classical P.B.s. […]'

The canonical quantization method hence instructs us to replace classical observables that satisfy a classical Poisson bracket $\{u, v\} = w$ by self-adjoint quantum operators on Hilbert space so that the commutation relation $[\hat{u}, \hat{v}] = i\hbar\hat{w}$ holds.

Throughout this thesis we will use the term 'conventional quantum mechanics' as a label for standard text book quantum mechanics over $\mathbb{R}^n$ (see [Sak94; Scho7; CDL99]).

As history has shown, Dirac's canonical quantization technique works reasonably well in the case of 'conventional' quantum mechanics over $\mathbb{R}^n$. The primary reason for this is, however, that the underlying configuration space $\mathbb{R}^n$ is so well-behaved. When we try to quantize classical systems with phase spaces other than the cotangent bundle $T^*\mathbb{R}^n$, the situation changes dramatically. Already in classical mechanics, phase spaces different from $T^*\mathbb{R}^n$ require a more elaborate mathematical formalism, but this is not the only difficulty. Because global and topological aspects play a much bigger role in quantum theory than in classical physics – visible in form of the non-local nature of wave functions – it happens frequently that certain (even basic) classical observables have no appropriate quantum counterparts. In those cases, a straightforward application of the canonical quantization recipe becomes impossible and a more sophisticated approach is needed.

In practice, a better understanding of the relation between classical and quantum mechanics is of particular interest in the light of ongoing attempts to quantize general relativity (see [Woo09] for a review), where a definitive answer to the question of the correct quantum theory of gravitation is still missing. The inherently nonlinear nature of general relativity makes a perturbative description problematic, if not impossible. A quantization procedure that is able to take the nonlinear structure into account right from the outset is a useful tool to construct and study possible candidates for a theory of quantum gravity. Quantum gravity was, in fact, Isham's key motive in [Ish83].

In addition, we expect a better knowledge of quantization on arbitrary phase spaces to be of practical advantage in calculations when physical systems satisfy constraints or boundary conditions. Similar to Lagrangian mechanics, our goal is to make the side conditions an *implicit* part of the formalism. Of course, all calculations could be done with side conditions enforced by explicit means, like in Newtonian mechanics, but it is much easier if the formalism handles them for us. With that in mind, quantization is studied not only for the sake of novel predictions; it is equally rewarding to reproduce existing results in a more illuminating manner. As an additional benefit, examples that

reproduce well-known results let us verify under controlled circumstances that a given quantization method really implements the side conditions in the correct manner before we enter new and uncharted territory.

Finally, being the main motivation for the present dissertation, the quantization of classical systems with non-trivial phase spaces provides an opportunity to learn more about the *structure* of quantum theory, like the geometric formulation of Hamiltonian mechanics on arbitrary manifolds lead to a better understanding of the structure of classical mechanics. Given only the example of standard quantum mechanics over $\mathbb{R}^n$ it is impossible to decide which properties are truly universal and which just 'artifacts' of the specific configuration space $\mathbb{R}^n$. When constructing new quantum theories, however, it is crucial to know which features really matter. In particular, the group-theoretical quantization scheme supports the idea of G-theory, originally formulated in [Hei+87; Hei+89] and later generalised by [Sła91], which emphasises the role of group actions and symmetries in a maximal way. Recent applications include [KPRL08; PRL10].

As a long-term goal, we believe the study of quantization might also be a step in the right direction to replace the set of mathematical axioms on which quantum theory is currently based by some well-motivated physical principles [Rov96; Fuc01]. As Blau phrases it [Bla92, section 1]: 'Perhaps, ultimately, the study of quantization will tell us enough about quantum theory itself to allow us to do away with the very concept of quantization.'

∞

Let us give a short outline of this thesis. In the first chapter, we will start with a brief review of the mathematical formalisation of canonical quantization in terms of the *Dirac quantization map*. The well-known theorem of Groenewold and Van Hove, however, states that such a 'full' quantization is impossible. Thus, we need to ask what a reasonable quantization method *can*, and what it can *not* accomplish. To this end, we will examine the individual mathematical assumptions that appear in the definition of the Dirac quantization map and try to determine their physical motivation. Our goal is to make clear what is sensible to aim for.

In addition, the first chapter clarifies some points of the mathematical formalism of quantum mechanics where standard physics text books are insufficient to our cause. Of crucial importance to an understanding of quantization on non-trivial phase spaces is the proper distinction between *self-adjoint* and merely *symmetric* operators, discussed in section 1.6. While it is usually excusable to neglect the difference in conventional quantum mechanics over the mathematically trivial configuration space $\mathbb{R}^n$, we will see that the existence of self-adjoint (but not of symmetric!) operators is linked directly to global and topological aspects, due to the Stone theorem. Stone's theorem explains why a straightforward application of Dirac's canonical quantization method which does not take these aspects into account becomes impossible in such cases. We will describe how the Canonical Group Quantization method with its focus on group structures is able to handle those difficulties in a very natural and systematic manner. A brief outline of the group-theoretical quantization scheme on a conceptual level is given in section 1.8.

The following chapters 2, 3 and 4 are more of a supplementary nature. In chapter 2 we provide a review of basic differential geometry using the calculus of differential forms. Chapter 3 deals with geometric aspects of (Lie) groups, Lie algebras and group actions on manifolds. As a more advanced topic, section 3.6 introduces semidirect products and general group extensions, necessary for a proper understanding of the quantization method. Chapter 4 is dedicated to symplectic manifolds and the geometric formulation

of Hamiltonian mechanics, and sections 4.4 and 4.5 discuss the important special case of cotangent bundles. A reader accustomed to those topics may quickly skim over these chapters to familiarise himself with our notation or jump straight to chapter 5.

Chapters 5 and 6 constitute the main part of this dissertation. In chapter 5 we will analyse Isham's group-theoretical quantization technique in a systematic and rigorous manner. Of particular interest to us is a certain exceptional case where the Canonical Group must be defined via a central Lie algebra extension. The existing construction is conceptually rather unsatisfying at this point and Isham himself calls this special case 'pathological'. Our intention is to resolve this irritating point of Isham's proposal and to shed some light on the physical meaning of this exceptional case. We will also clarify the role of the two distinct group structures that play their part in the group-theoretical quantization scheme.

Finally, chapter 6 is devoted to quantization in the presence of boundary conditions. Using the group-theoretical method, we will study the problem of direct quantization for a particle moving on a half-line with 'hard wall' boundary condition. This toy model of a simple contact interaction provides us with a unique opportunity to verify the correct realisation of the boundary condition at the end of the half-line as we expect a sensible quantum theory that implements the boundary condition implicitly to produce the same predictions as standard quantum mechanics restricted to the spatial region $\mathbb{R}^+ \subset \mathbb{R}$ by an explicit 'hard wall' potential.

The half-line model has recently gained some attention because of its mathematical intricacies [BFV01; FCT02; GK04; BW10]. It is well-known, for instance, that the usual momentum operator $\hat{p} = -i\hbar\partial_q$ is not self-adjoint on the Hilbert space $L^2(\mathbb{R}^+, \mathrm{d}q)$ and a common suggestion is to 'resolve the paradox by acknowledging the existence of the rest of the real line' [GK04]. Still, this is not a very satisfying solution as wave functions vanish on the negative part of the real line. Moreover, the missing momentum observable actually makes sense from a physical point of view: momentum eigenstates do not satisfy the boundary condition imposed by a 'hard wall' potential. Our goal is hence a *direct* quantization that does not require the forbidden part of the real line.

In section 6.1, we will discuss Isham's attempt to construct a quantum theory over the half-line based on the 'obvious' phase space $T^*\mathbb{R}^+$ [Ish83, section 4.5]. We will see that Isham's representation of the quantum operators is incompatible with an embedding of the half-line $\mathbb{R}^+ \subset \mathbb{R}$ as half of the full line. In section 6.2, we will therefore work out the relation between Isham's result and a different representation of the quantum operators that can be obtained from restricting the position space representation of conventional quantum mechanics over $\mathbb{R}$ to a sub Hilbert space over $\mathbb{R}^+$.

Although this fixes the representation problem, section 6.3 shows that the quantum theory constructed in this manner is incomplete, unable to describe the reflection at the end of the half-line which is an important trait of the reference model with the explicit 'hard wall' potential. To solve this shortcoming, we construct a different phase space that emerges naturally on closer examination and realises the boundary condition as a topological feature: the orbifold $\mathbb{R}^2/\mathbb{Z}_2$. In section 6.4 we study quantization on $\mathbb{R}^2/\mathbb{Z}_2$ via covering groups of the non-compact semisimple Lie group $\mathrm{SL}_2\mathbb{R}$. We will demonstrate that the proposed classical model when quantized using the group-theoretical method yields a quantum theory that correctly realises the 'hard wall' boundary condition and we will identify the specific Canonical Group that must be used.

1 The Quantization Problem

> To live in the world without becoming
> aware of the meaning of the world is
> like wandering about in a great library
> without touching the books.
>
> *Manly Palmer Hall*, The Secret Teachings of All Ages

We begin our discussion with the mathematical formalisation of Dirac's canonical quantization programme in terms of the so-called *Dirac quantization map*. However, it is well-known that such a 'full' quantization doesn't in general exist: the mathematical requirements are inconsistent. That being said, we will see that some of the supposedly 'natural' assumptions are actually rather poorly justified from a physical point of view, a matter more of mathematical convenience than fundamental physical principles. In an attempt to understand what a quantization method *can*, and what it *cannot* accomplish we will study the physical motivation behind the assumptions, as far as possible. As result, we will give a rough description of what we can expect from a quantization method founded on *physical* principles. In the last section, we will provide a short outline of Isham's *Canonical Group Quantization* scheme from a birds-eye perspective and discuss how the group-theoretical quantization method embraces these ideas.

A second goal of this chapter is to discuss the mathematical framework of quantum theory at points where standard physics textbooks are insufficient to our cause. Some simplifications that are more or less acceptable for the trivial configuration space $\mathbb{R}^n$ are no longer justified in case of physical systems with non-trivial phase spaces, where global and topological aspects enter the arena. Of great importance is the difference between *self-adjoint* and merely *symmetric* operators, discussed in section 1.6, which is the major motivation for a *group*-based approach to quantization.

1.1 Canonical Quantization à la Dirac

We already discussed Dirac's *canonical quantization* recipe in the introduction. For the following, however, it is convenient to give a more formal definition of Dirac's quantization method. A commonly agreed upon is the so-called *Dirac quantization map* (see e. g. [TAE05, section 1.1] or [Woo97, section 8.1]). The explicit definition includes an *irreducibility condition*, which we want to formulate using the notion of *complete sets* of observables [Bla92, section 2.2] in order to avoid unnecessary restrictions.

Definition 1.1.1 (Complete set) Given a Lie algebra $(\mathcal{A}, [\,\cdot\,,\,\cdot\,])$, consider a set of Lie algebra elements $C = \{a_i\} \subseteq \mathcal{A}$. We say that C is a ***complete set*** (***in the algebra*** $\mathcal{A}$) if for every $a \in \mathcal{A}$ we have:

$$[a_i, a] = 0 \quad \forall a_i \in C \qquad \Rightarrow \qquad a \propto 1_{\mathcal{A}}\,,$$

i. e. if a commutes with all elements of C it has to be proportional to the identity.

The motivation for complete sets comes from *Schur's lemma*. The variant given in [Sim96, theorem II.4.1] states that, if a *-representation U of the group algebra is irreducible, then the representing operators $U(f)$ form a complete set.

In other words, a set $C \subseteq \mathcal{A}$ is called complete if and only if there are no nontrivial subspaces $S \subseteq \mathcal{A}$, other than $\{0\}$ and $\mathcal{A}$ itself, which are invariant under the action of all elements in C (a subspace is said to be *invariant under* $c \in C$ when $cS = \{ca : a \in S\}$ is a subset of S). The classical observables of position and momentum, q_i and p_j, form

such a complete set under the Poisson bracket, as do the associated quantum operators $\hat{q}_i$ and $\hat{p}_j$ under the commutator.

Definition 1.1.2 (Dirac Quantization Map) Let M be a classical phase space and let $\mathcal{H}$ be the corresponding quantum-mechanical Hilbert space (a separable complex Hilbert space). The ***Dirac quantization map*** is then a mapping of classical observables (i. e. smooth, real-valued functions over M) to operators on Hilbert space:

$$\mathcal{Q} : C^\infty(M, \mathbb{R}) \to \mathrm{Op}(\mathcal{H}) ,$$

with the following properties:

(q1) the map $\mathcal{Q}$ is $\mathbb{R}$-*linear*,

(q2) the operators $\mathcal{Q}(f)$ for $f \in C^\infty(M, \mathbb{R})$ are *essentially self-adjoint*,

(q3) $\mathcal{Q}$ maps *Poisson brackets* to *commutators* such that:

$$[\mathcal{Q}(f), \mathcal{Q}(g)] = i\hbar\mathcal{Q}(\{f, g\}) , \tag{1.1}$$

(q4) the map $\mathcal{Q}$ is *irreducible* in the sense that any complete set of classical observables $\{f_1, \ldots, f_k\}$ is mapped to a complete set of operators $\{\mathcal{Q}(f_1), \ldots, \mathcal{Q}(f_k)\}$,

(q5) the constant function 1_M on M maps to the identity on $\mathcal{H}$, i. e. $\mathcal{Q}(1_M) = \mathbb{1}_{\mathcal{H}}$.

In the case of conventional quantum mechanics over the configuration space $\mathbb{R}^n$ the irreducibility condition (q4) guarantees that the well-known Schrödinger representation of the canonical commutation relations in position space (that is, the representation on $\mathcal{H} = L^2(\mathbb{R}^n, \mathrm{d}^n x)$ with operators realised as $\hat{q}_j = q_j$ and $\hat{p}_j = -i\hbar\,\partial/\partial q_j$) is *essentially unique*. A proof of this statement was first outlined by Stone [Sto30] and later elaborated on by von Neumann [Neu31]. A good contemporary presentation of the now famous *Stone–von Neumann theorem* can be found in [Ros04].

We will discuss the meaning of 'essentially unique' more thoroughly in the second part of section 1.6 but we do need some prerequisites before that.

Remark 1.1.3 (von Neumann rule) There is sometimes an additional requirement, the so-called *von Neumann rule*. It is a condition of the form:

$$\mathcal{Q}\big(\varphi(f)\big) = \varphi\big(\mathcal{Q}(f)\big) , \tag{1.2}$$

for some functions $\varphi : \mathbb{R} \to \mathbb{R}$ and some observables f. However, there seems to be no consensus in the literature on what is meant by 'some' functions and 'some' observables. In the original work of von Neumann [Neu68, section III.1, (F.)] equation (1.2) is supposed to hold for *all* physical observables and for *all* functions φ (as long as both sides of the equation are well-defined), while others have much weaker conditions. For example, they may only require $\mathcal{Q}(q^n) = \mathcal{Q}(q)^n$ for the position and an analogous condition for the momentum p.

In essence, the von Neumann rule preserves the *multiplicative* structure of the observables, whereas the condition (q3) preserves the Lie bracket.

1.2 Problems of the Dirac Quantization Map

Although quite successful in applications, the canonical quantization method has some severe shortcomings from a theoretical point of view. The most significant is that *a Dirac quantization map satisfying the above conditions doesn't exist*! This is the famous result of Groenewold [Gro46] and Van Hove [VH00]. Adapted to our definition of the Dirac quantization map it says the following:

Theorem 1.2.1 (Groenewold, Van Hove) The postulates (q1), (q3) and (q4) as they are given in definition 1.1.2 are inconsistent, even if we restrict the set of classical observables to polynomials in q and p of degree ≤ 4.

The following proof is essentially the one given in [TAE05, section 1.2]. We repeat their argument in this work as it gives some insights on how to proceed later on. To save us some indices we consider only a one-dimensional configuration space. Nevertheless, the argument translates naturally to higher dimensions, so this is no loss of generality.

Proof. We will employ two different Poisson brackets to quantize the classical observable p^2q^2 using the given conditions, namely:

$$\{q^3, p^3\} = 9p^2q^2 \,, \quad \{q^2p, qp^2\} = 3p^2q^2 \,. \tag{α}$$

First note, as a result of (q3) we still have the usual canonical commutation relations of $\hat{q} = \mathcal{Q}(q)$ and $\hat{p} = \mathcal{Q}(p)$ at our disposal:

$$[\hat{q}, \hat{p}] = i\hbar \,, \quad [\hat{q}, \hat{q}] = 0 \,, \quad [\hat{p}, \hat{p}] = 0 \,. \tag{β}$$

To apply the Poisson brackets in (α) we further need to quantize all the terms that occur *inside* the brackets. This is done in several consecutive steps:

'pq': Consider the Poisson brackets $\{pq, q\} = -q$ and $\{pq, p\} = p$. If we temporarily write $X = \mathcal{Q}(pq)$, the correspondence rule (q3) yields the commutators:

$$[X, \hat{q}] = -i\hbar\hat{q} \,, \quad [X, \hat{p}] = i\hbar\hat{p} \,.$$

Yet, these commutators are also fulfilled by the operator $X' = \frac{1}{2}(\hat{p}\hat{q} + \hat{q}\hat{p})$ if we apply the canonical commutation relations of $\hat{q}$ and $\hat{p}$ from (β). Hence the difference $X - X'$ *commutes* with both $\hat{q}$ and $\hat{p}$, and $X - X'$ must then be a multiple of the identity according to the irreducibility condition (q4). Thus we obtain:

The irreducibility of $\mathcal{Q}$ implies that $\hat{q}$ and $\hat{p}$ form a complete set (definition 1.1.1).

$$X = \mathcal{Q}(pq) = \tfrac{1}{2}(\hat{p}\hat{q} + \hat{q}\hat{p}) + c\mathbb{1} \,,$$

with some constant $c \in \mathbb{C}$.

'q^m and p^m': Pick the Poisson brackets $\{q^m, q\} = 0$ and $\{q^m, p\} = mq^{m-1}$ and let us write $Y_m = \mathcal{Q}(q^m)$. The correspondence rule (q3) together with the linearity (q1) yields:

$$[Y_m, \hat{q}] = 0 \,, \quad [Y_m, \hat{p}] = mi\hbar Y_{m-1} \,.$$

Since both commutators are also fulfilled by $Y'_m = \hat{q}^m$ we can use the same argument as above to see that Y_m and Y'_m differ only by some constants $d_m \in \mathbb{C}$:

$$Y_m = \mathcal{Q}(q^m) = \hat{q}^m + d_m\mathbb{1} \,.$$

In this case it is possible to show that the constants d_m vanish. Using the expressions for X and Y_m from above and the Poisson bracket $\{q^m, pq\} = mq^m$ we get:

$$\begin{aligned} mi\hbar Y_m &= [Y_m, X] = \left[\hat{q}^m + d_m\mathbb{1}, \tfrac{1}{2}(\hat{p}\hat{q} + \hat{q}\hat{p}) + c\mathbb{1}\right] \\ &= \left[\hat{q}^m, \tfrac{1}{2}(\hat{p}\hat{q} + \hat{q}\hat{p})\right] = \tfrac{1}{2}\left([\hat{q}^m, \hat{p}]\,\hat{q} + \hat{q}\,[\hat{q}^m, \hat{p}]\right) = mi\hbar\hat{q}^m \,, \end{aligned}$$

and hence $Y_m = \hat{q}^m$. An analogous calculation can be done for p^m; together we have:

$$\mathcal{Q}(q^m) = \hat{q}^m \,, \quad \mathcal{Q}(p^m) = \hat{p}^m \,. \tag{γ}$$

Observe that (γ) is a special case of a *von Neumann rule* (see remark 1.1.3).

'q^2p and qp^2': Given the Poisson brackets $\{q^3, p^2\} = 6q^2p$ and $\{q^2, p^3\} = 6qp^2$ these products can be reduced to expressions containing only the powers of q and p for which quantizations are known from (γ):

$$6i\hbar\mathcal{Q}(q^2p) = [\mathcal{Q}(q^3), \mathcal{Q}(p^2)] = [\hat{q}^3, \hat{p}^2] = [\hat{q}^3, \hat{p}]\,\hat{p} + \hat{p}\,[\hat{q}^3, \hat{p}] = 3i\hbar(\hat{q}^2\hat{p} + \hat{p}\hat{q}^2)\,.$$

After an analogous calculation involving the second bracket we obtain:

$$\mathcal{Q}(q^2p) = \tfrac{1}{2}(\hat{q}^2\hat{p} + \hat{p}\hat{q}^2)\,, \quad \mathcal{Q}(qp^2) = \tfrac{1}{2}(\hat{q}\hat{p}^2 + \hat{p}^2\hat{q})\,. \tag{δ}$$

Now we finally have everything we need to quantize the product p^2q^2 using the Poisson brackets mentioned in (α). Via the first one we get:

$$9i\hbar\mathcal{Q}(p^2q^2) = [\mathcal{Q}(q^3), \mathcal{Q}(p^3)] = [\hat{q}^3, \hat{p}^3]\,,$$

which can be further evaluated using the commutation relations (β) and then yields:

$$\mathcal{Q}(p^2q^2) = \hat{q}^2\hat{p}^2 - 2i\hbar\hat{q}\hat{p} - \tfrac{2}{3}\hbar^2\mathbb{1}\,. \tag{ε}$$

On the other hand, if we use the second Poisson bracket from (α) and the results from equation (δ), we obtain:

$$3i\hbar\mathcal{Q}(p^2q^2) = \big[\mathcal{Q}(q^2p), \mathcal{Q}(qp^2)\big] = \big[\tfrac{1}{2}(\hat{q}^2\hat{p} + \hat{p}\hat{q}^2), \tfrac{1}{2}(\hat{q}\hat{p}^2 + \hat{p}^2\hat{q})\big]\,,$$

which again by (β) results in:

$$\mathcal{Q}(p^2q^2) = \hat{q}^2\hat{p}^2 - 2i\hbar\hat{q}\hat{p} - \tfrac{1}{3}\hbar^2\mathbb{1}\,. \tag{ε'}$$

This is obviously not the same as (ε); both results differ by a term of $\frac{1}{3}\hbar^2\mathbb{1}$! This means the given conditions are inconsistent with each other and the quantization map is therefore ill-defined. ■

Remark 1.2.2 Other no-go theorems exist for various combinations of the conditions in definition 1.1.2 and the von Neumann rule from remark 1.1.3 (a review can be found in [TAE05]). The proofs usually look similar to the one given above: take some product of q's and p's and quantize it via the rules in question to obtain a contradiction.

The theorem shows that a full Dirac quantization map doesn't exist. More important, however, is the fact that the requirements turned out to be inconsistent even though they all 'look' kind of 'natural'. This means our intuition about the proper mathematical description is wrong. In the following sections we will thus study what the conditions actually *mean* from a *physical* point of view – which ones really *are* 'natural', and which ones are to be discarded.

That being said, the proof of the Groenewold–Van Hove theorem already gives some hints how a contradiction can be avoided from a mathematical point of view:

Remark 1.2.3 (Geometric and Deformation quantization)

- An obvious way to circumvent the conclusion from theorem 1.2.1 is to restrict the domain of $\mathcal{Q}$ to contain only *some quantizable observables* which form a *subset* of the functions on phase space:

$$\mathrm{Obs}(M) \subset C^\infty(M, \mathbb{R})\,.$$

The product p^2q^2 which appeared in the proof then simply ceases to be quantizable and hence the contradiction vanishes. Furthermore, it is possible to keep a larger set of quantizable observables if one in turn relaxes some of the other requirements. This is roughly the approach taken by **Geometric Quantization** (see [Woo97]). The method of Canonical Group Quantization also falls into this category. I will argue in section 1.4 why I understand this restriction of the set of quantizable observables to be a physically reasonable assumption.

- A different route is taken by **Deformation Quantization**. The underlying observation is that the two contradicting results in the proof of theorem 1.2.1 differ only by a term of order $\hbar^2$. The idea is hence to modify the correspondence (q4) between Poisson brackets and quantum commutators by adding terms in higher orders of $\hbar$ that disappear in the limit $\hbar \to 0$, i. e.

  $$[\mathcal{Q}(f), \mathcal{Q}(g)] = i\hbar\mathcal{Q}(\{f, g\}) + \mathcal{O}(\hbar^2) \,.$$

  This is usually accomplished by means of a 'deformed product' (called *star* or *Moyal product*), which is defined on the classical phase space. Details on this method can be found, for instance, in the textbook written by Waldmann [Wal07].

Notice, however, that $\hbar \to 0$ is not a true classical limit (see [DHS00, section 4.8]). A physically reasonable classical limit requires a more sophisticated approach (see e. g. [Lan98]).

Remark 1.2.4 There is actually another difficulty with canonical quantization when it comes to configuration spaces other than $\mathbb{R}^n$. Consider, for example, a particle that is restricted to move on the positive real line. The configuration space is $Q = \mathbb{R}^+$. It seems reasonable to use the position q and momentum p as classical observables, which satisfy the usual commutation relations. However, when we try to represent these by operators $\hat{q} = q$ and $\hat{p} = -i\hbar\,\partial/\partial q$, it turns out that the momentum operator $\hat{p}$ is *not self-adjoint* on the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^+, \mathrm{d}q)$. Thus, a straightforward application of Dirac's canonical quantization recipe is impossible.

See section 1.6 for the definition of self-adjointness.

The reason is that the exponentiated momentum operator $U(a) := e^{-ia\hat{p}}$ shifts wave functions to the negative numbers (for $a < 0$) according to:

$$(U(a)\psi)(x) = \psi(x - \hbar a) \,,$$

which in this case implies a 'loss' of probability as shown in figure 1.1. Given that unitary operators preserve probabilities it is clear that the operator $U(a)$ *cannot* be unitary. This result, however, contradicts *Stone's theorem* which states that the exponential $e^{ia\hat{x}}$ of any

This argument is quite useful as it allows us to reason about possible *unitarity* or *self-adjointness* of operators in an intuitive way.

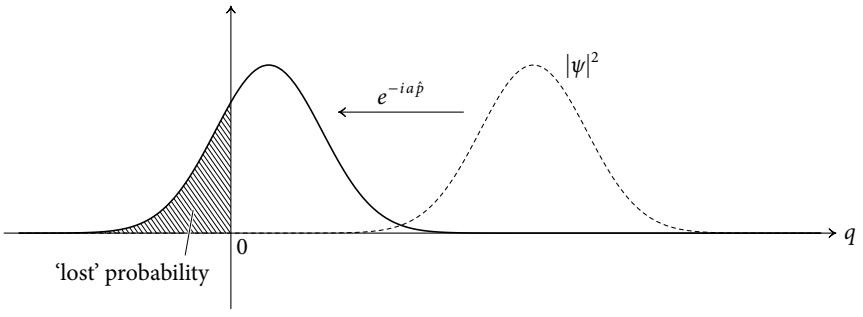


Figure 1.1 The operator $e^{-ia\hat{p}}$ can 'shift' the wave function ψ to the negative numbers, and hence probability is not preserved if we consider the Hilbert space $L^2(\mathbb{R}^+, \mathrm{d}q)$ over only the *positive* real numbers.

self-adjoint operator $\hat{x}$ is a one-parameter group of unitary transformations (see [Sto30] or theorem 1.6.15 on page 28). Consequently, the operator $\hat{p}$ cannot be self-adjoint.

We are left with the question *how to find* classical observables that can be quantized to yield meaningful self-adjoint operators. The canonical quantization programme of Dirac doesn't provide an answer. We will see later how the *Canonical Group Quantization* scheme deals with such phase spaces.

1.3 Some Words on 'Simplicity'

The requirements in the definition of the Dirac quantization map are based on an intuition of what looks 'natural'. As the last section has proven this intuition to be wrong, we need to ask ourselves about the causes of this misjudgement.

When modelling a physical theory there are several guiding principles involved. First of all, a physical theory should of course describe empirical observations. Nevertheless, if two competing theories describe the same matter in different ways we usually assume the *simpler* one has to be more fundamental. Yet how do we *decide* which theory is 'simpler'? What do we mean when we say that a theory is 'simpler' than another?

At this point some heuristic notions of *economy*, *elegance* and *beauty* enter the field. In particular the development of modern theories – like string theory, quantum gravity and extensions of the standard model of particle physics – is heavily influenced by some idea of *mathematical beauty*. Sometimes this principle of mathematical beauty even comes *before* other considerations. This was also Dirac's opinion during the time quantum mechanics was developed [Dir38]:

Rota discusses the concept of *mathematical beauty* in his amazing book [Rot97].

> '[…] The method is to begin by choosing that branch of mathematics which one thinks will form the basis of the new theory. One should be influenced very much in this choice by considerations of mathematical beauty. […]'

As is well-known, this 'mathematical approach' to physics had, and still has, quite some successes, in particular when symmetry principles are involved.

The *rule of simplicity* is also known as *law of parsimony* or as *Ockham's razor* in the literature.

Despite this, we should be careful how we apply this 'rule of simplicity'. Although it has proven useful when developing new concepts it can also lead to wrong decisions (some examples are given in [Ger07] and [Giu08]). This happens specifically if we look for the 'wrong kind' of simplicity. As such, the question whether a theory possesses some desirable mathematical properties doesn't necessarily say if the theory is correct or not. Perturbative renormalizability, for example, is certainly a desirable and useful property of a quantum field theory, as it allows us to do perturbative calculations. Nevertheless, we should always keep in mind that perturbation theory is a *mathematical tool* and not an actual physical process.

Another problem we will have to face in our study of quantization is the difference between what physicists *say* compared to what physicists *do*. To give an example, most textbooks on quantum mechanics claim that pure states can be described by vectors in the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^n, \mathrm{d}^n x)$ of square-integrable functions (see e. g. [CDL99, section 3.2.1], [Sak94, section 1.2]). Although in principle correct, the same books then use *plane waves* to describe free particles, which are clearly *not* square-integrable. We will demonstrate in section 1.5 how this inconsistent use of mathematical concepts can lead to contradictions, and how those issues can be avoided.

The goal of the remaining part of this chapter is two-fold. On one side, we want to have a closer look at the mathematical foundations of quantum mechanics. We will

try to clarify some finer points of the mathematical concepts relevant to quantization attempts and explain how the mathematical framework can be brought into accordance with real practical calculations. The other goal is to give some explanations – or at least good motivations – *why* the mathematical concepts we are using are necessary from a *physical* point of view.

1.4 Measurements and Quantizable Observables

To begin with, we already pointed out in section 1.2 that it is possible to obtain a self-consistent quantization method if we restrict the set of quantizable observables. We want to explain why this restriction is physically quite reasonable. First, though, we need to determine what the conceptual meaning of the term *quantizable observable* is. While words like 'observable' and 'measurements' directly spring to mind, it turns out that these are not the best associations to think of.

1.4.1 Observables and Measurements

We will start with a simple working definition of an *observable*, which can be used for both classical and quantum observables (see [Omn99, chapter 8, p. 96]):

> A physical observable is a physical quantity of a system that can in principle be observed and measured.

This definition immediately raises some new questions: What is meant by *physical quantity*? What is an *observation* or a *measurement*? – It is an interesting point that these questions are usually only posed for *quantum observables*. On the other hand, there seems to be a tacit agreement that *all functions on classical phase space* correspond to *classical observables* and can be measured as a matter of principle.

When it comes to how actual experiments work, however, it turns out that only a few quantities can really be measured *directly* (even in classical physics). Consider for instance how the *velocity* of a particle is 'measured'. In fact, we don't measure velocity directly. What we typically *do* measure is *positions* of the particle at distinct times, and then we *calculate* the velocity from this data. Another example: when we say that we 'measure' the *momentum* of a charged particle, we will most often *measure* the particle trajectory when subjected to a magnetic field (e. g. in a wire chamber; again a series of position measurements) and then *calculate* the momentum from the radius of curvature of the trajectory. Even time measurements can finally be tracked down to position measurements of an oscillating system (most obviously the pointer of analog clocks).

To go to the extreme, Penrose remarks in [Pen05, section 21.8] that 'some physicists have indeed taken the view that all measurements are ultimately measurements of position', and he points the reader to Goldstein [Gol87] and Bell [Bel04] for references. The latter writes in [Bel82]: 'it is always positions that we are in the end concerned with', and 'in physics the only observations we must consider are position observations, if only the positions of instrument pointers'. While certainly quite radical, this view is consistent with the way measurements were presented above, and as far as I can tell there aren't any counterexamples, neither for classical nor for quantum-mechanical measurements.

There is also a quite interesting discussion of measurements and the role of so-called *device observables* in the book of Dubin et al. [DHS00, chapter 16].

On the other end of the spectrum, von Neumann proposes a quite different interpretation of (quantum-mechanical) observables [Neu68, section IV.1]:

> 'Unter einer Größe ist eigentlich die Anweisung zu verstehen, wie sie zu messen ist – und wie ihr Wert aus den Zeigereinstellungen der Meßinstrumente abzulesen bzw. zu berechnen ist.'

('an [observable] quantity must rather be seen as a set of instructions how to measure it – and how its value is to be read off the indicator positions of measuring devices, or calculated therefrom.')

There are two interesting points to remark about this: First, von Neumann notes that what we actually measure *directly* are – again – indicator *positions*. The values of physical quantities are then *calculated* from these results. This perspective fits perfectly with our ongoing discussion if we make the distinction between ***primary observables***, for the quantities that can be measured directly, and ***secondary observables***, for the quantities that have been calculated from these, and so are measured only *indirectly*.

On the other hand, von Neumann assumes that (secondary) observables are really *nothing more* than measurement instructions. What speaks against this view is the fact that the same observables appear throughout the underlying theory, including parts which aren't directly related to measurements. One could argue, of course, that quantum mechanics is a theory *exclusively* about measurements. Still, while this perspective might be compatible with the Copenhagen interpretation, there are some alternative formulations of quantum mechanics where the measurement process doesn't play such a special role (e. g. the pilot-wave formulation of Bohmian mechanics [Boh52a; Boh52b], or ideas surrounding the concept of *decoherence* [Sch04]). Moreover, we were also searching for a definition of *classical* observables and it seems to me quite questionable to put such an emphasis on measurements in classical mechanics.

A good review of some alternative formulations and interpretations can be found in [Pen05, chapter 29].

1.4.2 Quantizable Observables

On the other hand, the basic observation that justifies the study of quantization is that *some quantities* of the classical theory have counterparts in a somehow associated quantum theory. In conventional quantization on $\mathbb{R}^n$, for instance, we have position q, momentum p and a constant, corresponding to quantum operators $\hat{q}$, $\hat{p}$ and a multiple of $\mathbb{1}$. Beyond these, some additional quantities can be quantized but the Groenewold–Van Hove theorem showed that we will quickly run into problems.

Quantization, however, isn't really about measurements. Hence, for the purposes of quantization I want to propose a classification of physical quantities as in figure 1.2. The underlying idea is to split up the physical theory into several distinct parts:

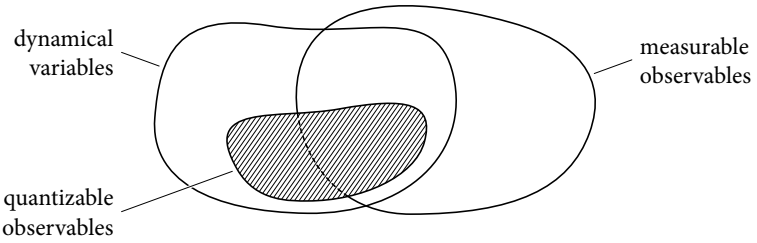


Figure 1.2 Quantizable observables seen as a subset of the dynamical variables. The difference between dynamical variables and measurable observables is first of all a conceptual one. The mathematical descriptions usually overlap.

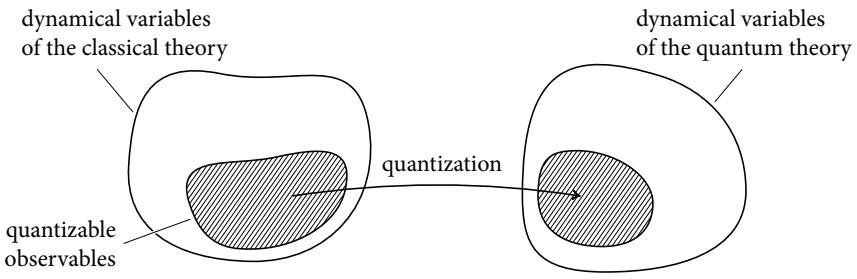


Figure 1.3 The correspondence established by a quantization method may not hold for all dynamical variables but only for a subset of *quantizable observables*.

$$\text{theory} = \text{kinematics} + \text{dynamics} + \text{measurements}\ .$$

Within the scope of this decomposition, a quantization procedure essentially allows us to construct the *kinematical* and the *dynamical* part of a quantum theory based on the kinematics and dynamics of a given classical theory. The probabilistic description of *measurements* in quantum theory is then built on top of this foundation.

The idea to make a distinction between *dynamical variables* and *measurable observables* is also supported by remark 6.2.3.

The term ***dynamical variables*** in this context describes all quantities that appear in the kinematical and dynamical part of the theory. In the mathematical formulation of classical mechanics the dynamical variables are basically smooth, real-valued functions on phase space. Nevertheless, I wouldn't expect that every arbitrarily crazy function on phase space necessarily plays a dynamical role. On the other hand, it may prove useful to include quantities beyond the real-valued functions. Electrodynamics, for instance, employs *complex* vector fields to facilitate calculations. It might even be a useful trick to introduce additional dynamical variables with absolutely no measurable impact for the classical theory if these dynamical variables pick up some physical role and have measurable consequences *after* the quantization method has been applied.

A well-known example for dynamical variables which are not measurable are *ghost fields*.

In contrast to dynamical variables, I want to use the term ***measurable observables*** as being conceptually related to measurements, just as the name implies. The example of ghost fields shows that not all dynamical variables have to be measurable. On the other hand, the arbitrarily crazy function on phase space, mentioned above, with no dynamical role to play is still measurable in the sense of von Neumann's definition of observables.

The conceptual difference between dynamical variables and measurable observables is probably best made clear in the context of quantum field theory where *bare* quantities are used as dynamical variables, whereas *renormalised* quantities represent measurable observables. This example also shows that, despite the *conceptual* difference, there is usually a big overlap when it comes to the *mathematical* description (bare and renormalised operators are both operators). This is shown in figure 1.2.

Finally, if classical mechanics is seen as a limiting case of quantum mechanics, which is the prevailing opinion, it is reasonable to expect some similarities in their mathematical formulations. In particular we would expect a correspondence between at least *some basic* dynamical variables of the classical and the quantum theory (there is a more thorough discussion of this point in Mackey's book [Mac63, section 2.4]). That being said, we can't ignore that there are also structural *differences* between classical and quantum theory. Consequently, the basic correspondence might not hold between *all* dynamical variables of the classical theory and all dynamical variables of the quantum theory.

Besides, two quantities with a similar *mathematical* appearance may still play different *physical* roles in classical and quantum theory (see [DHS00, section 4.8]).

A pragmatic approach is to expect a relation to exist for a *subset* of the dynamical

Deformation quantization, for example, allows for a bigger set of quantizable observables at the expense of a 'less strict' correspondence (remark 1.2.3).

variables, the ***quantizable observables*** (see figure 1.3). The size of this set of quantizable observables depends on how 'strict' we require the correspondence to be, and hence, ultimately, on the quantization procedure. It seems physically reasonable to *assume that at least some basic dynamical variables of the classical theory are quantizable, yet not to require this for every single (arbitrarily crazy) dynamical variable*.

1.4.3 Quantization on Non-Trivial Phase Spaces?

A QUESTION RELATED to the one of quantizable observables is to ask whether any classical phase space we can think of must produce a meaningful quantum theory. Again, it is certainly desirable from a mathematical point of view, yet there are some physical arguments why this need not be the case. Although this topic has already been discussed by Isham in [Ish83, section 2], the answer will help us to better understand the *aim* of quantization. In the process, the nature of the relationship that exists between a classical and the associated quantized system will also become clearer.

To understand the issue we want to raise, remember that more general phase spaces typically appear in classical mechanics if we consider systems subject to *constraints*. If we look closely at *concrete* physical systems, however, most of these classical constraints are valid only on a *macroscopic* level but 'soften' if we examine the same system from a *microscopic* perspective. Consider for example a particle restricted to move on a circle, say the realisation of a mathematical pendulum consisting of a weight attached to a rod of a given length. Of course, this length can be seen as constant in the macroscopic system, leading to the configuration space S^1. On a microscopic level, though, it will only be constant *on average* (see figure 1.4).

Seeing this discrepancy between macroscopic and microscopic view, we have to ask ourselves if the same kinds of constraints that are possible in classical mechanics still make sense for quantum systems. In an analogous quantum system which describes the circular motion of a particle it will probably be more appropriate to use a potential to implement the constraint, instead of a restricted configuration space.

Even if the constraint *is* rigidly enforced in a quantum system, we might sometimes run into problems with self-adjointness of operators if we try to restrict the configuration space (see remark 1.2.4 for an example).

In the following discussion we will use the name ***extrinsic constraints*** for the type of constraints like the one above, because these constraints are somehow imposed 'from the outside'. A typical indicator for extrinsic constraints is that the phase space appears, conceptually, as a *sub*manifold of some bigger space.

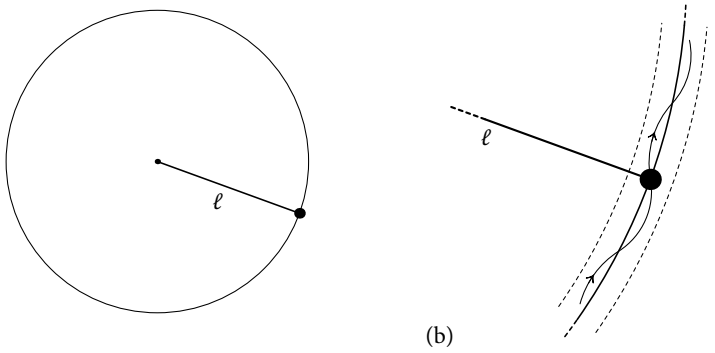


Figure 1.4 The macroscopic constraints (a) of a mathematical pendulum 'soften' if we look at them from microscopic perspective (b).

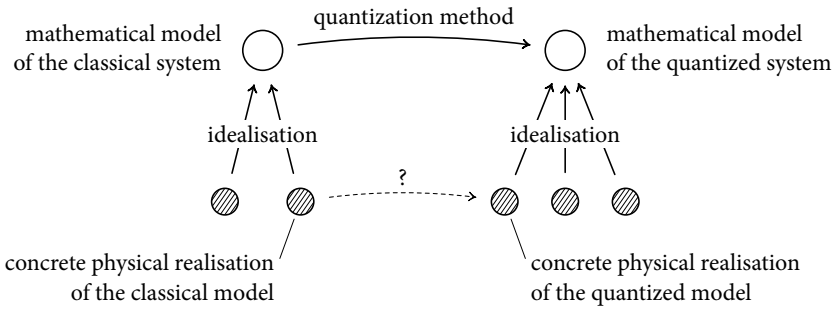


Figure 1.5 A quantization method establishes a relation between idealised mathematical models of concrete physical systems.

A different type of constraints arises from 'internal' properties of the system; hence we call them ***intrinsic constraints***. In contrast to extrinsic constraints, a phase space that occurs in the presence of intrinsic constraints appears conceptually not as subspace of a bigger space but is 'in its nature' a more general manifold.

An example featuring intrinsic constraints is a system of n indistinguishable particles moving in d spatial dimensions [LDM71; LM77; Reyo6; PRL10]. Leinaas and Myrheim pointed out in [LM77, section 2] that indistinguishability of identical particles is necessary already in the *classical* description to resolve Gibbs' entropy paradox. The proper classical configuration space would thus be:

$$Q = (\mathbb{R}^{nd} \setminus \Delta_n)/\mathbb{S}_n \,, \tag{1.3}$$

where the subtraction of Δ_n ensures that no particles occupy the same point in space:

$$\Delta_n = \left\{(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in (R^n)^d : \boldsymbol{x}_i = \boldsymbol{x}_j \text{ for at least one pair } i, j \in \{1, \ldots, n\}\right\} , \tag{1.4}$$

and dividing out the symmetric group $\mathbb{S}_n$ implements indistinguishability of particles. The dominating opinion is that quantization on such a configuration space makes sense also from a physical point of view, because the indistinguishability is expected to carry over to the quantum theory.

To look at this problem of quantizability from a different perspective, consider the conceptual model in figure 1.5. The key point is that a quantization method establishes a relation between idealised *mathematical models*. From a physical perspective, however, these models will be meaningless unless there are concrete physical systems which realise these mathematical models (at least approximately).

Think of the classical harmonic oscillator, for instance. The mathematical model in this case is an appropriate idealisation of an underlying, concrete physical system (say, a mass attached to a spring). Given the classical model, a quantization method may return a *quantized* mathematical model. To make sense of the quantized model, however, we need to look for actual, physically realised systems which the quantized model describes in some reasonable idealisation. The quantized model of the harmonic oscillator – the *quantum harmonic oscillator* – is realised by a particle moving in a harmonic potential, for instance. On the other hand, it is difficult to find an appropriate application for the quantized model of the mathematical pendulum.

That being said, *if there is no physical system which realises the quantized model, there is no physical reason for the classical model to be quantizable in the first place.*

This relation via a quantization procedure is the *reason* why we call some concrete physical system a quantum harmonic oscillator – not because of an actual *physically justified* limit that relates the two models.

As a result, for most typical examples of nontrivial configuration spaces discussed in the literature it doesn't make sense to expect them to be quantizable from a conceptual point of view because the classical system doesn't have a direct quantum equivalent. Still, it could be that the resulting quantum theory is realised by a system that has a quite different physical interpretation. In contrast, there *are* classical systems, like the model of indistinguishable particles, that should genuinely be quantizable because we expect an analogous interpretation in the quantum theory. Anyway, even if there is no sensible physical interpretation but quantization is possible mathematically, we can still use such examples as 'toy models' to learn something about the quantization method.

1.5 The Complex Structure and Quantum States

It is also possible to start with a C^*-algebra of observables. The Hilbert space, however, then appears as a derived notion.

When we introduced the Dirac quantization map in the beginning of this chapter we tacitly assumed that quantum mechanics has to be formulated in terms of operators acting on a complex Hilbert space. It also seems to be common knowledge, dating back to the birth of quantum mechanics in the 1920s, that pure states are vectors in this Hilbert space. Or was it rays?

This section is meant to shed some light on states in quantum mechanics. We will first investigate why the mathematical formulation of quantum theory seems to require a complex structure – in contrast to classical mechanics. Furthermore, we want to answer the question whether states are vectors or rays. We will also show how the concept of the Hilbert space needs be extended to accommodate all eigenstates if we want to study observables with continuous spectra (see also [Gie00] for this last part).

1.5.1 The Complex Structure in Quantum Mechanics

To begin with, the physical reason why the space of states should be a vector space is the *superposition principle*. Most textbooks on quantum mechanics then just assume that complex numbers are a necessity of the mathematical formulation, not worth any further consideration. Still, Cohen-Tannoudji et al. mention in [CDL99, section 1.1.2] that the quantum mechanical wave function is complex *in its nature*, unlike it is the case in classical electrodynamics where complex numbers are used merely for mathematical convenience. Unfortunately, the book provides no proper explanation, just some vague hints that it has something to do with the probabilistic interpretation.

An argument concerning the complex structure which is based on the probabilistic interpretation can indeed be found in Omnès' book [Omn99, chapter 5, p. 42]:

> 'A simple example may show that one cannot generally use real wave functions because of their probabilistic meaning. Consider a wave function for a particle having a definite momentum p in some reference frame. It must be a periodic function according to de Broglie, and if real, it can be written as $A\cos[(px - Et)/\hbar + \alpha]$. At time zero, the probability for observing the particle at a point x vanishes at the nodes of the cosine, which are separated by a distance $\lambda/2 = \pi\hbar/p$ [see figure 1.6]. In another reference system moving at a velocity V with respect to the first one, the momentum is $p' = p - mV$, and the points where the probability of observation vanishes are separated by a different distance $\lambda'/2 = \pi\hbar/p'$. This is obviously impossible, since distances are invariant in nonrelativistic physics. The difficulty disappears when the wave function is a complex exponential.'

Actually, the same arguments shows that there must not be *any* variations of the probability density which would allow us to deduce the momentum.

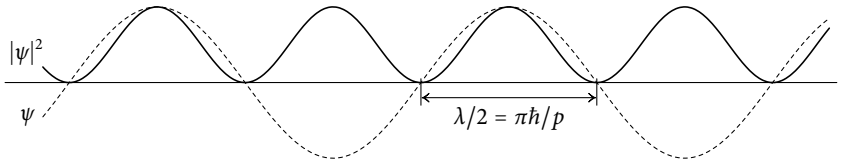


Figure 1.6 The issue with real-valued wave functions is that the probability density $\|\psi\|^2$ vanishes at distances determined by the momentum. However, this is incompatible with changes of reference frames.

This argument proves, in fact, that simple real-valued wave functions cannot be used (at least in nonrelativistic quantum mechanics). Nevertheless, it doesn't justify the need for *complex* numbers. Omnès remarks that the difficulty disappears if the wave function is allowed to be complex, yet this is not the only viable solution. It is equally possible to circumvent the problem *without* referring to a complex structure, say by use of wave functions that assume values in $\mathbb{R}^2$. This 'proof' certainly fails.

Sakurai, on the other hand, presents a completely different argument in favour of complex numbers. He discusses a Stern–Gerlach experiment featuring a beam of silver atoms and then remarks [Sak94, section 1.1]:

> '[…] we see that if we are allowed to make the coefficients preceding base kets complex, there is no difficulty in accommodating the $S_{y\pm}$ atoms in our vector space formalism […]. We thus see that the two-dimensional vector space needed to describe the spin states of silver atoms must be a *complex* vector space; […]
>
> […] we have already accomplished the main goal of this section: to introduce the idea that quantum-mechanical states are to be represented by vectors in an abstract complex vector space.'

It is true, in fact, that half-integer values of spin correspond to *complex* representations of the group $\mathrm{SU}(2)$ [Mac63]. Nevertheless, we can ask ourselves why the rest of quantum mechanics seems to require a complex structure *before* spin is even introduced. Actually, nonrelativistic quantum mechanics incorporates spin more or less like an afterthought, not as an essential ingredient. In the common formulation it requires some formal (and mathematically quite questionable) notation – the ubiquitous $\boldsymbol{\sigma} \cdot \boldsymbol{a} = \sigma_i a_i$, with Pauli matrices σ_i – which, as it turns out, *embeds* the 'spin-less' quantum mechanics in some Clifford algebra that comes with its *own* complex structure [DL03; Jun06]. Seen this way, Sakurai's argument is also not that terribly convincing. We would much prefer a reason that comes from plain quantum mechanics, without having to incorporate spin.

Another, frequently cited attempt to explain the complex structure has been made by Stueckelberg [Stu60]. He starts with a quantum theory build upon real numbers and then argues that an operator J with $J^2 = -\mathbb{1}$ needs to be introduced in order to have an *uncertainty principle*. More specifically, given two symmetric operators F and G acting on a *real* Hilbert space, he explains that the uncertainty principle, if one doesn't want to introduce a complex structure, has to be of the form $\langle(\delta F)^2\rangle\langle(\delta G)^2\rangle \geq \lambda^2\langle P\rangle$, with $P = -\left[F, G\right]^2$ and $\delta A = A - \langle A\rangle$, because the expectation value of the commutator of self-adjoint operators $\langle[F, G]\rangle$ vanishes in a real Hilbert space. In section 2 of the paper he then presents his argument, intending to rule out this possibility. This argument would in turn force us to introduce a complex structure. He argues that the only choice for λ is given by $\lambda = 0$, which yields the trivial equation $\langle(\delta F)^2\rangle\langle(\delta G)^2\rangle \geq 0$ that is obviously

For real Hilbert spaces we have $\langle F\psi|G\psi\rangle = \langle G\psi|F\psi\rangle$. In case of complex Hilbert spaces there is only a *conjugate* symmetry.

incompatible with a meaningful uncertainty principle.

However, in order to deduce $\lambda = 0$ he makes the innocent-looking assumption that the spectrum of G should be bounded. As consequence, the argument doesn't actually show that we need a complex structure to formulate an uncertainty principle; rather it results in the statement that only *unbounded* operators exhibit a nontrivial uncertainty, which is a well-established fact in functional analysis [Gro88, section 5.5]. This issue was already noted by Sharma and Coulson in [SC87]. Although they provide an idea how a proper proof following this train of thought could look like, in the end they conclude their paper with the words: 'An attempt is being made to prove this conjecture, but this is a problem in measure theory on a nondistributive lattice which is a touchy business'.

There actually exists a follow-up on this paper, [Sha88], written by Sharma alone. He doesn't, however, pursue his original proposal from [SC87] but presents a different argument:

> 'It is, of course, clear that the commutation relation
>
> $$[q, p] = i\hbar\mathbb{1}$$
>
> is the right one and can arrive at this commutation relation from a variety of physical considerations [...]. It is also clear that such a commutation relation cannot exist in a real Hilbert space unless one introduces the operator ι [with $\iota^2 = -\mathbb{1}$; a complex structure] to take the place of i.'

While true that the right-hand side of the equation features the imaginary unit i, the same i also appears on the left-hand side – although hidden for example in the expression $p = -i\hbar\partial_q$ if we choose the Schrödinger representation. It would thus be possible to circumvent Sharma's conclusion if we divide the whole equation by i and represent the momentum by $p = -\hbar\partial_q$. A related issue is that the author arrives at the conclusion that the usual commutation relation is 'of course, the right one' by citing only references that *a priori*(!) assume a *complex* Hilbert space. It is not even asked if it has to be replaced by a different equation on a real Hilbert space.

That p is then no longer a self-adjoint operator isn't that problematic because the momentum is not directly measurable (in the stricter sense of section 1.4).

In the end it seems Sharma wasn't that convinced by his own argument either. He doesn't go into the details but puts forward yet another proposal, now by means of the probabilistic interpretation of wave functions. It is in fact the same approach that we have already shown to be wrong above.

After seeing all the attempts to justify a complex structure fail for one reason or another, it seems there is currently no actual proof left (as far as I could find out, at least). Nevertheless, there are some good motivations to require a complex structure.

First, there is currently no formulation of quantum mechanics that doesn't need a complex structure at some point or another, although it has been searched for. While this could simply be dismissed as a lack of creativity, the failed attempts show that there is at least no simple way to avoid a complex structure.

Furthermore, we know from the discussion above that a simple real Hilbert space is incompatible with the probabilistic interpretation of wave functions. We already mentioned that this doesn't prove the necessity of complex numbers, but there is an argument based on the characteristic properties of quantum logic that may show that a complex Hilbert space, even if not the only option, may be the simplest choice. The idea goes back to Mackey's work [Mac63, section 2.2]:

Instead of starting out with a Hilbert space, he first establishes a correspondence between observables and 'questions' concerning the results of measurements, and then

proposes a set of physically motivated axioms for these questions in quantum mechanics. One of the main differences between classical and quantum theory is then that in the latter there exist questions that are not simultaneously answerable. In consequence, he concludes, the partially ordered set of all questions in quantum mechanics forms a so-called *σ-complete orthomodular lattice* (please refer to Mackey's book for a definition). As his 7th axiom, he then assumes [Mac63, section 2.2, Axiom VII]:

> 'The partially ordered set of all questions in quantum mechanics is isomorphic to the partially ordered set of all closed subspaces of a separable, infinite dimensional Hilbert space.'

Mackey notes that, while it may seem a little *ad hoc* at first, the axiom is actually quite natural if one looks at all the choices that are left. His main point in favour of a *complex* Hilbert space is that this gives a 'natural correspondence between observables and one-parameter groups of symmetries', like Noether's theorem does in classical mechanics.

A further study of Mackey's axioms can be found in a paper of Mączyński [Mac72]. There, the author argues (without the help of the above axiom) that quantum mechanics needs to be build upon a Hilbert space over a *division ring*. Seen this way, a complex Hilbert space is then the simplest possibility after having ruled out a real Hilbert space with the help of the 'probabilistic interpretation' argument discussed before.

Anyway, we will in the following assume that quantum mechanics should be build upon a complex Hilbert space.

1.5.2 Vectors or Rays?

Our next objective is to answer the question whether pure states are *vectors* or *rays* in this complex Hilbert space $\mathcal{H}$. The obvious physical motivation to use *rays* is that two state *vectors* $|\tilde{\psi}\rangle, |\psi\rangle \in \mathcal{H}$ yield the same expectation value of an operator $\hat{A}$ if they are related by $|\tilde{\psi}\rangle = z|\psi\rangle$, where z is an invertible complex number $z \in \mathbb{C}^\times$:

It is common to write $\mathbb{F}^\times$ for the set of invertible elements in $\mathbb{F}$. Here we have $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$.

$$\langle \hat{A} \rangle_{\tilde{\psi}} = \frac{\langle \tilde{\psi} | \hat{A} | \tilde{\psi} \rangle}{\langle \tilde{\psi} | \tilde{\psi} \rangle} = \frac{|z|^2 \langle \psi | \hat{A} | \psi \rangle}{|z|^2 \langle \psi | \psi \rangle} = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle} = \langle \hat{A} \rangle_{\psi} \, .$$

The conclusion is that the space of ***physical states*** is a projective Hilbert space.

Definition 1.5.1 (Projective Hilbert space) Given a complex Hilbert space $\mathcal{H}$, the associated ***projective Hilbert space*** $P\mathcal{H}$ is the quotient space:

An analogous definition can be made, of course, for Hilbert spaces using other fields.

$$P\mathcal{H} := (\mathcal{H} \setminus \{0\})/\sim \, ,$$

where $\sim$ is the equivalence relation:

$$|\tilde{\psi}\rangle \sim |\psi\rangle \;:\Leftrightarrow\; \exists z \in \mathbb{C}^\times \text{ such that } |\tilde{\psi}\rangle = z|\psi\rangle \, .$$

The equivalence classes $[\psi] \in P\mathcal{H}$ are called ***rays***.

The story doesn't end here, however. Rays are incompatible with the *superposition principle*. The reason is that the superposition principle requires addition of states, but addition of rays isn't well-defined. To see this, consider two rays $[\psi_1]$, $[\psi_2]$ and try to define the superposition $[\psi] = [\psi_1] + [\psi_2] := [|\psi_1\rangle + |\psi_2\rangle]$ based upon the addition of representatives. We could just as well have represented the rays $[\psi_i]$ by some equivalent vectors $|\tilde{\psi}_i\rangle$, though, with $|\psi_i\rangle = z_i|\tilde{\psi}_i\rangle$ for some arbitrary numbers $z_i \in \mathbb{C}^\times$. These

representatives, however, yield the superposition $|\psi'\rangle = z_1|\psi_1\rangle + z_2|\psi_2\rangle$, which doesn't in general represent the same ray as the one generated by $|\psi\rangle$. Hence addition of rays is ill-defined.

A concrete situation where this problem surfaces is for interference effects. We can change an overall phase without influencing the outcome, but the *relative phases* of the individual wave functions in a superposition are important. Since rays have forgotten all phase information (including relative phases), we can't use them to properly describe interference effects. That said, it could make sense to see *vectors* as more fundamental states and interpret *rays* as something like 'measurement states'. The relation between these two kinds of states is then given by Wigner's theorem, which shows that symmetry transformations of the projective Hilbert space can be realised as unitary or anti-unitary operators on the Hilbert space of wave functions [KPRL08; Kel06].

1.5.3 Actual Calculations and Gel'fand Triples

One last point we want to discuss in this section is how states are used in actual calculations. The point is that we can't just construct some idealised quantum mechanics based on wishful thinking; we need to see quantum mechanics the way it is really used in practice. With this in mind, already Dirac mentions in his book that a Hilbert space is often too limiting for practical calculations [Dir58, section 10]:

> 'In our work up to the present it has been implied that our bra and ket vectors are of finite length and their scalar products are finite. We see now the need for relaxing this condition when we are dealing with eigenvectors of an observable whose eigenvalues form a range. If we did not relax it, the phenomenon of ranges of eigenvalues could not occur and our theory would be too weak for most practical problems.'

With 'ranges of eigenvalues' he means the *continuous spectrum* of an operator.

His conclusion is that 'The bra and ket vectors that we now use form a more general space than a Hilbert space'. Although he remarks later that 'It may be [...] that all realizable states correspond to ket vectors that can be normalized and that form a Hilbert space', he also mentions that 'such [generalised] eigenstates play a very useful role in the theory and one could not very well do without them' [Dir58, section 12].

While Dirac uses generalised eigenstates routinely in his formalism, he doesn't even waste a thought on an underlying mathematical framework. Generalised eigenstates are treated just as if they were part of the Hilbert space, ignorant of any differences that may be necessary. Consequently, Dieudonné comments: 'When one gets to the mathematical theories which are at the basis of quantum mechanics, one realizes that the attitude of certain physicists in the handling of these theories truly borders on the delirium.' (see [Gie00, chapter 1]), and von Neumann accuses him in the preface of his book [Neu68] that he introduces 'mathematical fictions' to make the formalism work.

Unfortunately, even today, most textbooks on quantum theory simply ignore the problem and assume the standpoint that the formalism works just fine for all practical purposes. Objections are often dismissed by the comment that wave packets can be used if necessary. Fact is, however, that operators with continuous spectra, and thus generalised eigenstates, appear quite frequently in practical calculations – think for instance of position and momentum of a free particle – and that we cannot obtain *reliable* results by some inconsistent use of mathematical machinery. Using wave packets, on the other hand, isn't very desirable either, as that would complicate calculations considerably.

For example, it is known today that a lot of the infinities in quantum field theory come from such wrong mathematical presumptions (see the work of Epstein and Glaser [EG73]).

Thankfully, the problem can be solved in quite an elegant way if we use so-called *rigged Hilbert spaces*, also known as *Gel'fand triples* [GV64; Böh78; Gro88]. The general

idea behind Gel'fand triples is to let operators act on *functionals* instead of wave functions and understand all equations in a distributional sense – to be evaluated on a set of test functions. From a physical point of view this evaluation on test functions is effectively the same as forming wave packets, yet without the inconvenience of using *concrete* wave packets during actual calculations. The precise definition of a Gel'fand triple is the following:

Definition 1.5.2 (Gel'fand triple) Let $\mathcal{H}$ be a Hilbert space, its topology induced by the inner product, and let $\mathcal{S} \subseteq \mathcal{H}$ be a dense subspace that carries a finer topology such that the inclusion $\iota : \mathcal{S} \hookrightarrow \mathcal{H}$ is continuous. Then, a ***Gel'fand triple*** (also called ***rigged***, or ***equipped Hilbert space***) is the triple:

$$\mathcal{S} \subseteq \mathcal{H} \subseteq \mathcal{S}^* \,, \tag{1.5}$$

where $\mathcal{S}^* \supseteq \mathcal{H}^*$ is the dual of $\mathcal{S}$, and the dualised Hilbert space $\mathcal{H}^*$ is identified with $\mathcal{H}$ itself canonically via $J : \mathcal{H} \to \mathcal{H}^* \,,\ f \mapsto \langle f | \cdot \rangle$.

According to Riesz' representation theorem, J is an isometric isomorphism.

The map $J : \mathcal{H} \to \mathcal{H}^*$ from above is the well-known duality between bras and kets, often written suggestively in Dirac notation as $J : |\psi\rangle \mapsto \langle\psi|$. Since J is an isomorphism, we know that to any ket there has to exist a bra and vice versa. In practical calculations, however, this is often not the case. More precisely, only the eigenstates of an operator $\hat{A}$ that belong to *discrete* eigenvalues are really elements of $\mathcal{H}$ and fulfil an eigenvalue equation of the type:

$$\hat{A}|\psi\rangle = \lambda|\psi\rangle \,, \quad \lambda \in \mathbb{C} \,,\ |\psi\rangle \in \mathcal{H} \setminus \{0\} \,. \tag{1.6}$$

We already mentioned that there are no eigenstates associated to *continuous* eigenvalues if we stay in the Hilbert space framework, but using an appropriate Gel'fand triple, we can define *generalised* eigenstates, i. e. functionals $\psi \in \mathcal{S}^*$ that satisfy a more general eigenvalue equation:

$$\hat{A}\psi = \lambda\psi \,,$$

which holds in the sense of distributions, that is:

$$(\hat{A}\psi)(f) = (\lambda\psi)(f) \quad \forall f \in \mathcal{S} \setminus \{0\} \,. \tag{1.7}$$

Here, f plays the role of a test function.

Let us finish this section with a short example of how the mathematical formalism based on Gel'fand triples works in practice.

Example 1.5.3 (Free particle) Consider a free particle moving in 1-dimensional space, so $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$. The position operator $\hat{q}$ acts on wave functions $\psi \in \mathcal{H}$ in the usual manner, $\hat{q}\psi(q) = q\psi(q)$, and the corresponding eigenvalue equation (where ψ_0 is the eigenfunction associated to q_0) is:

$$\hat{q}\psi_0(q) = q_0\psi_0(q) \;\Leftrightarrow\; (q - q_0)\psi_0(q) = 0 \,, \qquad \forall q \in \mathbb{R} \,.$$

Accordingly, $\psi_0(q) = 0$ for all $q \neq q_0$, and thus ψ_0 vanishes almost everywhere. Consequently, there are no solutions other than the null vector $\psi_0 = 0$. In the Hilbert space setting we conclude that $\hat{q}$ has no eigenstates.

The situation changes completely if we allow for distributions. Then we have the well-known solution:

$$\int q\delta_{q_0}(q)\, f(q)\, \mathrm{d}q = \int q_0\delta_{q_0}(q)\, f(q)\, \mathrm{d}q \quad \forall f \in \mathcal{S} \,, \tag{1.8}$$

with $\delta_{q_0}(q) := \delta(q - q_0)$. The space $\mathcal{S}$ of test functions in this example is the *Schwartz space* $\mathcal{S}(\mathbb{R})$ of differentiable functions that decrease faster than any inverse power of q. More precisely (see [Gro88, section 2.3.2.c] or any other book on functional analysis):

The elements of the dual space $\mathcal{S}(\mathbb{R})^*$ are known as *tempered distributions*.

$$\mathcal{S}(\mathbb{R}) := \{f \in C^\infty(\mathbb{R}) : \|f\|_{i,j} < \infty \ \forall i, j \in \mathbb{N}_0\} , \tag{1.9}$$

where $\|\cdot\|_{i,j}$ is the norm family:

$$\|f\|_{i,j} := \sup_{q\in\mathbb{R}} |q^i \partial_j f(q)| . \tag{1.10}$$

This is exactly the setting of a Gel'fand triple $\mathcal{S} \subset \mathcal{H} \subset \mathcal{S}^*$, and $\hat{q}$ now properly has a *generalised* eigenstate $\delta_{q_0} \in \mathcal{S}^*$ associated to each eigenvalue $q_0 \in \mathbb{R}$. To demonstrate that equation (1.8) is really an eigenvalue equation of the type like (1.7) it is possible to rewrite (1.8) in the form:

$$(\hat{q}\delta_{q_0})(f) = (q_0\delta_{q_0})(f) ,$$

where we use the notation $\delta_{q_0}(f) := f(q_0)$ and $(\hat{q}\delta_{q_0})(f) := \delta_{q_0}(qf) = q_0 f(q_0)$.

For our further discussion this is already enough to know about Gel'fand triples, so we want to stop here and refer the interested reader to the literature mentioned above. What should be kept in mind, however, is that Gel'fand triples provide the necessary foundation for Dirac's bra-ket formalism.

Concerning the quantization programme the result is that we can always start with representations on Hilbert spaces and extend them into Gel'fand triples later on when needed. In formal calculations involving states, however, we cannot in general assume states to be ordinary functions. Multiplication of states, for example, may be ill-defined if the states later turn out to be distributions.

1.6 Self-Adjoint and Unitary Operators

Stone's theorem 1.6.15 shows that there is an intimate link between self-adjoint operators and unitary transformations. A self-adjoint Hamiltonian is therefore required in order to preserve probabilities.

Self-adjoint operators are used to describe quantum-mechanical observables because their eigenvalues are real numbers and their eigenvectors orthogonal. There are, however, some peculiarities that, while often ignored during concrete calculations, have to be accounted for by quantization attempts. The first is that unbounded self-adjoint operators are defined only on a dense subset of the full Hilbert space. The other is the subtle yet quite important difference between just *symmetric* and *self-adjoint* operators. We base our discussion of these issues primarily on the paper by Gieres [Gieoo] and Großmann's book [Gro88]. Another source of information is the exhaustive work of Reed and Simon [RS.I; RS.II; RS.III; RS.IV].

When we later talk about irreducible representations of the canonical commutation relations, the peculiarities explain why self-adjoint representations are usually studied in terms of associated *unitary* representations. In the context of conventional quantum mechanics on $\mathbb{R}^n$, a lesser known variant of the Stone–von Neumann theorem then states that there are actually infinitely many *non-equivalent* representations. While this first seems like an obstacle, we will argue that it is exactly this feature that is the source of the scale-dependence (the fundamental constant $\hbar$) in quantum mechanics.

1.6.1 The Difference between Symmetric and Self-Adjoint Operators

Up front, we should remark that the mentioned peculiarities of self-adjoint operators reveal themselves only in the case of *infinite*-dimensional Hilbert spaces. There is,

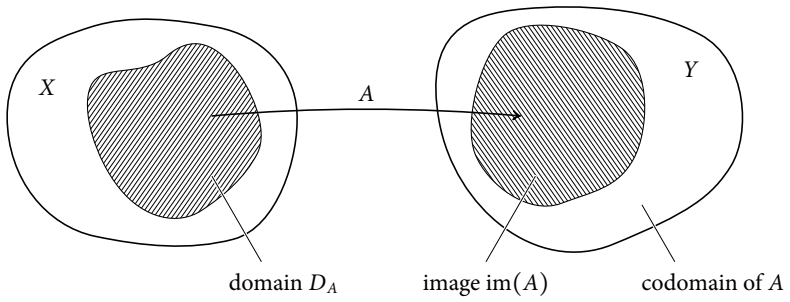


Figure 1.7 Concerning the definition 1.6.1 of an operator.

however, a simple argument which shows that Hilbert spaces of infinite dimension are necessary to describe quantum mechanics – at least in most cases. To this end, assume that the position and momentum operators $\hat{q}$ and $\hat{p}$ are defined for a *finite*-dimensional Hilbert space $\mathcal{H}$, say of dimension $\dim \mathcal{H} = n$. If we then calculate the trace of their commutator $[\hat{q}, \hat{p}] = i\hbar\mathbb{1}$, we obtain a contradiction because the left-hand side of the relation vanishes (the trace is cyclic) while the right-hand side yields $\mathrm{tr}(i\hbar\mathbb{1}_n) = i\hbar n \neq 0$. The problem disappears for infinite-dimensional Hilbert space because the trace of $\mathbb{1}$ is then no longer defined. While there are situations where only *traceless* operators appear as the results of commutators, we will hence in general need to consider Hilbert spaces of infinite dimension.

An example of traceless operators where a finite-dimensional Hilbert space is sufficient is given by Pauli's spin matrices.

Now, back to self-adjoint operators. Unfortunately, the terms *symmetric*, *Hermitian* and *self-adjoint* are often mixed up, so let us start with some proper definitions. The crucial point to observe is that an operator always comes together with a *domain* that may be smaller than the full Hilbert space. (The following definitions are loosely based on the ones given in [Gro88].)

Definition 1.6.1 (Operator) Consider two sets X and Y. An ***operator*** A consists of a ***domain*** $D_A \subseteq X$ and an ***operating prescription*** $A : D_A \to Y$ which maps elements in the domain D_A to elements in the ***target space*** (or ***codomain***) Y (see figure 1.7).

If X is a topological space, an operator A on D_A is said to be ***densely defined*** if its domain D_A is a dense subset of X (i. e. $\overline{D}_A = X$).

If X and Y are vector spaces, an operator A is called ***linear operator*** if the operating prescription $A : D_A \subseteq X \to Y$ is a linear map.

When we talk about equality of operators, the domains have to be taken into account as well. While it may seem a bit overly accurate at this point, the idea is nevertheless crucial to the understanding of self-adjoint operators.

Definition 1.6.2 (Equality of operators) Two operators A and B are ***equal***, $A = B$, if they have the same domain and the same operating prescription:

$$A = B \quad :\Leftrightarrow \quad D_A = D_B \text{ and } A\psi = B\psi \text{ for all } \psi \in D_A \,.$$

We say that A is a ***restriction*** of B, written $A \subset B$, if the domain of A is *smaller* than the domain of B but the operating prescriptions are compatible:

$$A \subset B \quad :\Leftrightarrow \quad D_A \subset D_B \text{ and } A\psi = B\psi \text{ for all } \psi \in D_A \,.$$

In turn, the operator B is in this case called an ***extension*** of A, written as $B \supset A$.

Next we define *symmetric* operators. Symmetric operators are the ones also called ***Hermitian*** by Reed and Simon [RS.I]. However, since the term Hermitian is used with ambiguous meaning throughout the literature we prefer the name *symmetric*.

Definition 1.6.3 (Symmetric operator) Let $\mathcal{H}$ be a Hilbert space and let A be a densely defined linear operator on $\mathcal{H}$. We say A is ***symmetric*** if:

$$\langle \varphi | A\psi \rangle = \langle A\varphi | \psi \rangle \quad \forall \psi, \varphi \in D_A \,. \tag{1.11}$$

ψ^* denotes the *complex conjugate* of ψ. Mathematicians will usually write $\overline{\psi}$ instead.

In the case of $\mathcal{H} = L^2(\mathbb{R}^n, \mathrm{d}^n x)$ this property is written as:

$$\int \varphi^*(\boldsymbol{x})\,(A\psi)(\boldsymbol{x})\,\mathrm{d}^n x = \int (A\varphi)^*(\boldsymbol{x})\psi(\boldsymbol{x})\,\mathrm{d}^n x \quad \forall \varphi, \psi \in D_A \,.$$

However, an operator that is merely symmetric is not guaranteed to have real eigenvalues. If we try to define the momentum operator $-i\hbar\partial_x$ on the interval $[0;1]$ it turns out to have *complex* eigenvalues (see also [Gie00, example 4]):

Example 1.6.4 (Particle restricted to a finite interval) We want to describe a particle that is constrained to the 1-dimensional interval $Q = [0;1]$. The appropriate Hilbert space is $L^2([0;1], \mathrm{d}x)$. To check whether the momentum operator $\hat{p} = -i\hbar\partial_x$ is symmetric we integrate by parts (for $\varphi, \psi \in \mathcal{H}$):

$$\int_0^1 \varphi^*(x)(\hat{p}\psi)(x)\,\mathrm{d}x = \int_0^1 (\hat{p}\varphi)^*(x)\psi(x)\,dx - i\hbar\big[\psi^*(x)\varphi(x)\big]_{x=0}^1 \,.$$

Großmann discusses some other options in his book but also shows that only *periodic* boundary conditions yield real eigenvalues [Gro88, section 17.2.1].

Accordingly, the operator $\hat{p}$ is symmetric if and only if the surface term vanishes. There are different types of boundary conditions to achieve this. Periodic boundary conditions don't make sense from a physical perspective because this would mean identifying both ends of the interval, yet this would describe movement on a *circle*. The most obvious choice is thus to require $\psi(0) = \psi(1) = 0$. Consequently, we restrict the domain of $\hat{p}$ to continuously differentiable functions in $\mathcal{H}$ which satisfy these boundary conditions. The eigenvalue equation then is:

$$\hat{p}\psi_p(x) = p\psi_p(x) \quad \forall x \in [0;1] \,,$$

which is solved by:

$$\psi_p(x) = c\, e^{(i/\hbar)px} \,, \quad c \in \mathbb{C} \setminus \{0\} \,,$$

for any value of $p \in \mathbb{C}$. None of these functions respects the boundary conditions, though, and so it first looks as if the spectrum of $\hat{p}$ were empty. Nevertheless, a more careful treatment reveals that the spectrum of $\hat{p}$ is actually the *entire complex plane* (see [Gro88, section 17.2.1]).

The spectrum in this example is purely *residual*, i. e. there are no eigenstates of $\hat{p}$ – not even generalised ones – but there are eigenstates of its *adjoint* $\hat{p}^\dagger$.

The example shows that symmetry is not enough to ensure real eigenvalues. The definition of *self-adjoint* operators fixes this problem [RS.I, section VIII.1]:

Definition 1.6.5 (Adjoint operator) Let A be a densely defined linear operator on some Hilbert space $\mathcal{H}$, with domain D_A. Its ***adjoint*** $A^\dagger$ is then uniquely determined by:

$$D_{A^\dagger} := \{\varphi \in \mathcal{H} : \exists \tilde{\varphi} \in \mathcal{H} \text{ such that } \langle \varphi | A\psi \rangle = \langle \tilde{\varphi} | \psi \rangle \ \forall \psi \in D_A\} \,, \tag{1.12}$$

and:

$$A^\dagger \varphi := \tilde{\varphi} \,, \tag{1.13}$$

where $\tilde{\varphi}$ depends on both A and φ. Accordingly, we have:

$$\langle A^\dagger \varphi | \psi \rangle = \langle \varphi | A\psi \rangle \quad \forall \psi \in D_A \,, \ \forall \varphi \in D_{A^\dagger} \,. \tag{1.14}$$

Remark 1.6.6 The operator A needs to be densely defined on $\mathcal{H}$ so that equation (1.12) uniquely determines its adjoint $A^\dagger$. The domain of $A^\dagger$, however, may not be dense (an example is given in [RS.I, section VIII.1, example 4]).

Further, it is absolutely crucial to note that the domain of the adjoint $A^\dagger$ is *explicitly given* by the definition. The only way to change it is to alter the domain of A. Observe that – according to equation (1.12) – $B \supset A$ implies $B^\dagger \subset A^\dagger$, while $B \subset A$ implies $B^\dagger \supset A^\dagger$. Hence, if we *enlarge* the domain of A, the domain of $A^\dagger$ gets *smaller*, and vice versa.

Definition 1.6.7 (Self-adjoint operator) A densely defined linear operator A on $\mathcal{H}$ is called ***self-adjoint*** if $A = A^\dagger$ in the sense of definition 1.6.2, that is:

$$D_A = D_{A^\dagger} \quad \text{and} \quad A\psi = A^\dagger \psi \quad \forall \psi \in D_A \,.$$

Both the domains *and* the operating prescriptions have be equal.

Note that neither the sum $A + B$ nor the product AB of self-adjoint operators has to be self-adjoint again.

The relation between symmetric and self-adjoint operators is that self-adjointness requires $A = A^\dagger$ while symmetric operators only satisfy $A \subseteq A^\dagger$. This raises the question whether a symmetric operator can be 'tuned' to produce a self-adjoint one. In the light of remark 1.6.6 we see this can only be achieved by *extending* A, because the domain of $A^\dagger$ is initially too large and needs to be made *smaller* to obtain equality.

Definition 1.6.8 (Essentially self-adjoint) A symmetric operator A is said to ***admit self-adjoint extensions*** if there exist (possibly many) self-adjoint extensions $B \supseteq A$.

A densely defined, symmetric operator A is ***essentially self-adjoint*** if the (unique) closure $\overline{A}$ (the smallest closed extension) of A is a self-adjoint operator.

Remark 1.6.9 (Deficiency indices) A common way to check if a given closed, symmetric linear operator A has self-adjoint extensions is to calculate its ***deficiency indices***:

$$n_\pm(A) = \dim\big(\mathfrak{r}_\pm(A)\big) \,, \tag{1.15}$$

where $\mathfrak{r}_\pm(A)$ are the so-called ***deficiency subspaces***:

$$\mathfrak{r}_\pm(A) := \ker(A^\dagger \mp i) = \mathrm{im}(A \pm i)^\perp \,. \tag{1.16}$$

An operator A is called ***closed*** if its ***operator graph*** $\Gamma(A)$, the set $\Gamma(A) := \{(\psi, A\psi) : \psi \in D_A\}$, is a closed subset. Any symmetric operator can be assumed closed without loss of generality.

The operator A admits self-adjoint extensions if and only if the deficiency indices are equal, $n_+ = n_-$, and A is self-adjoint if and only if $n_+ = n_- = 0$.

The idea underlying the method of deficiency indices is quite elegant. It turns out that if any complex number $z \in \mathbb{C} \setminus \mathbb{R}$ is an eigenvalue of a closed, symmetric operator A, then all points in the *whole closed (upper or lower) half-plane* $H_\pm$ containing z are eigenvalues as well. It is hence sufficient to check whether $z = \pm i$ is an eigenvalue of A, which is accomplished by looking at the dimension of the deficiency subspaces (1.16). In particular, if both deficiency indices vanish, the operator A admits only real eigenvalues and therefore is self-adjoint.

A thorough treatment of the method of deficiency indices can be found, for example, in [RS.II, chapter X] or [Gro88, section 16.5]. The method has been developed originally by von Neumann [Neu30].

The other crucial observation is that any symmetric extension of A changes both deficiency indices by the same amount. If the deficiency indices of A differ from each other, there can hence exist no self-adjoint extensions. On the other hand, there is an explicit formula for the self-adjoint extensions if n_+ and n_- are equal [RS.II, theorem X.2]. This finally yields the second statement according to which the operator A has self-adjoint extensions if and only if the deficiency indices are equal.

In some cases there are easier criteria for self-adjointness, like in the following:

Lemma 1.6.10 Let A be a symmetric linear operator that is defined on the whole Hilbert space $\mathcal{H}$, that is, $D_A = \mathcal{H}$. Then A is self-adjoint.

Proof. Since the adjoint of a symmetric operator fulfils $D_{A^\dagger} \supseteq D_A$ but also $D_{A^\dagger} \subseteq \mathcal{H}$, we immediately obtain $D_{A^\dagger} = D_A = \mathcal{H}$. Moreover, $A\psi = A^\dagger\psi$ for all $\psi \in D_A = \mathcal{H}$, and hence A is self-adjoint. ∎

Although it thus looks as if all the obscure problems with self-adjointness vanish if we simply use symmetric operators that are defined on the whole Hilbert space, there is a catch. Everywhere defined, symmetric linear operators have another property: they are automatically *bounded* as a result of the Hellinger–Toeplitz theorem.

Theorem 1.6.11 (Hellinger, Toeplitz) If A is an everywhere defined, symmetric linear operator on a Hilbert space $\mathcal{H}$, then A is bounded.

Proof. See [RS.I, section III.5] or [Gro88, section 10.1.2]. ∎

A *Banach space* is a *complete normed* vector space. Any Hilbert spaces is also a Banach space, with $\|\psi\|^2 := \langle\psi|\psi\rangle$.

Definition 1.6.12 (Bounded operator) A linear operator A on some Banach space $\mathcal{B}$ is ***bounded*** if there exists a constant c such that:

$$\|A\psi\| \leq c\,\|\psi\| \qquad \forall\psi \in D_A\ .$$

Under these circumstances, the ***norm*** of A is defined as:

$$\|A\| := \sup_{\psi\in D_A} \frac{\|A\psi\|}{\|\psi\|}\ . \tag{1.17}$$

Bounded operators have a number of nice properties. To give an example, a linear operator is bounded if and only if it is continuous, and it can always be defined on the whole Hilbert space. Nevertheless, it turns out that most physically interesting operators *cannot be bounded* because only *unbounded* operators may have a nontrivial uncertainty relation (see section 1.5 and [Gro88, section 5.5]). Read in reverse, the Hellinger–Toeplitz theorem then states that the domain of unbounded, self-adjoint operators *cannot* be the whole Hilbert space – the domain must be a proper subset $D_A \subset \mathcal{H}$. This leads to problems such as the following (see [Gie00, example 5]):

Example 1.6.13 (Angle and angular momentum variables) Consider polar coordinates in the plane, with angle φ and angular momentum ℓ, then $\mathcal{H} = L^2([0;2\pi], \mathrm{d}\varphi)$ is our Hilbert space. The associated quantum operators $\hat\varphi$ and $\hat\ell$ act on wave functions via:

$$(\hat\varphi\psi)(\varphi) = \varphi\psi(\varphi)\ , \quad (\hat\ell\psi)(\varphi) = -i\hbar\partial_\varphi\psi(\varphi)\ .$$

If we choose the domains as:

$$D_{\hat\varphi} = \mathcal{H}\ , \quad D_{\hat\ell} = \left\{\psi \in \mathcal{H} : \psi' \in \mathcal{H}\ ,\ \psi(0) = \psi(2\pi)\right\}\ ,$$

both are self-adjoint operators (see [Gro88, section 16.1.3]), and their commutator is $[\hat\varphi, \hat\ell\,] = i\hbar\mathbb{1}$. Furthermore, eigenstates of $\hat\ell$ are given by:

$$\hat\ell|\psi_m\rangle = m\hbar|\psi_m\rangle\ , \quad \psi_m(\varphi) = (2\pi)^{-1/2}e^{im\varphi}\ , \quad (m \in \mathbb{Z})$$

and they are orthonormal with respect to the standard scalar product:

$$\langle \psi_1 | \psi_2 \rangle := \int_0^{2\pi} \psi_1^*(\varphi)\psi_2(\varphi)\,\mathrm{d}\varphi \,.$$

Now, the expectation value of the commutator in eigenstates of $\hat{\ell}$ is:

$$\langle \psi_m | \,[\hat{\varphi}, \hat{\ell}\,]\, | \psi_m \rangle = \langle \psi_m | i\hbar \mathbb{1} | \psi_m \rangle = i\hbar \langle \psi_m | \psi_m \rangle = i\hbar \,. \tag{1.18}$$

Nevertheless, using that $\hat{\varphi}$ and $\hat{\ell}$ are self-adjoint, we can calculate this in a different way:

$$\begin{aligned} \langle \psi_m | \,[\hat{\varphi}, \hat{\ell}\,]\, | \psi_m \rangle &= \langle \psi_m | (\hat{\varphi}\hat{\ell} - \hat{\ell}\hat{\varphi}) | \psi_m \rangle = \langle \hat{\varphi}^\dagger \psi_m | \hat{\ell}\psi_m \rangle - \langle \hat{\ell}^\dagger \psi_m | \hat{\varphi}\psi_m \rangle \\ &= \langle \hat{\varphi}\psi_m | \hat{\ell}\psi_m \rangle - \langle \hat{\ell}\psi_m | \hat{\varphi}\psi_m \rangle = (m\hbar - m\hbar) \langle \psi_m | \hat{\varphi}\psi_m \rangle = 0 \,. \end{aligned} \tag{1.19}$$

This second result obviously contradicts (1.18), however. What went wrong?

The solution is to look closely at the domains of all the operators involved in this calculation. The important observation is that *sums* and *products* of operators have their *own* specific domains, different from the domains of the individual operators. The rules to calculate them are:

$$D(A + B) = D(A) \cap D(B) \,, \tag{1.20}$$

$$D(AB) = \{\psi \in D(B) : B\psi \in D(A)\} \,. \tag{1.21}$$

Applied to the example we obtain:

$$\begin{aligned} D(\hat{\varphi}\hat{\ell}) &= \{\psi \in D_{\hat{\ell}} : \hat{\ell}\psi \in D_{\hat{\varphi}}\} = \{\psi \in D_{\hat{\ell}} : \hat{\ell}\psi \in \mathcal{H}\} = D_{\hat{\ell}} \,, \\ D(\hat{\ell}\hat{\varphi}) &= \{\psi \in D_{\hat{\varphi}} : \hat{\varphi}\psi \in D_{\hat{\ell}}\} = \{\psi \in \mathcal{H} : \hat{\varphi}\psi \in D_{\hat{\ell}}\} \,, \\ D([\hat{\varphi}, \hat{\ell}\,]) &= D(\hat{\varphi}\hat{\ell}) \cap D(\hat{\ell}\hat{\varphi}) = \{\psi \in D_{\hat{\ell}} : \hat{\varphi}\psi \in D_{\hat{\ell}}\} \,. \end{aligned}$$

We see, what went wrong in calculations (1.18) and (1.19) is that the eigenstates ψ_m are not in the domain of the commutator since:

$$\hat{\varphi}\psi_m(\varphi) = \varphi\,(2\pi)^{-1/2} e^{im\varphi} \;\Rightarrow\; \hat{\varphi}\psi_m(0) = 0 \neq (2\pi)^{1/2} = \hat{\varphi}\psi_m(2\pi) \,,$$

yet $D_{\hat{\ell}}$ contains only *periodic* functions, and hence $\psi_m \notin D([\hat{\varphi}, \hat{\ell}\,])$. The contradicting results are thus no longer so entirely surprising, as we formally applied operators to wave functions outside their domain.

There are two points to remember about this:

- If we want reliable results, the domain of operators has to be taken into account. As such, the subset on which an identity holds – like the commutation relations – may be much smaller than the whole Hilbert space and it may not contain important classes of functions. In particular this is the case for the *uncertainty principle*, which if not trivial can only be true on a subset of functions, as it depends on the commutator of two necessarily unbounded operators.
- Moreover, the given example shows that domain problems are not restricted to some obscure pathological cases but may occur in quite ordinary and physically relevant situations.

There exist concrete examples of wave functions that violate the uncertainty principle (see e. g. [Gieoo, example 6]).

1.6.2 Stone's Theorem and One-Parameter Unitary Groups

Concerning quantization we see that problems with the domains of self-adjoint operators cannot simply be dismissed. On the other hand, it is nearly impossible to study representations if we always had to take care of the domains, so it would be helpful if there were an alternative. Fortunately there is. The resolve comes in the form of *Stone's theorem* that establishes a correspondence between self-adjoint operators and strongly continuous one-parameter groups of unitary transformations.

Definition 1.6.14 (One-parameter unitary groups) An operator-valued mapping

$$U : \mathbb{R} \to \mathrm{Op}(\mathcal{H}) , \quad t \mapsto U_t$$

is called ***one-parameter unitary group*** if:

- U_t is a unitary operator (i. e. $U_t^\dagger U_t = U_t U_t^\dagger = \mathbb{1}$) for all $t \in \mathbb{R}$,
- $U_{t+s} = U_t U_s$ for all $t, s \in \mathbb{R}$.

A theorem of von Neumann [RS.I, theorem VIII.9] implies that every weakly continuous *unitary*(!) one-parameter group is actually strongly continuous.

In addition, U is said to be ***strongly continuous*** if:

- $U_t|\psi\rangle \to U_{t_0}|\psi\rangle$ for $t \to t_0$ and any $|\psi\rangle \in \mathcal{H}$.

Theorem 1.6.15 (Stone) There is a correspondence between self-adjoint operators and strongly continuous one-parameter unitary groups. More precisely:

(i) Let A be a self-adjoint operator on a Hilbert space $\mathcal{H}$, then $U(t) := e^{itA}$ defines a strongly continuous one-parameter unitary group.
(ii) In reverse, if U is a strongly continuous one-parameter group there always exists a self-adjoint operator A on $\mathcal{H}$ such that $e^{itA} = U(t)$.

Note that the usual power series cannot be used to calculate the exponential for unbounded self-adjoint operators since the series might not converge in this case [RS.I, section VIII.4].

Proof. The original proof can be found in [Sto30]. Alternatively, you can refer to [RS.I], where part (i) is implied by theorem VIII.7 and part (ii) is theorem VIII.8. ∎

Unlike self-adjoint operators, unitary operators are always bounded and can hence be assumed as being *everywhere defined* on the whole Hilbert space $\mathcal{H}$, without any loss of generality. Instead of studying self-adjoint representations directly it is hence common to examine *unitary representations* and introduce self-adjoint generators afterwards. In more geometric terms, this amounts to studying unitary representations of a *Lie group* G to obtain self-adjoint representations of a *Lie algebra* $\mathfrak{L}G$.

1.6.3 Weyl Form of the Commutation Relations and the Heisenberg Group

In order to study the unitary representations we will have to rewrite the canonical commutation relations in terms of the associated unitary operators. The result is usually called ***multiplicative***, or ***integrated Weyl form*** of the relations. In the case of conventional quantum mechanics (for simplicity we restrict ourselves to one dimension) we obtain:

Lemma 1.6.16 (Weyl form of the CCR) The usual canonical commutation relations of conventional quantum mechanics on $\mathbb{R}$:

$$[\hat{q}, \hat{p}] = i\hbar\mathbb{1} , \quad [\hat{q}, \hat{q}] = 0 , \quad [\hat{p}, \hat{p}] = 0 ,$$

have the integrated Weyl form:

$$U(a)V(b) = V(b)U(a)e^{i\hbar ab}\,,$$
$$U(a_1 + a_2) = U(a_1)U(a_2)\,, \quad V(b_1 + b_2) = V(b_1)V(b_2)\,, \tag{1.22}$$

with $U(a) := e^{-ia\hat{p}}$ and $V(b) := e^{-ib\hat{q}}$.

Proof. The equivalence is most easily shown if we use that the exponential map satisfies the important relation $\mathrm{Ad} \circ \exp = \exp \circ \mathrm{ad}$ (for a proof see section 3.4, equation (3.16)), where 'Ad' and 'ad' are the adjoint actions of Lie group and Lie algebra, respectively, given in definitions 3.3.8 and 3.3.9. Moreover, there is a close relation between the adjoint action ad and the commutator, namely $[A, B] = \mathrm{ad}(A)(B)$. Thus we have:

$$\begin{aligned} U(a)\hat{q}U(a)^{-1} &= \mathrm{Ad}\big(U(a)\big)\hat{q} = \mathrm{Ad}\big(\exp(-ia\hat{p})\big)\hat{q} = \exp\big(\mathrm{ad}(-ia\hat{p})\big)\hat{q} \\ &= \big(\mathbb{1} + \mathrm{ad}(-ia\hat{p}) + \tfrac{1}{2}\,\mathrm{ad}(-ia\hat{p}) \circ \mathrm{ad}(-ia\hat{p}) + \cdots\big)\hat{q} \\ &= \hat{q} + [-ia\hat{p}, \hat{q}] + \tfrac{1}{2}\,[-ia\hat{p}, [-ia\hat{p}, \hat{q}]] + \cdots \\ &= \hat{q} - \hbar a\,, \end{aligned}$$

where we first expanded the exponential as a Taylor series, and then the last equality is due to the fact that all higher commutators of $\hat{q}$ and $\hat{p}$ vanish, since $\mathbb{1}$ commutes with both $\hat{q}$ and $\hat{p}$, and hence $[\hat{p}, [\hat{p}, \hat{q}]] \propto [\hat{p}, \mathbb{1}] = 0$ and so on. By an analogous calculation we obtain $V(b)\hat{p}V(b)^{-1} = \hat{p} + \hbar b$.

Finally, for any analytical function f there is the relation:

$$U(a)f(\hat{q})U(a)^{-1} = f(\hat{q} - \hbar a)\,,$$

and since the exponential is analytical we obtain:

$$U(a)V(b)U(a)^{-1} = e^{-ib(\hat{q}-\hbar a)} = V(b)e^{i\hbar ab}\,,$$

which is the first of the relations in equation (1.22). The remaining ones follow directly from the definitions of $U(a)$ and $V(b)$. ∎

The group behind the commutation relations (1.22) is the *Heisenberg group* $H(1)$.

Definition 1.6.17 (Heisenberg group) The ***Heisenberg***, or ***Weyl–Heisenberg group*** is given as:

$$H(n) = (\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}, \cdot)\,,$$

with the group law:

$$(\boldsymbol{u}_1, \boldsymbol{v}_1, s_1) \cdot (\boldsymbol{u}_2, \boldsymbol{v}_2, s_2) := \big(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2, s_1 + s_2 + \tfrac{1}{2}(\boldsymbol{v}_1 \cdot \boldsymbol{u}_2 - \boldsymbol{v}_2 \cdot \boldsymbol{u}_1)\big)\,. \tag{1.23}$$

Sometimes the name ***polarised Heisenberg group*** is used because $H(n)$ results from a more general version of the Heisenberg group, which acts on a symplectic manifold, after choosing a Darboux chart (see chapter 4, in particular theorem 4.1.8).

The representation theory of the Heisenberg group $H(n)$ is dictated by the famous *Stone–von Neumann theorem*. It can be formulated in the following way:

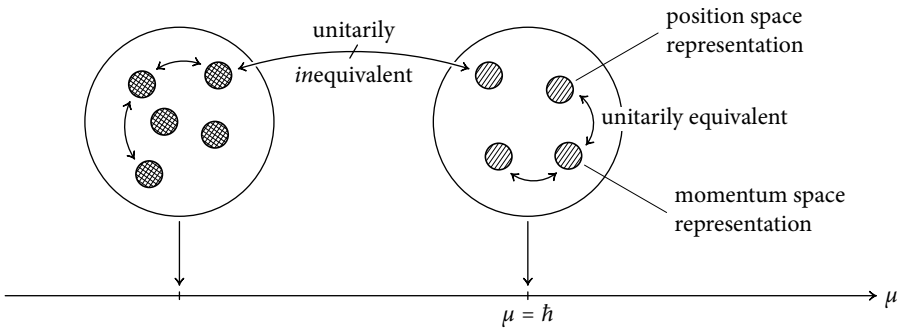


Figure 1.8 According to the Stone–von Neumann theorem representations of the Heisenberg group that belong to the *same* value of μ are unitarily equivalent. Representations belonging to *different* values of μ are unitarily *in*equivalent.

Note that the constant $\hbar$ from equation (1.22) has now been replaced by a variable μ which parametrises representations of the central subgroup of the Heisenberg group.

Theorem 1.6.18 (Stone, von Neumann) Consider the Weyl commutation relations of the Heisenberg group:

$$U(a_1 + a_2) = U(a_1)U(a_2)\,, \quad V(b_1 + b_2) = V(b_1)V(b_2)\,,$$
$$U(a)V(b) = V(b)U(a)e^{i\mu ab}\,, \quad \mu \in \mathbb{R}\,,$$

where $e^{i\mu s}\mathbb{1}$ is a unitary representation of the central subgroup $Z(H(n)) \cong (\mathbb{R}, +)$. Then all non-trivial, strongly continuous, unitary, irreducible representations partition into equivalence classes such that:

The additional statement about inequivalent representations is often omitted, yet we will see below that it is at least equally important from a physical point of view.

(i) representations that belong to the *same* value of μ are *unitarily equivalent*,

(ii) representations that belong to *different* values of μ are unitarily *in*equivalent.

Proof. See [Sto30] and [Neu31], or [Ros04]. ∎

Corollary 1.6.19 If we assume the parameter μ to be fixed beforehand, say to the value $\mu = \hbar$, the Schrödinger representation in position space:

$$\hat{x} = x\,, \quad \hat{p} = -i\hbar\partial_x\,, \quad \mathcal{H} = L^2(\mathbb{R}^n, \mathrm{d}^n x)\,,$$

is *unique up to unitary equivalence*. In other words, after fixing the value of μ all unitary, irreducible representation of the Heisenberg group are *indistinguishable from a physical perspective* because unitary equivalences preserve the outcomes of measurements.

1.6.4 Physical Meaning of the Stone–von Neumann Theorem

Historically, the Stone–von Neumann theorem showed the equivalence between Heisenberg's matrix mechanics and Schrödinger's wave function formalism. Apart from that, the corollary guarantees that calculations in position space yield the same physical results as calculations in momentum space or any other representation we may choose. Used this way, the corollary to the Stone–von Neumann theorem justifies the concept of an *abstract* Hilbert space. In contrast, if an analogous statement doesn't hold, calculations using different representations may well produce different physical results, so that an abstract Hilbert space doesn't make sense.

Still, there seems to be a problem if we don't fix the parameter μ in advance. In fact, theorem 1.6.18 then states that *different* values of μ give rise to unitarily *inequivalent* representations, resulting in different physical predictions. So, unless we agree to fix some value of μ we obtain *infinitely many different quantizations*! Is this acceptable?

First, I would say that it doesn't pose a problem because quantization is in its nature not a properly defined mathematical mapping (remember, first quantization is a mystery, not a functor). In fact, it is better seen as a 'recipe' to construct candidate theories that could be useful to describe quantum mechanical systems. While it might be possible to introduce additional requirements to make quantization mathematically well-defined, Isham points out that such attempts should be considered with great care [Ish83, p.1155]:

We already discussed this 'recipe-nature' of quantization at the end of section 1.4.

> 'Common sense must be used here to avoid embarking on an overaxiomatized, and hence misguided, piece of theoretical physics. We are still trying to feel our way towards an understanding of the quantization of intrinsically nonlinear systems and should not be trapped into axiomatizing theoretical ideas out of existence.'

Even more, though, I want to argue that the inequivalence of representations for different values of μ is actually quite important from a physical point of view. It is the feature responsible for the **fundamental scale** $\hbar$ of quantum theory. Unlike classical mechanics, quantum theory behaves differently for *microscopic* and *macroscopic* systems. As such, a physical system exhibits genuinely microscopic quantum behaviour if it has an action that is less or approximately equal to $\hbar$, whereas classical behaviour emerges if the effective action is much bigger than $\hbar$.

That being said, if representations belonging to different values of μ were unitarily equivalent, the value of μ would be irrelevant and could be chosen at will, because different values would produce physically indistinguishable predictions (this is what happens for the parameter α in section 6.1). Only inequivalent representations make different physical predictions:

> It is the unitary *inequivalence* between representations that says that μ is a quantity with physical impact; a quantity that can be determined by physical experiment and is able to fix a fundamental scale.

Finally, it is important to remember that the theorem of Stone and von Neumann is a statement *specifically* about representations of the Heisenberg group, which we will see is deeply connected to conventional quantum mechanics. In general, however, there will be other groups taking this role and then we have *no analogous statement* about their representation theory. There are examples where several distinct equivalence classes of representations appear, each describing a *genuinely different* physical setting, not just a scale dependence as above. It is then again a matter of common sense and physical insight to select the ones that describe reasonable quantum systems and find sensible interpretations.

Still, there are powerful methods to study representations of the occurring groups, most prominently Mackey's theory of induced representations (see e. g. [Mac49] and [Mac76]).

1.7 What Is Quantization?

Following this lengthy investigation of the fundamentals we are now in the position to formulate a pragmatic foundation for quantization. To this end, let us first recall what we initially started with: the Dirac quantization map.

In definition 1.1.2 the Dirac quantization map was defined as:

$$\mathcal{Q} : C^\infty(M, \mathbb{R}) \to \mathrm{Op}(\mathcal{H}) \,,$$

mapping functions over the classical phase space M to operators on some associated separable complex Hilbert space $\mathcal{H}$, such that $\mathcal{Q}$ has the following properties:

(q1) the map $\mathcal{Q}$ is $\mathbb{R}$-*linear*,

(q2) the operators $\mathcal{Q}(f)$ for $f \in C^\infty(M, \mathbb{R})$ are *essentially self-adjoint*,

(q3) $\mathcal{Q}$ maps *Poisson brackets* to *commutators* such that:

$$[\mathcal{Q}(f), \mathcal{Q}(g)] = i\hbar\mathcal{Q}(\{f, g\}) ,$$

(q4) the map $\mathcal{Q}$ is *irreducible* in the sense that any complete set of classical observables $\{f_1, \ldots, f_k\}$ is mapped to a complete set of operators $\{\mathcal{Q}(f_1), \ldots, \mathcal{Q}(f_k)\}$,

(q5) the constant function 1_M on M maps to the identity on $\mathcal{H}$, i. e. $\mathcal{Q}(1_M) = \mathbb{1}_{\mathcal{H}}$.

The theorem of Groenewold and Van Hove showed that such a 'full' quantization is impossible as the requirements are inconsistent. Any feasible quantization approach hence needs to discard, or at least modify, one or the other of the conditions. To make an informed choice, our goal in this chapter was to reveal which of the requirements have some actual *physical meaning* and which ones are assumed just for mathematical convenience.

First, we should remark that condition (q5) is meant to rule out some trivial solutions, yet is not really independent of the other requirements. It can be deduced from (q3), for example, and is hence only necessary if we restrict some of the conditions, say like in Deformation quantization.

In section 1.4 we argued that it makes sense to **restrict the domain** of $\mathcal{Q}$. Although there are similarities between classical and quantum mechanics (think of the classical limit) and hence some basic variables of the theory should be quantizable, there are also some obvious differences, so the correspondence cannot be assumed to hold for any arbitrarily crazy variable. Moreover, we explained that not all classical phase spaces must yield reasonable quantum theories, and that the physical interpretations of the classical and the quantized system may be different since the abstract mathematical models might be realised in different ways.

Further, with respect to quantum states it is indeed possible to initially describe quantum mechanics based upon some **complex separable Hilbert space**. While for most practical calculations we need Gel'fand triples over this Hilbert space in order to include eigenstates associated to continuous eigenvalues (section 1.5), this is an additional step which can be implemented later. Only when we need specific properties of states we have to keep in mind that states may in general be distributions rather than functions.

That said, there is another problem hidden in the term **associated** Hilbert space. For conventional quantum mechanics on $\mathbb{R}^n$ the Stone–von Neumann theorem tells us that the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^n, \mathrm{d}^n x)$ of square-integrable functions over configuration space is essentially unique – that is, up to some unitary equivalence that doesn't influence the physical results. In other situations, however, there is no such statement and it becomes an open question what 'associated' means. In general we will obtain for a single classical system a whole family of quantized systems that may have genuinely different physical interpretations.

In addition, Isham remarks in [Ish83] that when we are dealing with topologically nontrivial spaces we should also consider 'twisted' representations, meaning representations on *sections of vector bundles*, not just representations on functions. To handle

these, the Serre–Swan theorem might be used, which establishes a nice correspondence between sections of vector bundles over compact manifolds and functions in finitely generated projective modules [Swa62].

Another controversial question is the **linearity** of $\mathcal{Q}$. It can be broken down into two individual properties: *addition* and *multiplication by real numbers*. Physically, a change of units (say, from metres to centimetres) corresponds to a multiplication by a real number. Since such a change of units should have the same effect in classical as in quantum mechanics it seems reasonable to require the quantization map to respect this operation. There are, however, working quantization approaches with a *nonlinear* prescription (see the remark about Blattner–Kostant–Sternberg (BKS) kernels in [TAE05, section 1.2]). We thus see that a linear quantization map can't be strictly necessary.

Despite, there is a good motivation to require linearity when regarded in the context of condition (q3) – the mapping of Poisson brackets to commutators. Both conditions taken together ensure that $(i\hbar)^{-1}\mathcal{Q}$ is a Lie algebra morphism, so that the **Lie algebra structure** is preserved. Typically, it is the Lie algebra structure that is characteristic to physical systems and so should remain unaffected during the transition from classical to quantum theory. The reason is on one side the deep connection between *observables* and *symmetry transformations* according to Noether's theorem. On the other side, there is a relation between symmetries and the **geometry** of phase space; a relation that will become more visible in the course of chapter 3. The rules (q1) and (q3) can thus be interpreted such that they ensure quantization somehow respects the geometry.

This relation between geometry, symmetry transformations and observables provides one of the cornerstones of Canonical Group Quantization.

The requirement for essentially **self-adjoint** operators – property (q2) – ensures real eigenvalues, and thus real numbers as the results of measurements. That being said, we explained in section 1.6 that there is a difference between merely *symmetric* and *self-adjoint* operators, based on the domains. As a result, unbounded self-adjoint operators are in practice quite difficult to treat correctly because their domains have to be taken into account at all times. This is an additional point that a serious quantization approach has to respect if it means to produce reliable results.

Concerning the **irreducibility condition** (q4) the idea is, roughly speaking, that it ensures that the number of observables doesn't change. This is best illustrated by means of an example: For classical mechanics on $Q = \mathbb{R}^3$ the set $\{q^i, p_j\}$ consisting of position and momentum is complete (i. e. there are no observables other than the identity that Poisson-commute with all q^i and p_j) and it is mapped to the set $\{\hat{q}^i, \hat{p}_j\}$ under the quantization map. In spin-less quantum mechanics all operators are expressible as functions of $\hat{q}^i$ and $\hat{p}_j$ due to Schur's lemma, so this set is complete and hence the quantization map irreducible. If we include spin, however, we know that the spin operators commute

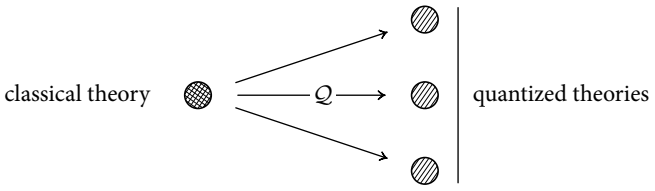


Figure 1.9 For conventional quantization the Stone–von Neumann theorem guarantees that the result is essentially unique. In general, however, a single classical theory may yield several *genuinely different* quantized theories.

with all the $\hat{q}^i$ and $\hat{p}_j$ but are *not* proportional to the identity. In this case the set $\{\hat{q}^i, \hat{p}_j\}$ is not complete, and hence the quantization map *not* irreducible.

There are two conclusions: The irreducibility ensures that no additional, unmotivated observables arise from the quantization process, nor do we lose any. In principle this is just what we need because the same basic set of variables should play a role in the classical as in the quantized theory (remember the classical limit argument from section 1.4). Nevertheless, the example of spin shows that it may sometimes be reasonable to relax this irreducibility condition a little bit to allow for quantum observables with no classical analogue.

Finally, let me summarise how a pragmatic approach to the quantization problem could look like. First, we explained that quantization isn't an actual physical process but rather a *recipe* how to construct quantum theories based on some classical ingredients. Because of this 'recipe-nature', a quantization method doesn't have to produce a *unique* quantized system. Until the recipe is better understood it makes more sense to exclude some of the results afterwards based on a case-by-case analysis rather than to dismiss theoretical possibilities beforehand by introducing physically ill-justified axioms.

Instead of a single quantization map we will hence assume that there exists in general a *collection* of different possible quantizations. For each individual quantization map $\mathcal{Q}$ in this collection the following assumptions seem physically reasonable:

- Quantization is restricted to a (possibly small) subset of *quantizable observables*. This set $\mathrm{Obs}(M) \subseteq C^\infty(M, \mathbb{R})$ needs to be complete in the sense of definition 1.1.1 in order to 'feel out' the whole classical phase space M.
- Functions in $\mathrm{Obs}(M)$ are mapped to essentially self-adjoint operators on a complex separable Hilbert space $\mathcal{H}$. With regard to the operator domains, most difficulties can be circumvented if we make use of strongly continuous one-parameter unitary groups instead of studying self-adjoint operators directly.
- The quantization map should preserve the Lie algebra structure in the sense that $\mathcal{Q}$ is linear and maps Poisson brackets to commutators according to the rule (q3).
- The quantization map is assumed to be irreducible in the sense of (q4). However, this condition may be relaxed if necessary.

1.8 A Short Outline of Canonical Group Quantization

The aim of this section is to *convey the idea* of Canonical Group Quantization. We will fill in the technical details in chapter 5.

The main difficulties of a quantization attempt as outlined above are to determine a reasonable subset of quantizable observables and to find suitable Hilbert spaces. Canonical Group Quantization, proposed by Christopher Isham in [Ish83], tries to solve these problems in a way that is particularly transparent from a physical point of view. The idea underlying Isham's approach is that the canonical commutation relations arise from a group $\mathcal{G}$ which acts on the classical phase space M in some natural way.

The first step is to understand the canonical commutation relations as a self-adjoint representation of a Lie algebra $\mathrm{Obs}(M)$, the Lie algebra that is generated by what we will call a set of ***fundamental observables*** (a complete set of quantizable observables, big enough to uniquely determine the quantization map; see also figure 1.10). Still, it is more convenient to replace this *concrete* Lie algebra $\mathrm{Obs}(M)$ consisting of functions on classical phase space M by an *abstract* Lie algebra, the *Canonical Lie algebra* $\mathfrak{LC}$:

For quantization on $\mathbb{R}^n$ such a set of *fundamental observables* is given by $\{q^i, p_i, 1_M\}$, for instance. The Lie algebra $\mathrm{Obs}(M)$ is then the smallest Lie algebra containing these functions.

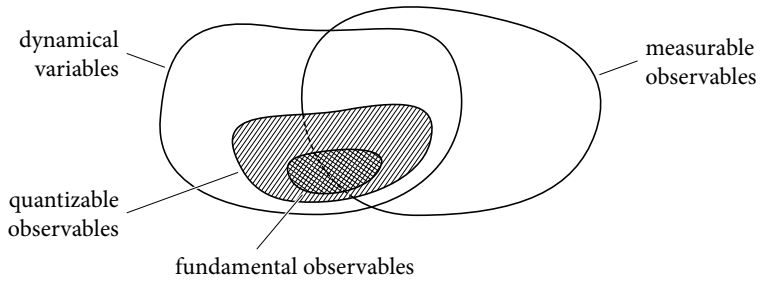
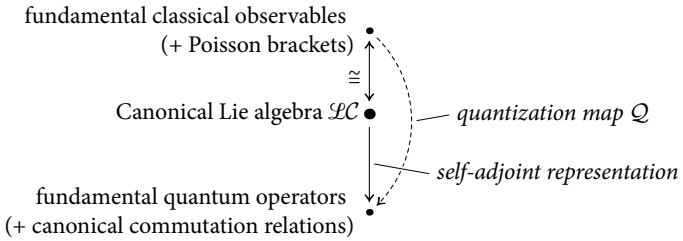
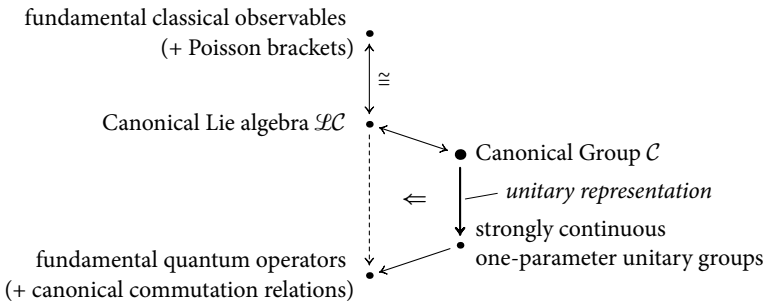


Figure 1.10 The fundamental observables form a subset of the set of quantizable observables, big enough to uniquely determine a quantization map.



Definition 1.8.1 (Canonical Lie algebra) The ***Canonical Lie algebra*** $\mathcal{LC}$ is an abstract Lie algebra that is isomorphic to the Lie algebra $\mathrm{Obs}(M)$ that is generated by a complete set of fundamental observables.

Furthermore, we explained earlier that it is easier to study self-adjoint representations in terms of strongly continuous unitary representations of a Lie group associated to $\mathcal{LC}$. This group is the ***Canonical Group*** $\mathcal{C}$, and we replace the former construction by the following:



Once we have a Canonical Group we can study its unitary representations on all possible Hilbert spaces $\mathcal{H}$. Furthermore, a quantization map $\mathcal{Q}$ can be defined by sending fundamental classical observables to the self-adjoint generators of strongly continuous one-parameter unitary groups.

For conventional quantum mechanics over the configuration space $Q = \mathbb{R}^n$ this method reproduces the well-known results (see section 5.4). The Canonical Group of conventional quantum mechanics is the *Weyl–Heisenberg group* $H(n)$ from section 1.6

and it yields the usual canonical commutation relations of $\hat{q}^i$, $\hat{p}_j$ and $\mathbb{1}$ thanks to the Stone–von Neumann theorem.

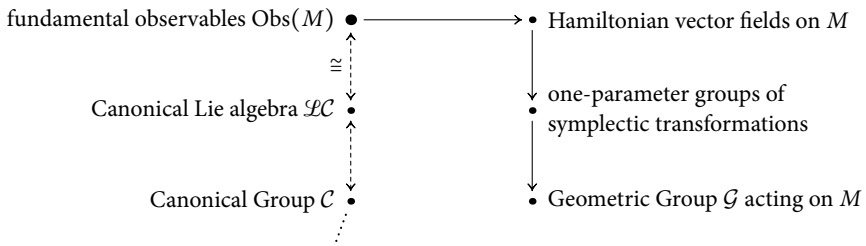
In general, however, we have neither a Canonical Group, nor is there a 'natural' set of fundamental quantizable observables. If the configuration space is the circle S^1, for instance, the most obvious choice of observable is the angle φ. Nevertheless, the angle φ is not a continuous function on S^1 and this attempt fails. Another case is example 1.6.4, a particle constrained to an interval, where the usual momentum operator $-i\hbar\partial_x$ is not self-adjoint. The exact same problem appeared in remark 1.2.4 where we considered the configuration space $Q = \mathbb{R}^+$ of the positive real line.

A more detailed discussion of why the angle φ doesn't work can be found in Kastrup's paper [Kas06, appendix A].

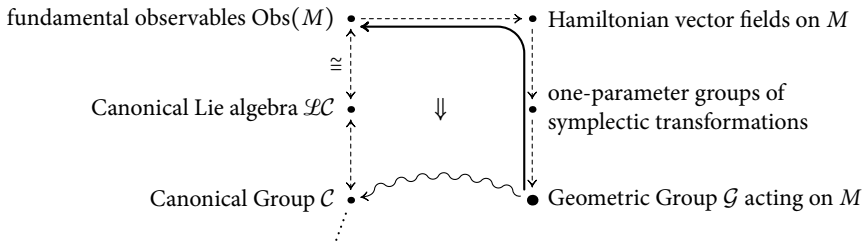
So, in the general case, where do the Canonical Group and the set of the fundamental observables come from?

Isham's idea is the following: Let us assume for the moment that we *already have* a complete set of fundamental observables, generating the Lie algebra $\mathrm{Obs}(M)$. Each function $f \in \mathrm{Obs}(M)$ produces a Hamiltonian vector field X_f and the corresponding flow $\Phi : \mathbb{R}\times M \to M$ if complete is a one-parameter group of symplectic transformations (more precisely, of *Hamiltonian symplectomorphisms*) of phase space M. Furthermore, all these one-parameter groups taken together form a Lie group $\mathcal{G}$ which acts on phase space by symplectic transformations (we will call this group together with its action on M the ***Geometric Group*** $\mathcal{G}$ for reasons that will become clear in a moment). Thus, we have the following picture:

A 'symplectic transformation', or 'symplectomorphism', is the mathematical term for what is better known in physics as 'canonical transformation' (see also chapter 4).



Isham's important observation in [Ish83] is that the above procedure can be *reversed*. Thus, it becomes possible to determine an appropriate Canonical Group $\mathcal{C}$ by looking for a group $\mathcal{G}$ of (global) symplectic transformations on M:



The *geometry* and the global *topological* structure of phase space will largely restrict which symplectic transformations are possible, and will often hint at a reasonable choice. As such, if we once again consider conventional quantum mechanics on $Q = \mathbb{R}^n$, the most obvious choice for $\mathcal{G}$ is the additive group $\mathcal{G} = (\mathbb{R}^{2n}, +)$ acting on the classical phase space $M = \mathbb{R}^{2n}$ by translations, and, in fact, the Canonical Group $\mathcal{C}$, which in this case is known to be the Weyl–Heisenberg group $H(n)$, can be constructed from $\mathcal{G}$ by a so-called *central extension* (the general construction is given in section 5.3).

We will discuss this example in detail in section 5.4.

Apart from these mathematical considerations there will be some additional criteria to be kept in mind which are necessary for an adequate *physical interpretation*:

The first point is that there might already *be* a group structure. As an example, when we discussed conventional quantum mechanics the vector space $\mathbb{R}^n$ as configuration space of a particle was a poor choice from a conceptual point of view. The actual physical setting is more accurately described by an *affine space* A^n, which comes pre-equipped with an action of the group $(\mathbb{R}^n, +)$, acting on A^n by translations. An adequate Geometric Group will usually have to be compatible with such a pre-given group action.

The idea of an affine space is generalised by the concept of G-spaces (see section 3.2).

Secondly, there is a connection between the realisation of the *group law* of the Geometric Group $\mathcal{G}$ and the physical interpretation of observables. As an example, suppose we have some observable that describes an *angle*. The 1-parameter subgroup of $\mathcal{G}$ that is related to the observable then has to act on phase space by a transformation that can be interpreted as a *rotation*. In terms of the subgroup this means that there will be some kind of periodicity, and this periodicity will reflect itself in the realisation of the group law of $\mathcal{G}$. Accordingly, the other way around, if the group action of the Geometric Group doesn't provide such a periodicity, the associated observable *cannot* be interpreted as an angle.

In chapter 5 we will discuss these issues more thoroughly after we have seen some concrete examples of how the quantization scheme works in practice. First, however, it is now the time to give an overview of the mathematical concepts we need to formulate the quantization method in more rigorous terms.

2 Manifolds and Fibre Bundles

> Although it is generally accepted that physics is written in the language of mathematics, there are disagreements on how much mathematical background is needed to give a proper description of physical phenomena.
>
> *Garbaczewski and Karwowski*, [GK04]

In this chapter we are going to summarise some facts from differential geometry. A good introductory textbook is the one written by Jänich [Jän01b; Jän01c]. For further reference I can highly recommend Michor's lectures [Mic08] as well as the books of Nakahara [Nak03] and Lee [Lee03]. An excellent review, particularly suitable for our purposes, is given by Waldmann [Wal07, chapters 2–4] (in German). Another useful presentation of differential geometry, strongly biased towards applications in classical mechanics, is given in the book of Marsden and Ratiu [MR99].

2.1 Differentiable Manifolds

Let us start with some basic definitions.

Definition 2.1.1 (Chart, transition map, atlas) Let M be a topological space. Then, an n-dimensional ***chart*** (U, x) of M is an open subset $U \subseteq M$ (the ***chart domain***) together with a homeomorphism $x : U \to \tilde{U}$ onto an open subset $\tilde{U} \subseteq \mathbb{R}^n$. A chart locally 'flattens' the space M.

Given two overlapping charts (U, x) and (V, y), that is with $U \cap V \neq \varnothing$, the map $y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$ is called ***transition map*** between these charts. By definition, it is always at least a homeomorphism between open subsets of $\mathbb{R}^n$.

An n-dimensional ***atlas*** $\mathfrak{A}$ of M is a collection of n-dimensional charts $\{(U_\alpha, x_\alpha)\}_\alpha$ such that the sets U_α constitute an open cover of M. A C^k-***atlas*** is an atlas whose transition maps are all C^k.

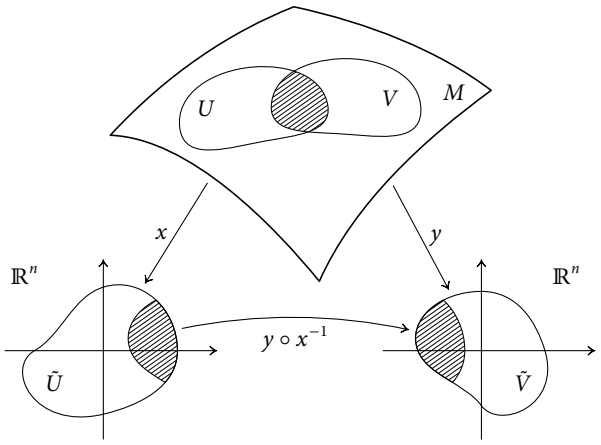


Figure 2.1 Transition map $y \circ x^{-1}$ between two charts (U, x) and (V, y).

Definition 2.1.2 (Equivalent atlases, differential structure) Two C^k-atlases $\mathfrak{A}_1$ and $\mathfrak{A}_2$ of a topological space M are said to be ***equivalent*** (or ***compatible***) if their union $\mathfrak{A}_1 \cup \mathfrak{A}_2$ is again a C^k-atlas of M.

An n-dimensional C^k-***differential structure*** of M is an equivalence class of n-dimensional C^k-atlases. Alternatively, a differential structure can be defined as a ***maximal atlas***, which is the union of all atlases in an equivalence class.

Definition 2.1.3 (Manifold) An n-***dimensional*** C^k-***manifold*** M is a second countable topological Hausdorff space equipped with an n-dimensional C^k-differential structure.

As special cases, a C^0-manifold is called ***topological manifold***; a C^∞-manifold is said to be a ***smooth manifold***.

While the demand for the differential structure is clear, it isn't immediately obvious why M needs to be second countable. The reason behind this requirement is the existence of a *partition of unity* which enables us to define integration on manifolds. The Hausdorff property, on the other hand, ensures that the manifold doesn't 'branch' (see the remark in [Peno5, section 12.2]).

Definition 2.1.4 (Differentiable map) A continuous map $\varphi : M \to N$ between two manifolds M and N is ***differentiable at*** $p \in M$ if it is differentiable at this point 'in charts', this is, if there are charts (U, x) around $p \in M$ and (V, y) around $\varphi(p) \in N$ such that $y \circ \varphi \circ x^{-1}$ is differentiable at $x(p) \in \mathbb{R}^n$ in the customary sense (see figure 2.2). The map φ is said to be ***differentiable*** (globally) if it is differentiable at every point $p \in M$.

Since $\mathbb{R}$ is a manifold itself – with global chart $(\mathbb{R}, \mathrm{id})$ – the definition also applies to differentiable functions $f : M \to \mathbb{R}$ and differentiable curves $\alpha : \mathbb{R} \to M$.

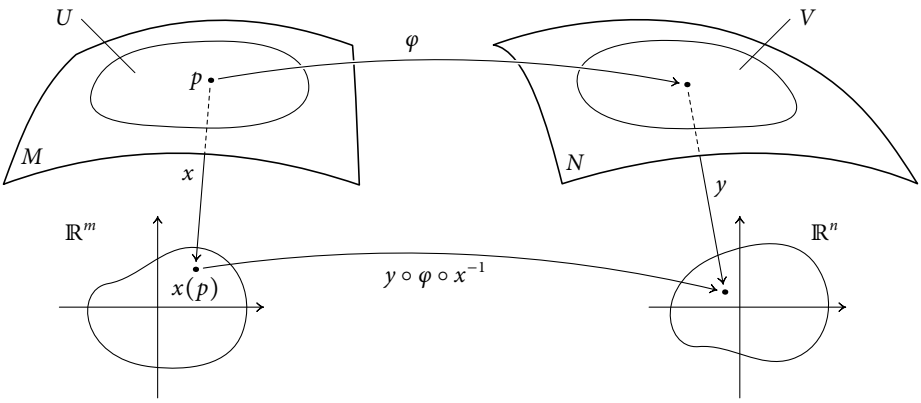


Figure 2.2 A map φ between manifolds is *differentiable* if it is differentiable 'in charts'.

2.2 Tangent Vectors and Tangent Spaces

Intuitively, we know how a plane tangent to a surface looks like. Although usable for submanifolds of $\mathbb{R}^n$, the visual picture of 'tangent to' fails when the surface isn't embedded in an ambient space, as it critically depends on it. For the general case we therefore need to characterise tangent vectors *without* reference to the surrounding space.

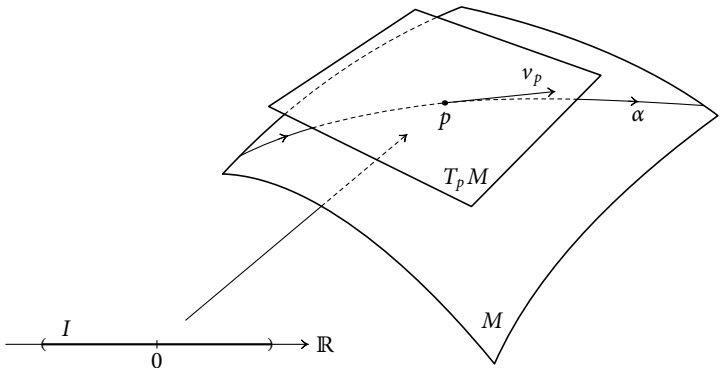


Figure 2.3 Differentiable curve α through $p = \alpha(0)$ and the corresponding tangent vector $v_p = [\alpha]$. Mention that the tangent vector v_p doesn't live *on* the manifold and can only be pictured this way if M is embedded in a surrounding space (in which case v_p is the *velocity* of the curve at time 0).

There are several different ways to define a tangent vector: via the velocity of a curve, as a derivation or by its transformation law. For finite-dimensional manifolds the three approaches are equivalent (a proof can be found in [Jäno1c, section 2.3]), only for infinite-dimensional manifolds one has to be more careful ([Ish83, section 3.1.2]).

Definition 2.2.1 (Curve) Let $I \subseteq \mathbb{R}$ be an open subset of $\mathbb{R}$ and let M be a manifold. A ***parameterised curve*** (or just ***curve***) α on M is a map:

$$\alpha : I \to M\,,\ t \mapsto \alpha(t)\,.$$

The curve is called ***smooth*** if α is differentiable (usually C^∞).

Using curves we can define tangent vectors in the following 'geometric' way:

Definition 2.2.2 (Tangent vector, tangent space) Let M be a manifold and let K_p be the set of differentiable curves α through $p \in M$, with $\alpha(0) = p$. Two such curves $\alpha, \beta \in K_p$ are ***tangentially equivalent***, $\alpha \sim \beta$, if they are tangent to each other 'in charts', i. e.:

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} x \circ \alpha(t) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} x \circ \beta(t)\,, \tag{2.1}$$

for any chart (U, x) around p. The equivalence classes $v_p = \dot{\alpha}(0) := [\alpha]$ for $\alpha \in K_p$ are called ***tangent vectors*** at p and $T_pM := K_p/\!\sim = \{[\alpha] : \alpha \in K_p\}$ is the ***tangent space*** at p.

In addition to the geometric viewpoint, tangent vectors can also be characterised *algebraically* by their action on smooth functions. That said, since a tangent vector is defined locally at a point $p \in M$ we don't need the functions to be defined on the whole manifold, just in a small neighbourhood of p. The following definition of *function germs* implements this idea.

Definition 2.2.3 (Germs of functions) Let f, g be two real-valued functions, defined and C^k-differentiable in a neighbourhood of a point $p \in M$. We will consider them *germ equivalent* if there exists an open neighbourhood U of p such that $f|_U = g|_U$. The

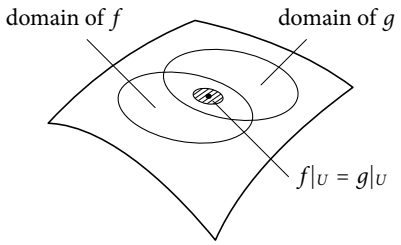


Figure 2.4 The two functions f and g define the same *germ*.

equivalence classes $[f]_p$ are called (C^k*-differentiable*) ***germs of functions at*** p, and the set of all function germs $[f]_p$ at p is denoted by $C^k_p(M)$.

Germs of functions at p can be added and multiplied with each other, thereby giving $C^k_p(M)$ the structure of an associative ring with unity. We can identify the real numbers $\mathbb{R}$ with constant functions to induce a scalar multiplication compatible with the ring structure. The space of germs $C^k_p(M)$ is hence not just a ring but a *real associative algebra with unity*.

Definition 2.2.4 (Tangent vector as a derivation) Given a C^k-manifold M, an ***'algebraic' tangent vector*** v_p at $p \in M$ is a ***derivation*** of the algebra of function germs at p, namely a linear map:

$$v_p : C^k_p(M) \to \mathbb{R} ,$$

that satisfies *Leibnitz's law*:

$$v_p(fg) = v_p(f)g(p) + f(p)v_p(g) .$$

The 'geometric' tangent vectors $[\alpha] \in T_pM$ from above become 'algebraic' ones when we define the action on differentiable germs of functions $f \in C^1_p(M)$ as:

$$[\alpha](f) = \dot{\alpha}(0)f := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (f \circ \alpha)(t) . \tag{2.2}$$

Finally, we want to work our way towards the definition of tangent vectors 'in coordinates', the way they are commonly used in most physics textbooks. Given an n-dimensional manifold M and a chart (U, x) around $p \in M$, a basis of the tangent space T_pM is given by the derivations:

$$\left.\frac{\partial}{\partial x^i}\right|_p : C^1_p \to \mathbb{R} , \quad f \mapsto \left.\frac{\partial}{\partial x^i}\right|_p f := \left.\frac{\partial (f \circ x^{-1})}{\partial x^i}\right|_{x(p)} , \tag{2.3}$$

where $f \circ x^{-1} : \tilde{U} \subseteq \mathbb{R}^n \to \mathbb{R}$ is the *coordinate representation* of f. A general tangent vector $v_p \in T_pM$ can then be expressed as a linear combination:

$$v_p = v^i_p \left.\frac{\partial}{\partial x^i}\right|_p , \quad \text{with } v^i_p = v_p(x^i) \in \mathbb{R} . \tag{2.4}$$

Unfortunately, since the coordinates v^i_p depend on this specific chart (U, x) they can't really characterise v_p in an invariant way. Expressing v_p in a different chart (V, y)

around p, we obtain:

$$\begin{aligned} v_p(f) &= v_p^i \left.\frac{\partial}{\partial x^i}\right|_p (f) = v_p^i \left.\frac{\partial (f \circ x^{-1})}{\partial x^i}\right|_{x(p)} = v_p^i \left.\frac{\partial (f \circ y^{-1} \circ y \circ x^{-1})}{\partial x^i}\right|_{x(p)} \\ &= v_p^i \left.\frac{\partial (f \circ y^{-1})}{\partial y^j}\right|_{y(p)} \left.\frac{\partial (y \circ x^{-1})^j}{\partial x^i}\right|_{x(p)} = v_p^i \left.\frac{\partial (y \circ x^{-1})^j}{\partial x^i}\right|_{x(p)} \left.\frac{\partial}{\partial y^j}\right|_p (f) \\ &\overset{!}{=} \tilde{v}_p^i \left.\frac{\partial}{\partial y^j}\right|_p (f) \,. \end{aligned}$$

Thus, when changing charts, the *coordinates* have to be transformed via the *Jacobian matrix* of the transition map, the *basis vectors* via the inverse:

$$\tilde{v}_p^j = v_p^i \left.\frac{\partial (y \circ x^{-1})^j}{\partial x^i}\right|_{x(p)} , \quad \left.\frac{\partial}{\partial y^j}\right|_p = \left.\frac{\partial (x \circ y^{-1})^i}{\partial y^j}\right|_{x(p)} \left.\frac{\partial}{\partial x^i}\right|_p . \tag{2.5}$$

Using this, we can make the vague notion of '*a tangent vector is a tangent vector if it transforms like one*' more explicit by the following definition.

Definition 2.2.5 (Tangent vector 'in coordinates') Consider pairs of coordinates and basis vectors of T_pM (see equation (2.3)):

$$\left(v_p^i \,;\, \left.\frac{\partial}{\partial x^i}\right|_p\right) .$$

Two pairs are said to be *equivalent* if they are related to each other by a transformation as in equation (2.5). A ***tangent vector 'in coordinates'*** is then an equivalence class with respect to this relation:

$$\left[v_p^i \,;\, \left.\frac{\partial}{\partial x^i}\right|_p\right] .$$

Intuitively, one can think of such a tangent vector as a collection of the coordinate representations in all possible bases.

Definition 2.2.6 (Tangent bundle) The ***tangent bundle*** TM of an n-dimensional manifold M is the *disjoint union* of all tangent spaces:

$$TM := \coprod_{p \in M} T_pM := \bigcup_{p \in M} \{p\} \times T_pM \,.$$

The set TM is – in a canonical way – a $2n$-dimensional manifold, where the charts are given by maps $TM \ni (p, v_p) \mapsto \left(x^1(p), \ldots, x^n(p), v_p(x^1), \ldots, v_p(x^n)\right) \in x(U) \times \mathbb{R}^n$, when (U, x) is a chart of M around the point p. In addition, there is a *natural projection* $\pi : TM \to M \,,\ (p, v_p) \mapsto p$, turning TM into a *vector bundle*, a special case of a *fibre bundle*. The preimage $\pi^{-1}(p)$ of $p \in M$ is called the *fibre over* p, in this specific case $\pi^{-1}(p) = \{p\} \times T_pM \cong T_pM$.

We will take a closer look at fibre bundles in section 2.5.

2.3 The Tangent Map

A MAPPING $\varphi : M \to N$ between two manifolds can be used to 'push' tangent vectors from one manifold to the other. The idea is – when viewing a tangent vector at the point $p \in M$ as represented by a curve α on M – that φ maps α to a curve $\varphi \circ \alpha$ on N. The resulting curve then defines a new tangent vector at the point $\varphi(p) \in N$ as shown in figure 2.5 on the following page. The following definition formalises this idea.

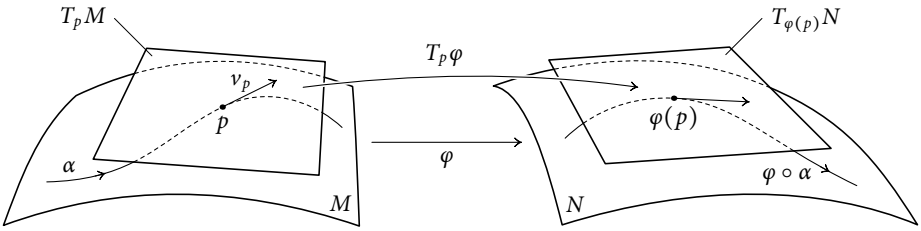


Figure 2.5 Tangent map $T_p\varphi$ between two manifolds. The function $\varphi : M \to N$ can be used to push the curve α on M to the manifold N and thereby induces a transformation of tangent vectors $v_p \mapsto T_p\varphi(v_p)$.

Definition 2.3.1 (Tangent map) Let $\varphi : M \to N$ be a differentiable map between two manifolds. The ***tangent map*** $T_p\varphi$ ***of*** φ ***at the point*** $p \in M$ (also called ***pushforward*** or ***differential***) is:

$$T_p\varphi : T_pM \to T_{\varphi(p)}N\,, \quad [\alpha] \mapsto T_p\varphi([\alpha]) := [\varphi \circ \alpha]\,. \tag{2.6}$$

In addition, we define the (global) ***tangent map*** $T\varphi : TM \to TN$ between the tangent bundles by $T\varphi|_{T_pM} := T_p\varphi$. The tangent map satisfies the commutative diagram:

$$\begin{array}{ccc} TM & \xrightarrow{T\varphi} & TN \\ {\scriptstyle \pi_M}\downarrow & & \downarrow{\scriptstyle \pi_N} \\ M & \xrightarrow{\varphi} & N \end{array}\,,$$

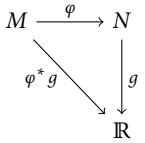
where $\pi_M : TM \to M$ and $\pi_N : TN \to N$ are the natural projections of the tangent bundles TM and TN, respectively (see definition 2.2.6).

We write $\varphi^* v_p := (\varphi^{-1})_* v_p$ for the pushforward of tangent vectors via the inverse of φ.

There exists a variety of alternative notations. In common use are $\mathrm{d}\varphi_p$ or $\mathrm{D}\varphi_p$, since the tangent map at a point generalises the usual differential from linear algebra. To emphasise the relation to the *pullback*, the tangent map is also frequently written as $\varphi_{*,p}$ and then usually called *pushforward*. Moreover, in this case, the point p and brackets around the argument are often omitted, resulting in the notation $\varphi_*[\alpha]$ or $\varphi_* v_p$. The same notation, φ_*, is widely used for the global tangent map.

Remark 2.3.2 If we take the algebraic viewpoint and consider tangent vectors as *derivations* (as in definition 2.2.4), the tangent map assumes the form:

$$T_p\varphi(v_p) = v_p \circ \varphi^*\,, \tag{2.7}$$



where the so-called *pullback* φ^* acts on germs of functions like:

$$\varphi^* : C^k_{\varphi(p)}(N) \to C^k_p(M)\,, \quad g \mapsto \varphi^*(g) := g \circ \varphi\,.$$

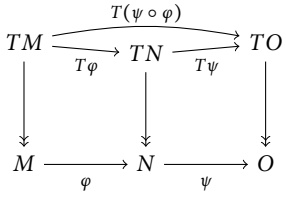
Evaluating equation (2.7) on functions then yields the explicit formula:

$$\big(T_p\varphi(v_p)\big)(g) = v_p \circ \varphi^*(g) = v_p(g \circ \varphi)\,. \tag{2.8}$$

Remark 2.3.3 The *coordinate representation* of the tangent map – using charts (U, x) around $p \in M$ and (V, y) around $\varphi(p) \in N$ – is given by:

$$T_p\varphi\Big(v_p^i \frac{\partial}{\partial x^i}\Big|_p\Big) = v_p^i \frac{\partial(y^j \circ \varphi \circ x^{-1})}{\partial x^i}\Big|_{x(p)} \frac{\partial}{\partial y^j}\Big|_{\varphi(p)} . \tag{2.9}$$

Lemma 2.3.4 (Chain rule) Given two differentiable maps $\varphi : M \to N$ and $\psi : N \to O$ between manifolds, the following *chain rule* holds:



$$T(\psi \circ \varphi) = T\psi \circ T\varphi . \tag{2.10}$$

Furthermore, we have:

$$T\,\mathrm{id}_M = \mathrm{id}_{TM} . \tag{2.11}$$

Proof. The proof is a straightforward application of formula (2.8). For $v_p \in T_pM$ and a differentiable function germ $g \in C^k_{\psi\circ\varphi(p)}(O)$ we obtain:

$$\begin{aligned}\big(T_p(\psi \circ \varphi)(v_p)\big)(g) &= v_p\big(g \circ (\psi \circ \varphi)\big) = v_p\big((g \circ \psi) \circ \varphi\big) \\ &= \big(T_p\varphi(v_p)\big)(g \circ \psi) \\ &= \big(T_{\varphi(p)}\psi \circ T_p\varphi(v_p)\big)(g) ,\end{aligned}$$

where, in the last step, $T_p\varphi(v_p)$ is simply a tangent vector in $T_{\varphi(p)}N$.

The second identity follows from:

$$\big(T\,\mathrm{id}_M(v_p)\big)(f) = v_p(f \circ \mathrm{id}_M) = v_p(f) = \big(\mathrm{id}_{TM}(v_p)\big)(f) ,$$

for all function germs $f \in C^k_p(M)$ and all $p \in M$. ∎

Equations (2.10) and (2.11) are the defining properties of a *covariant functor*, a map between *categories*. Categories combine objects (e. g. sets, groups, vector spaces, …) with compatible 'relations' between them (e. g. equivalence relations, group homomorphisms, linear maps, …). More precisely:

Definition 2.3.5 (Category) A ***category*** $\mathcal{C}$ consists of:

- a class of ***objects*** $\mathrm{Obj}(\mathcal{C})$,
- to each pair (A, B) of objects a set of ***morphisms*** (or ***arrows***) $\mathrm{Hom}(A, B)$ between these objects. An element $f \in \mathrm{Hom}(A, B)$ is typically written as $f : A \to B$ (yet f doesn't have to be a function in general). The object A is called the ***source***, and B the ***target*** of f. We will denote the class of *all* morphisms by $\mathrm{Hom}(\mathcal{C})$ but some authors prefer to write $\mathrm{Mor}(\mathcal{C})$ instead.
- a binary operation $\circ : \mathrm{Hom}(A, B) \times \mathrm{Hom}(B, C) \to \mathrm{Hom}(A, C)$, $(f, g) \mapsto f \circ g$, called ***composition*** of morphisms,

such that the following axioms hold:

- the composition is *associative*, $(f \circ g) \circ h = f \circ (g \circ h)$,
- for each object A, there exists an ***identity morphism*** $\mathrm{id}_A \in \mathrm{Hom}(A, A)$ that satisfies $\mathrm{id}_A \circ f = f$ and $g \circ \mathrm{id}_A = g$.

It should be noted that the sets $\mathrm{Hom}(A, B)$ of morphisms are disjoint unless both domain *and* codomain match.

More details on categories and functors can be found in the book of Hilton and Stammbach [HS71, chapter II]. The standard reference is Mac Lane's work [Mac98].

Example 2.3.6 Some important examples of categories include:

- The category Grp of groups, where the objects are groups and the morphisms are the group homomorphisms.
- A group G itself can be seen as a category with a single object x, where all the morphisms are invertible, i. e. isomorphisms. The morphisms then correspond to the group *elements*; the group properties are provided by the category laws. The *neutral element* of the group is the identity morphism $\mathrm{id}_x : x \to x$.
- A slight generalisation of the last example: If we have a set of multiple objects and all the morphisms are isomorphisms we obtain a so-called *groupoid*. The difference is that not all morphisms can be composed with each other. In terms of groups this means that multiplication is only a *partial functions*, i. e. not all pairs of groupoid elements can be multiplied with each other.
- The category Man^k of differentiable C^k-manifolds where the objects are C^k-manifolds and the morphisms are C^k-differentiable maps between them.
- The category $\mathrm{O}(M)$ where the objects are open subsets of a topological space M and the morphisms are inclusions. The set $\mathrm{Hom}(A, B)$ thus contains a single morphism if $A \subseteq B$ and otherwise is empty.
- An abstract example: A commutative diagram can be interpreted as a category where the objects are the vertices of the diagram and the morphisms are the arrows. Note that the morphisms in this case are not functions but they really are seen as the *edges* in a directed graph.

Groupoids will appear again later in section 3.2 where we talk about group actions. In fact, so-called G-sets can be seen as *action groupoids*.

Definition 2.3.7 (Functor) Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A (***covariant***) ***functor*** $F : \mathcal{C} \to \mathcal{D}$ is a mapping that associates:

- to each object $A \in \mathrm{Obj}(\mathcal{C})$ an object $F(A) \in \mathrm{Obj}(\mathcal{D})$,
- to each morphism $f : A \to B$ in $\mathcal{C}$ a morphism $F(f) : F(A) \to F(B)$ in $\mathcal{D}$,

such that:

- identities are preserved: $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$,
- for any two morphisms $f \in \mathrm{Hom}(A, B)$ and $g \in \mathrm{Hom}(B, C)$ we have the chain rule: $F(f \circ g) = F(f) \circ F(g)$.

A ***contravariant functor*** $G : \mathcal{C} \to \mathcal{D}$ is defined the same way except that it *reverses* morphisms, namely $f : A \to B$ is mapped to $G(f) : G(B) \to G(A)$ and the chain rule changes to $G(f \circ g) = G(g) \circ G(f)$.

Remark 2.3.8 (Tangent functor) The ***tangent functor*** $T : \mathrm{Man}^k \to \mathrm{VB}$, acting on differentiable manifolds (objects) by $M \mapsto TM$ and on differentiable maps (morphisms) by $\varphi \mapsto T\varphi$, is a *covariant functor* from the category Man^k of C^k-manifolds (with differentiable maps) to the category VB of vector bundles (with vector bundle morphisms).

In contrast to the tangent functor, T_p can not be a functor starting in the category of manifolds because T_pM only makes sense if p is a point in M. Nevertheless, it *is* a functor from the category of *pointed manifolds* (the objects are pairs (M, p) consisting of a manifold M and a chosen base point $p \in M$; the morphisms are differentiable maps between manifolds that preserve the base points) to the category $\mathrm{Vect}_{\mathbb{R}}$ of real vector spaces, with linear maps as morphisms.

Lie groups can be thought of as pointed manifolds. The *Lie functor* $\mathfrak{L}$ (definition 3.3.7) is essentially equal to T_e.

2.4 Vector Fields, Flow, and the Lie Derivative

Definition 2.4.1 (Vector field) Given some manifold M, a ***vector field*** X on M is a continuous map:

$$X : M \to TM\,, \quad p \mapsto (p, X_p)\,,$$

that attaches a tangent vector $X_p \in T_pM$ to each point $p \in M$ (see figure 2.6).

Alternatively, we can interpret a vector field as a ***section*** of the tangent bundle TM, i. e. a map $X : M \to TM$ with $\pi \circ X = \mathrm{id}_M$. The condition hereby ensures that $X(p)$ ends up in the fibre $\pi^{-1}(p)$ over p, such that the tangent vector X_p lies in the correct tangent space T_pM.

If this viewpoint of vector fields as sections should be emphasised they are usually denoted by $\Gamma(TM)$. The set of ***smooth vector fields*** on M (where X is required to be a C^∞-differentiable function) will be written as $\Gamma^\infty(TM)$ or simply as $\mathfrak{X}(M)$.

Remark 2.4.2 The vector fields $\Gamma^k(TM)$ possess the structure of a $C^k(M)$-module over the algebra of functions thanks to the definition:

$$(fX + gY)(p) := f(p)X(p) + g(p)Y(p)\,,$$

with $X, Y \in \Gamma^k(TM)$ and $f, g \in C^k(M)$.

Definition 2.4.3 (Vector field as derivation) A vector field $X \in \Gamma^\infty(TM)$ naturally acts on functions $f \in C^\infty(M)$ via:

$$(Xf)(p) := X_p(f) \quad \forall p \in M\,. \tag{2.12}$$

Moreover, for any two functions f and g, we have the product rule:

$$X(fg) = (Xf)g + f(Xg)\,.$$

Thus, vector fields can be seen as *derivations* of the algebra of functions $C^\infty(M)$.

For finite-dimensional, differentiable C^k-manifolds with $k \geq 2$, the vector fields are actually *isomorphic* to the vector space of derivations of functions on M (see [AMR88, section 4.2]).

This interpretation of vector fields coincides with the *Lie derivative of functions*:

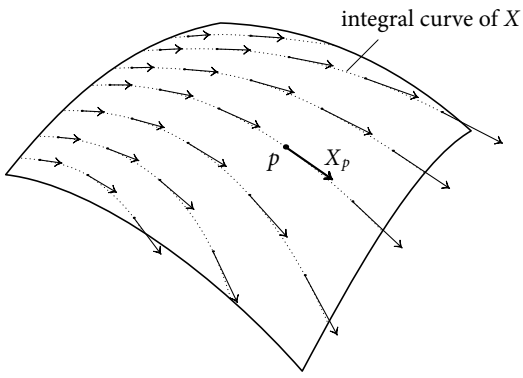


Figure 2.6 Vector field on a manifold. The field X is a prescription that attaches a tangent vector X_p to each point $p \in M$, located in the tangent space T_pM. Integral curves are tangent to the vector field at each point.

A geometrical interpretation of the Lie derivative is given in the remark after lemma 2.4.11.

Definition 2.4.4 (Lie derivative) For functions, the ***Lie derivative*** $\mathcal{L}_X f$ ***along*** X is:

$$\mathcal{L}_X : C^\infty(M) \to C^\infty(M) , \quad \mathcal{L}_X f := Xf . \tag{2.13}$$

The product rule then looks like:

$$\mathcal{L}_X(fg) = \mathcal{L}_X(f)g + f\mathcal{L}_X(g) . \tag{2.14}$$

In this definition we used the derivational viewpoint of *tangent vectors* to define the same concept for the *vector field*. Similarly, since we can transport tangent vectors from one manifold to another by means of the tangent map, it seems as if we could push whole vector fields between manifolds as well. Unfortunately, there are problems:

Consider a differentiable map $\varphi : M \to N$ between manifolds and a vector field X on M. Pointwise application of the tangent map yields $T_p\varphi(X_p) \in T_{\varphi(p)}(N)$. We can see immediately that – since φ is in general not surjective – the result doesn't define a vector field on the whole manifold N. The vector field wouldn't even be well-defined unless φ were injective. Moreover, the result still depends on $p \in M$ and not on a point in N as it should be the case.

Still, it makes sense that two vector fields are kind of 'related' in this way:

Definition 2.4.5 (φ-related vector fields) Let M, N be manifolds, X a vector field on M, Y a vector field on N. The vector fields X is said to be φ-***related*** to Y if there exists a continuous map $\varphi : M \to N$ such that the following diagram commutes:

$$\begin{array}{ccc} TM & \xrightarrow{T\varphi} & TN \\ {\scriptstyle X}\uparrow & & \uparrow{\scriptstyle Y} \\ M & \xrightarrow{\varphi} & N \end{array} ,$$

that is:

$$T\varphi \circ X = Y \circ \varphi , \quad \text{or} \quad T_p\varphi(X_p) = Y_{\varphi(p)} \quad \forall p \in M . \tag{2.15}$$

If, on the other hand, φ is a *diffeomorphism* it is possible to define the pushforward and pullback of vector fields in the following way:

Instead of $\varphi_\# X$, a lot of authors use the notation $\varphi_* X$ also for the pushforward of vector fields and leave it to the reader to figure things out.

$$\begin{array}{ccc} TM & \xrightarrow{T\varphi} & TN \\ {\scriptstyle X}\uparrow & \overset{\varphi_\#}{\Rightarrow} & \uparrow{\scriptstyle \varphi_\# X} \\ M & \xrightarrow{\varphi} & N \end{array}$$

Definition 2.4.6 (Pushforward of a vector field) Let $\varphi : M \to N$ be a diffeomorphism between manifolds and let X be a vector field on M. The ***pushforward of*** X is the vector field $\varphi_\# X \in \Gamma(TN)$, given by (see also section 4.5 for a further discussion):

$$(\varphi_\# X)_q := \varphi_*(X_{\varphi^{-1}(q)}) = T\varphi(X_{\varphi^{-1}(q)}) , \tag{2.16}$$

for each point $q \in N$. The field X is then φ-related to $\varphi_\# X$.

In addition, we define the ***pullback of*** $Y \in \mathfrak{X}(N)$ as the pushforward via the inverse φ^{-1}:

$$\varphi^\# Y := (\varphi^{-1})_\# Y .$$

In this case, the pullbacked field $\varphi^\# Y$ is φ-related to Y.

VECTOR FIELDS CAN be used to formulate ordinary differential equations on manifolds. The solutions are given by integral curves:

Definition 2.4.7 (Integral curve) Let X be a vector field on M. A differentiable curve $\gamma : I \to M$, defined on an interval $I \subseteq \mathbb{R}$, is called ***integral curve*** (or ***flow curve***) ***of*** X ***through*** $p \in M$ if:

$$\dot{\gamma}(t) = X_{\gamma(t)} , \quad \gamma(0) = p . \tag{2.17}$$

This tells us that the velocity of γ at time t is equal to the tangent vector $X_{\gamma(t)}$ at this point (see figure 2.6 on page 47).

If X is Lipschitz continuous, the Picard–Lindelöf theorem guarantees both existence and uniqueness of maximal integral curves.

We call γ a ***maximal*** integral curve if there exists no integral curve through p with a domain $I' \supset I$ larger than the domain I of γ. If γ is defined for all times $t \in \mathbb{R}$, i. e. the domain is equal to $I = \mathbb{R}$, the integral curve is said to be ***complete***.

If we take the point $p \in M$ to be variable we obtain the *flow* of X:

Definition 2.4.8 (Flow) The ***flow*** Φ of a vector field X on M is the continuous map:

$$\Phi : \mathbb{R} \times M \supseteq U \to M , \quad (t, p) \mapsto \Phi(t, p) ,$$

where $\Phi(\,\cdot\,, p) : I_p \subseteq \mathbb{R} \to M$ is the maximal integral curve through p and U is the maximal open subset $U \subseteq \bigcup_{p\in M}\{p\} \times I_p$. Explicitly, the flow satisfies:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, p) = X_{\Phi(t,p)} , \quad \Phi(0, p) = p . \tag{2.18}$$

If $U = M \times \mathbb{R}$, the flow is called ***complete*** or ***global flow***. If X possesses a complete flow, it is said to be a ***complete vector field***.

As a special case, compactly supported vector fields on a manifold are always complete (see [Mico8, lemma 3.8]).

Furthermore, we will write $\Phi_t := \Phi(t, \,\cdot\,)$.

Lemma 2.4.9 ('Group law' for the flow) The flow satisfies:

$$\Phi_r \circ \Phi_s = \Phi_{r+s} , \tag{2.19}$$

for all $r, s \in \mathbb{R}$ where Φ_r, Φ_s and Φ_{r+s} is defined.

Proof. Given $f(r) := \Phi_r \circ \Phi_s(p)$, equation (2.18) yields:

$$\frac{\mathrm{d}}{\mathrm{d}r} f(r) = X_{f(r)} , \quad f(0) = \Phi_0 \circ \Phi_s(p) = \Phi_s(p) .$$

On the other hand, using $g(r) := \Phi_{r+s}(p)$, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}r} g(r) = X_{g(r)} , \quad g(0) = \Phi_{0+s}(p) = \Phi_s(p) .$$

Since both f and g satisfy the same differential equation and initial condition, the theorem of Picard–Lindelöf tells us they have to be identical. Thus, equation (2.19) holds for all times where the solution to the differential equation exists. ∎

If the flow Φ is complete the set $\{\Phi_t : t \in \mathbb{R}\}$ is a *one-parameter group* of transformations of M (via diffeomorphisms). Otherwise the transformations only form a one-parameter *groupoid* (see example 2.3.6 on page 46 for the definition; the resulting object is also called *local pseudo group* in [CDD77, section III.C]).

Definition 2.4.10 (Generator of a transformation) Let X be a smooth vector field on a manifold M and let Φ be its flow. Then X is called the ***generator*** of the one-parameter groupoid $\{\Phi_t\}$ of local transformations of M.

If X is compactly supported it generates a one-parameter subgroup of global transformations of M [CDD77, p. 143].

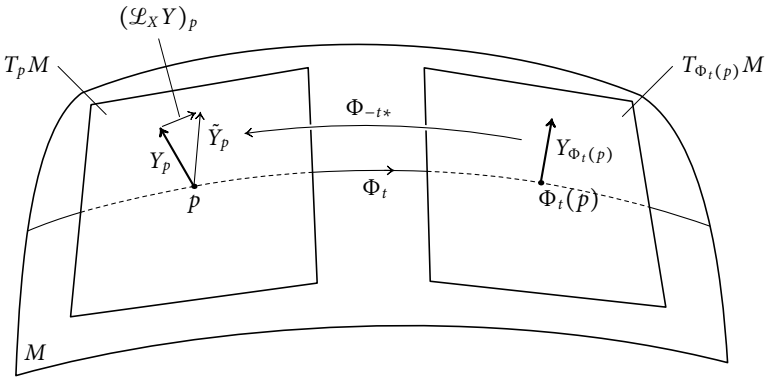


Figure 2.7 Lie derivative of a vector field. The Lie derivative $\mathcal{L}_X Y$ at the point p is effectively the difference between $Y_{\Phi_t(p)}$ and Y_p. To compare them, the former has to be transported back to T_pM using the flow Φ of X, thereby giving $\tilde{Y}_p$.

Using the flow, we can reformulate the Lie derivative of a function:

Lemma 2.4.11 (Lie derivative) Let M be a smooth manifold, X a smooth vector field on M and Φ the flow generated by X. Then, the Lie derivative $\mathcal{L}_X f$ of a function f can be written as:

$$\mathcal{L}_X f = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Phi_t^* f = \lim_{t\to 0} \frac{1}{t}\Big(\Phi_t^* f - f\Big)\,. \tag{2.20}$$

Proof. Given that the curve $\Phi(\,\cdot\,, p)$ defines the tangent vector X_p we can use the action of tangent vectors on functions from equation (2.2) to calculate the Lie derivative of f at the point p:

$$\mathcal{L}_X f(p) = (Xf)(p) = X_p f = [\Phi(\,\cdot\,, p)]f = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\Phi(t,p) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\Phi_t(p)\,.$$

Using the pullback notation $\Phi_t^* f = f\circ\Phi_t$, this is exactly what had to be shown. ∎

Equation (2.20) allows a geometrical interpretation of the Lie derivative: $\mathcal{L}_X f$ at the the point p is effectively the difference of $f(\Phi_t(p))$ and $f(p)$ when t goes to zero. This means that the Lie derivative $\mathcal{L}_X f$ measures the *change* of f *along the flow* of X. For this reason, $\mathcal{L}_X$ it is sometimes called the 'fisherman's derivative'.

Another important thing about equation (2.20) is that it can be used to define the Lie derivative of vector fields:

Definition 2.4.12 (Lie derivative of a vector field) Let M be a smooth manifold, let X and Y be vector fields on M and let Φ be the flow of X. Then, the ***Lie derivative*** $\mathcal{L}_X Y$ ***of*** Y ***along*** X is defined as:

$$(\mathcal{L}_X Y)_p = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\Phi_t^{\#} Y)_p = \lim_{t\to 0} \frac{1}{t}\Big(\Phi_t^* Y_{\Phi_t(p)} - Y_p\Big)\,, \tag{2.21}$$

where $\Phi_t^* = \big(\Phi_t^{-1}\big)_* = \Phi_{-t*}$ is the pushforward via the inverse of Φ_t (the definition of the Lie derivative is illustrated in figure 2.7).

Definition 2.4.13 (Jacobi–Lie bracket) Given vector fields $\mathfrak{X}(M)$ on a smooth manifold M, the ***Jacobi–Lie bracket*** is the $\mathbb{R}$-bilinear mapping:

$$\begin{aligned}[\,\cdot\,,\cdot\,] : \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathfrak{X}(M)\,, \\ (X, Y) &\mapsto [X, Y](f) := X\big(Y(f)\big) - Y\big(X(f)\big)\,. \end{aligned} \tag{2.22}$$

Lemma 2.4.14 (Jacobi–Lie bracket and commutator) The Jacobi–Lie bracket of two vector fields X and Y is the same as the Lie derivative of Y along X:

$$[X, Y](f) = \mathcal{L}_X Y\,.$$

Proof. A proof is given in [Mico8, lemma 3.13]. ∎

Lemma 2.4.15 (Properties of the Jacobi–Lie bracket) The Jacobi–Lie bracket has the following properties:

(i) $[X, Y] = -[Y, X]\,,$ (*antisymmetry*)

(ii) $\big[X, [Y, Z]\big] = \big[[X, Y], Z\big] + \big[Y, [X, Z]\big]\,,$ (*Jacobi identity*)

(iii) $[fX, Y] = f[X, Y] - (Yf)X\,,$

(iv) $[X, fY] = f[X, Y] + (Xf)Y\,.$

Proof. The first and second property are simple consequences of the definition via the commutator. The third one uses that vector fields form a $C^\infty(M)$-module and that they act on functions as derivations:

$$\begin{aligned}[fX, Y](g) &= (fX)\big(Y(g)\big) - Y\big((fX)(g)\big) = f\,X\big(Y(g)\big) - Y\big(f\,X(g)\big) \\ &= f\,X\big(Y(g)\big) - f\,Y\big(X(g)\big) - Y(f)\,X(g) \\ &= f\,[X, Y](g) - (Yf)X(g)\,. \end{aligned}$$

The last property is just a rewrite of the third one, because $[fX, Y] = -[Y, fX]$. ∎

Remark 2.4.16 (Vector fields as Lie algebra) The vector fields $\mathfrak{X}(M)$ together with the Jacobi–Lie bracket $[\,\cdot\,,\cdot\,]$, that satisfies the first two properties of lemma 2.4.15, form the prototype of a *Lie algebra*. The Jacobi identity says that $[X, \cdot\,]$ is a derivation of the Lie algebra. To see this, use the notation $\mathrm{ad}(X) = [X, \cdot\,]$ and write $X * Y = [X, Y]$ for the bracket. The Jacobi identity then reads:

$$\mathrm{ad}(X)(Y * Z) = \mathrm{ad}(X)(Y) * Z + Y * \mathrm{ad}(X)(Z)\,,$$

which we recognise as the usual product rule.

An important result is that the pullback (pushforward via the inverse) of vector fields via a diffeomorphism is a Lie algebra homomorphism:

Lemma 2.4.17 (Pullback and Jacobi–Lie bracket) Let $\varphi : M \to N$ be a diffeomorphism and let $Y_1, Y_2 \in \mathfrak{X}(N)$ be vector fields. The pullback $\varphi^\# : \mathfrak{X}(N) \to \mathfrak{X}(M)$ is then a Lie algebra homomorphism, that is:

$$\varphi^\#\,[Y_1, Y_2] = [\varphi^\# Y_1, \varphi^\# Y_2]\,. \tag{2.23}$$

Note that the proof is even a bit more general than necessary because only *φ-related* vector fields X_i and Y_i are assumed.

Proof (see [Mico8, lemma 3.10 and corollary 3.11]). Let us write $X_i := \varphi^\# Y_i$, then X_i is φ-related to Y_i, i. e. $T\varphi \circ X_i = Y_i \circ \varphi$. Given a smooth function $g \in C^\infty(N)$ we have:

$$\begin{aligned} X_i(g\circ\varphi)(p) &= X_{i,p}(g\circ\varphi) = \big(T_p\varphi(X_{i,p})\big)(g) = (T\varphi\circ X_i)_p(g) \\ &= (Y_i\circ\varphi)_p(g) = Y_{i,\varphi(p)}(g) = (Y_i g)\big(\varphi(p)\big) \quad \forall p\in M\,, \end{aligned}$$

and therefore $X_i(g\circ\varphi) = (Y_i g)\circ\varphi$. Using this we obtain:

$$\begin{aligned} \big(T\varphi\circ[X_1,X_2]\big)(g) &= [X_1,X_2]\,(g\circ\varphi) = X_1\big(X_2(g\circ\varphi)\big) - X_2\big(X_1(g\circ\varphi)\big) \\ &= X_1\big((Y_2 g)\circ\varphi\big) - X_2\big((Y_1 g)\circ\varphi\big) = Y_1\big((Y_2 g)\big)\circ\varphi - Y_2\big((Y_1 g)\big)\circ\varphi\,, \end{aligned}$$

which means $[X_1, X_2]$ is φ-related to $[Y_1, Y_2]$, in other words $\varphi^\#[Y_1, Y_2] = [X_1, X_2]$. Reinserting the definition of X_i from above gives the relation we need. ∎

2.5 Fibre Bundles

Informally speaking, a fibre bundle is a space that *locally* 'looks' like a product space $B \times F$ but *globally* may be more complicated. An example is the tangent bundle TM – introduced in definition 2.2.6 – which locally looks like $M \times \mathbb{R}^n$ but globally can have a nontrivial topological structure.

Some authors (e. g. [Nako3]) prefer to include a 'structure group' in the definition of fibre bundles. We call such bundles *G-bundles* (definition 2.5.5).

Definition 2.5.1 (Fibre bundle) A ***fibre bundle*** (E, B, π, F), which we will denote by $\pi : E \to B$ or just E for short, consists of:

- a manifold E (the *total space*),
- a manifold B (the *base space*),
- a manifold F (the *typical fibre* or *standard fibre*),
- a smooth mapping $\pi : E \to B$ (the *bundle projection*).

We sometimes use the notation $E_p := \pi^{-1}(p)$ for the *fibre over the point* $p \in B$.

Furthermore, for any point $z \in E$ there has to be an open neighbourhood U of $\pi(z) \in B$ (a *trivialising neighbourhood*) such that $E \supseteq \pi^{-1}(U)$ is diffeomorphic to $U \times F$ via a fibre-preserving diffeomorphism ψ, i. e. such that the following diagram commutes:

$$\begin{array}{ccc} E \supseteq \pi^{-1}(U) & \xrightarrow{\ \psi\ } & U\times F\,, \\ \Big\downarrow{\scriptstyle \pi} & \swarrow{\scriptstyle \mathrm{pr}_1} & \\ U & & \end{array}$$

where $\mathrm{pr}_1 : U\times F \to U\,,\ (u,f)\mapsto u$ is the projection on the first factor (see also figure 2.8 on the facing page).

If a fibre bundle E is *globally* diffeomorphic to the product space $B \times F$, we say that E is a ***trivial bundle***, otherwise E is called a *nontrivial bundle* (examples are given in figure 2.9 on the next page).

Definition 2.5.2 (Fibre chart, local trivialisation, transition functions) The pair (U, ψ) from above is called a ***fibre chart***. A collection (U_α, ψ_α) of fibre charts such that (U_α) is an open cover of B is called a ***fibre bundle atlas*** or ***local trivialisation*** of the bundle.

Given a local trivialisation (U_α, ψ_α) of a fibre bundle E and using the shorthand notation $U_{\alpha\beta} := U_\alpha \cap U_\beta$, we define ***transition functions*** $t_{\alpha\beta}$ by:

$$\psi_\alpha \circ \psi_\beta^{-1}(u,f) =: \big(u, t_{\alpha\beta}(u,f)\big)\,,$$

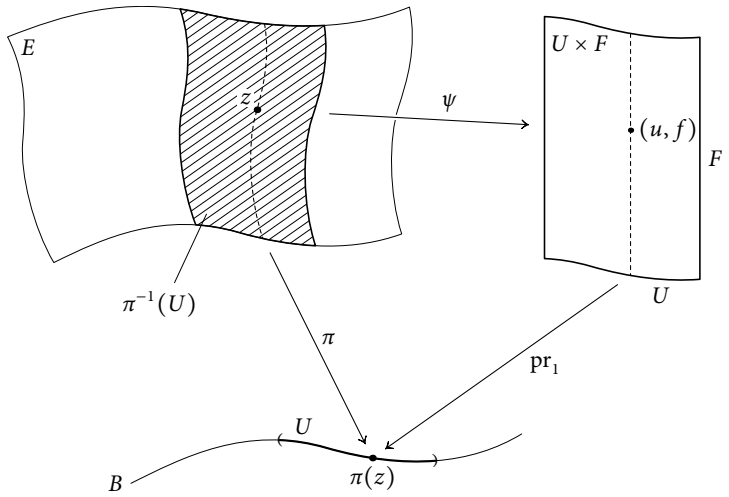


Figure 2.8 A fibre chart (U, ψ) locally trivialises the bundle E.

where $t_{\alpha\beta}$ can be seen either as a map $t_{\alpha\beta} : U_{\alpha\beta} \times F \to F$, $(u, f) \mapsto t_{\alpha\beta}(u, f)$ or (equivalently) as a map $t_{\alpha\beta} : U_{\alpha\beta} \to \mathrm{Diff}(F)$, $u \mapsto t_{\alpha\beta}(u)$, where $t_{\alpha\beta}(u) := t_{\alpha\beta}(u, \cdot)$ is a diffeomorphism of the typical fibre F. The transition functions satisfy the *cocycle conditions* (they form a cocycle in the *Čech cohomology* [Mico8, section 8.3]):

$$\begin{aligned} t_{\alpha\beta}(u) \circ t_{\beta\gamma}(u) &= t_{\alpha\gamma}(u) \qquad \forall u \in U_\alpha \cap U_\beta \cap U_\gamma \,, \\ t_{\alpha\alpha}(u) &= \mathrm{id}_F \qquad\quad \forall u \in U_\alpha \,, \\ t_{\alpha\beta}(u) &= t_{\beta\alpha}(u)^{-1} \quad \forall u \in U_{\alpha\beta} \,. \end{aligned} \tag{2.24}$$

Accordingly, the collection $(t_{\alpha\beta})$ is called a *cocycle of transition functions*.

The definition of vector fields as sections of the tangent bundle can be generalised to the case of fibre bundles. The sections of a fibre bundle take on values in the fibre over

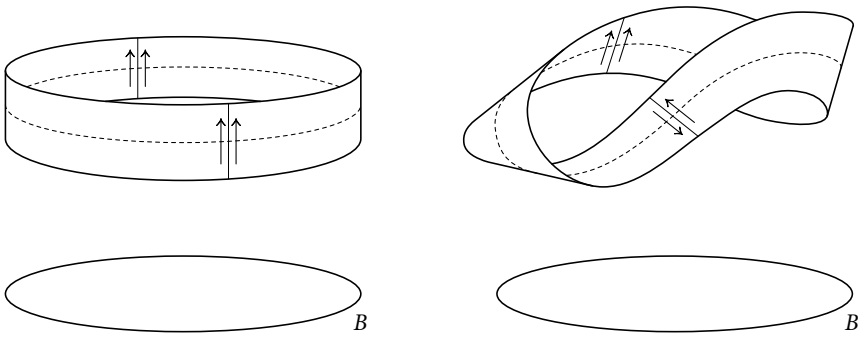


Figure 2.9 Different fibre bundles over the same base manifold B. The cylinder is a *trivial bundle*, the Moebius strip a *nontrivial bundle*. The latter can be constructed by gluing two pieces of paper together where one joint is straight, the other has a half-twist as shown in the figure. The *structure group* of the Moebius strip would therefore be $G \cong \{+1, -1\}$.

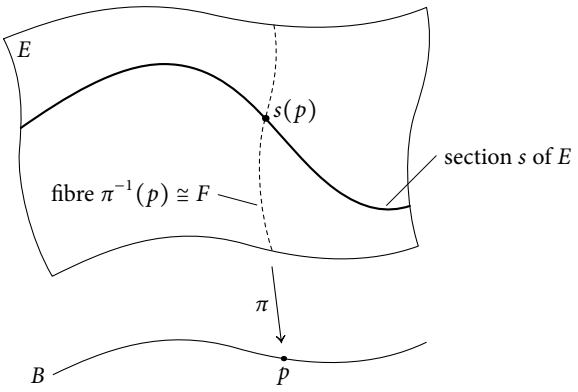


Figure 2.10 Global section of a fibre bundle E.

each point; a kind of 'F-valued fields'.

Definition 2.5.3 (Section) A (*global*) ***section*** (also called ***cross section***) of a fibre bundle $\pi : E \to B$ is a C^k-differentiable map $s : B \to E$ such that $\pi \circ s = \mathrm{id}_B$. The C^k-differentiable sections are denoted by $\Gamma^k(E)$ or $\Gamma^k(B, E)$, *smooth* sections by $\Gamma^\infty(E)$.

In analogy, a ***local section*** is a map $s : U \to E$ with $\pi \circ s = \mathrm{id}_U$, where $U \subseteq B$ is an open subset of B. The C^k-differentiable local sections will be written as $\Gamma^k(U, E)$.

Bundles in general don't have to admit globally defined sections (see for example [Nako3, theorem 9.2]). In contrast, local sections always do exist, supported by the local trivialising neighbourhoods U_α of a fibre bundle atlas. There actually is a connection between the existence of local sections and local trivialisations of the bundle.

Definition 2.5.4 (G-bundle atlas, G-bundle structure) Let (E, B, π, F) be a fibre bundle, then a G-***bundle atlas*** on E consists of:

In gauge theories, the fields are sections of a G-bundle and the structure group is then equal to the *gauge group*.

- a Lie group G (the ***structure group***) together with a *left action* ℓ on the typical fibre, i. e. a smooth mapping $\ell : G \times F \to F$ such that $\ell_{gh} f = \ell_g \ell_h f$ and $\ell_e = \mathrm{id}_F$, where we write $\ell_g f = \ell(g, f)$,
- a fibre bundle atlas (U_α, ψ_α) with a cocycle $(t_{\alpha\beta})$ of transition functions that act on F via the G-action. This means that there exists a family of smooth mappings $\varphi_{\alpha\beta} : U_{\alpha\beta} \to G$ such that $t_{\alpha\beta}(u, f) = \ell(\varphi_{\alpha\beta}(u), f)$, where $\varphi_{\alpha\beta}(u) \in G$.

A G-***bundle structure*** is then an equivalence class of G-bundle atlases, where two G-bundle atlases are considered equivalent if their union is again a G-bundle atlas.

Definition 2.5.5 (G-bundle) A G-***bundle*** is a fibre bundle (E, B, π, F) together with a G-bundle structure on E.

Usually, the G-bundle structure will be given by specifying an explicit fibre bundle atlas together with transition functions. An example is the Moebius strip in figure 2.9, where the open sets U_1 and U_2 correspond to the two pieces of paper, the structure group is $G \cong \mathbb{Z}_2 \cong \{+1, -1\}$ (in this case, as an exception, G is not a Lie group) where $+1$ produces the straight join and -1 does the half-twist mentioned in the description.

As another example, the tangent bundle TM of an n-dimensional manifold M is a $\mathrm{GL}(n, \mathbb{R})$-bundle, where the typical fibre is the vector space $\mathbb{R}^n$.

Definition 2.5.6 (Principal bundle) A ***principal fibre bundle*** (or just *principal bundle*) is a G-bundle $\pi : P \to M$ with typical fibre identical to G, where the structure group G acts on the fibre G (a Lie group) by left translation $G \times G \to G\,,\ \ell_g(h) := gh$.

Remark 2.5.7 (Principal right action) A principal bundle $\pi : P \to M$ admits a unique right action $r : P \times G \to P$ (the *principal right action*), determined by:

$$\psi_\alpha\big(r(\psi_\alpha^{-1}(u,h),g)\big) := (u,hg)\,,$$

where ψ_α are fibre bundle charts.

Definition 2.5.8 (Vector bundle) A ***vector bundle*** is a G-bundle (E,B,π,V) where the standard fibre V is a vector space and the structure group is $G = \mathrm{GL}(V)$, that is the transition functions $t_{\alpha\beta}(u) \in \mathrm{GL}(V)$ are linear maps.

A vector bundle with 1-dimensional fibre is also called ***line bundle***.

Remark 2.5.9 The space of smooth sections $\Gamma^\infty(E)$ of a vector bundle E naturally has the structure of a $C^\infty(M,\mathbb{R})$-module with fibre-wise addition and multiplication induced by the vector space structure of the typical fibre V. Moreover, vector bundles always admit a global section, the ***zero section*** $0_E : B \to E\,,\ p \mapsto (p,0_p)$ where 0_p denotes the zero in the fibre E_p (the zero element is special since the structure group is linear and 0_p is thus *invariant* under the group action). As an example, for the Moebius strip build out of $B = S^1$ and $V = \mathbb{R}$, the zero section is the *only* global section that is admitted.

Definition 2.5.10 (Bundle map) Let $\pi : E \to M$ and $\pi' : E' \to M'$ be fibre bundles. A continuous map $f : E \to E'$ is called ***bundle map*** or ***bundle morphism*** if it is *fibre-preserving*, that is if there exists a continuous map $f_\mathrm{B} : M \to M'$ such that the following diagram commutes:

$$\begin{array}{ccc} E & \xrightarrow{\ f\ } & E' \\ {\scriptstyle\pi}\big\downarrow & & \big\downarrow{\scriptstyle\pi'} \\ M & \xrightarrow[f_\mathrm{B}]{} & M' \end{array}\ .$$

In this case, we say that f is a bundle map *covering* f_B. If f is invertible, we speak of a *bundle isomorphism*. Two bundles over the same manifold M are called *isomorphic* if there exists a bundle isomorphism, with the additional property that this isomorphism covers the identity id_M.

If the bundle in question has some additional structure we want a bundle morphism to preserve this structure as well. For instance in the case of vector bundles we require that f is fibrewise linear, i. e. the restriction $f_p : E_p \to E'_{f_\mathrm{B}(p)}$ is a linear map between the fibres E_p over $p \in M$ and $E'_{f_\mathrm{B}(p)}$ over $f_\mathrm{B}(p) \in M'$.

The pushforward $\varphi_* : TM \to TN$ between tangent bundles is an example of a vector bundle map covering $\varphi : M \to N$.

2.6 Constructions With Fibre Bundles

There are several important constructions that create new fibre bundles out of existing ones. The basic observation is the following *construction theorem*:

Theorem 2.6.1 (Construction theorem for G-bundles) Given a manifold M, let (U_α) be an open cover of M, suppose we have a left G-action ℓ of a Lie group G on some manifold F, and let $(\varphi_{\alpha\beta} : U_{\alpha\beta} \to G)$ be a cocycle of transition functions. Then we can construct a G-bundle, which depends only on the cohomology class of the cocycle.

Proof. A proof is given by Michor in [Mico8, sections 8.3 and 18.1]. The general idea is sketched in [Nako3, section 9.2.2]: Construct a collection of trivial bundles $U_\alpha \times F$ out of the open cover (U_α) and let X be their disjoint union:

$$X = \coprod_\alpha U_\alpha \times F \,.$$

Introduce an equivalence relation $\sim$ on X where $(u, f) \in U_\alpha \times F$ is considered equivalent to $(u', f') \in U_\beta \times F$ if $u = u'$ and $f' = \ell(\varphi_{\beta\alpha}(u), f)$, that is if the base points match and fibre elements transform via the given transition functions. The total space of the G-bundle can then be defined by $E := X/\!\sim$.

If the elements of E (the equivalence classes) are written as $[(u, f)]$, the bundle projection π is then given by:

$$\pi : E \to M \,, \quad [(u, f)] \mapsto u \,,$$

and fibre charts by:

$$\psi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times F \,, \quad [(u, f)] \mapsto (u, f) \,,$$

where f can be taken from a fixed but freely chosen representative (u, f) of the equivalence class $[(u, f)]$. The cocycle conditions together with the equivalence classes ensure that the fibre charts are compatible with each other. ∎

The construction theorem can be applied to vector bundles in the following way (see [Mico8, section 8.8]): Consider the category $\mathrm{Vect}_\mathbb{R}$ of finite-dimensional real vector spaces with linear mappings and let $F : \mathrm{Vect}_\mathbb{R} \to \mathrm{Vect}_\mathbb{R}$ be a covariant functor. The functor is called *smooth* if $F : \mathrm{Hom}(V, W) \to \mathrm{Hom}(F(V), F(W))$ is a smooth mapping of the morphisms.

Now, given a vector bundle $\pi : E \to B$ together with vector bundle atlas (U_α, ψ_α) and a cocycle of transition functions $\varphi_{\alpha\beta} : U_{\alpha\beta} \to \mathrm{GL}(V)$, we can use F to obtain a new cocycle of transition functions:

$$F(\varphi_{\alpha\beta}) : U_{\alpha\beta} \to \mathrm{GL}(F(V)) \,, \quad u \mapsto F(\varphi_{\alpha\beta})(u) := F\big(\varphi_{\alpha\beta}(u)\big) \,,$$

where $F(\varphi_{\alpha\beta})$ satisfies the cocycle condition because F is a covariant functor. Finally, we define $F(E)$ to be the vector bundle over B with fibre $F(V)$ constructed out of this data according to the construction theorem.

Example 2.6.2 Important examples (most of them can be found in Michor's book):

- Let $F = \bigotimes^k$ be the k-th tensor power, which is a smooth functor as is known from linear algebra. Given a vector bundle E with typical fibre V, the resulting bundle is the ***tensor product bundle*** $F(E) = \bigotimes^k E$ with typical fibre $\bigotimes^k V$.
- The k-th ***exterior product bundle*** $\bigwedge^k E$ with typical fibre $\bigwedge^k V$ can be constructed by taking the smooth functor $F = \bigwedge^k$ producing the k-th exterior power.
- The ***dual bundle*** E^* is constructed via the duality functor $F(V) = V^*$. Since F is *contra*variant (not *co*variant) the new cocycle of transition functions has to be defined as $F(\varphi_{\alpha\beta})(u) := F(\varphi_{\alpha\beta}(u)^{-1})$.

- If we consider the k-functor $F(\cdot, \ldots, \cdot) = \cdot \otimes \cdots \otimes \cdot$, which takes k vector spaces $V_1, \ldots, V_k$ and produces a new vector space $V_1 \otimes \cdots \otimes V_k$ (and does the same for linear maps), we can construct the ***tensor product*** of k different vector bundles $E_1, \ldots, E_k$ over the same base manifold. This gives $F(E_1, \ldots, E_k) = E_1 \otimes \cdots \otimes E_k$. The new transition functions are given by $F(\varphi_{1,\alpha\beta}, \ldots, \varphi_{k,\alpha\beta})$.
- The direct sum is another smooth k-functor $F(V_1, \ldots, V_k) = V_1 \oplus \ldots \oplus V_k$, which is covariant in all arguments. The ***direct sum*** of k vector bundles over the same base manifold is then given by $F(E_1, \ldots, E_k) = E_1 \oplus \ldots \oplus E_k$.

The above constructions can be combined, for instance to define the Grassmann algebra bundle which will later provide the foundation for differential forms:

Definition 2.6.3 (Grassmann algebra bundle) Given some vector bundle (E, B, π, V), the ***Grassmann algebra bundle*** over E is:

$$\bigwedge E := \bigoplus_{k=0}^{n} \left(\bigwedge^k E \right),$$

where n is the dimension of the real vector space V and $\bigwedge^0 E := M \times \mathbb{R}$.

Definition 2.6.4 (Cotangent bundle) The ***cotangent bundle*** T^*M of a manifold M is the dual of the tangent bundle, $T^*M := (TM)^*$.

Remark 2.6.5 (Explicit construction of the cotangent bundle) We want to give a direct construction of the cotangent bundle. For this, let $\pi : TM \to M$ be the tangent bundle of some manifold M and let (U_α, ψ_α) be a fibre bundle atlas of TM. Since the fibres T_pM are vector spaces they can be dualised. Like in definition 2.2.6 of the tangent bundle we take the disjoint union to obtain the total space:

The elements of the dualised space T_p^*M are linear functions ω_p that map tangent vectors to real numbers, $\omega_p : T_pM \to \mathbb{R}$.

$$T^*M := \coprod_{p \in M} T_p^*M = \bigcup_{p \in M} \{p\} \times T_p^*M \,.$$

The bundle maps of the tangent bundle satisfy the following diagram:

$$\begin{array}{ccccc} TM \supseteq TU_\alpha & \xrightarrow{\psi_\alpha} & U_\alpha \times V & & \\ \downarrow{\scriptstyle \pi} & \swarrow{\scriptstyle \mathrm{pr}_1} & & \searrow{\scriptstyle \mathrm{pr}_2} & \\ M \supseteq U_\alpha & & & & V \end{array},$$

where $V = \mathbb{R}^n$ is the typical fibre. Now we can transpose the map $\mathrm{pr}_2 \circ \psi_\alpha : T_pM \to V$ to first obtain a map between the dual spaces $(\mathrm{pr}_2 \circ \psi_\alpha)|_p^{\mathrm{T}} : V^* \to T_p^*M$ and then define the fibre bundle charts of T^*M via the inverse:

$$\psi_\alpha^* : T^*U_\alpha \to U_\alpha \times V^* \,, \quad z \mapsto \left(\pi(z), \left((\mathrm{pr}_2 \circ \psi_\alpha)|_{\pi(z)}^{\mathrm{T}} \right)^{-1}(z) \right),$$

where $\pi : T^*M \to M$ arises naturally from the definition as disjoint union.

2.7 Differential Forms

Apart from differential equations we need a way how to integrate on manifolds. The problem with the conventional integral is that it has coordinate dependencies both in

the integrand $f(x)$ and in the measure $\mathrm{d}x$ independently:

$$\int f(x)\,\mathrm{d}x\,.$$

The coordinate invariant approach to integration is based on *differential forms*. We will see – in corollary 2.7.7 to be specific – that they somehow combine $f(x)$ and $\mathrm{d}x$ into a single, coordinate invariant entity. A straightforward definition is the following:

Definition 2.7.1 (Differential form) Let M be a smooth manifold. A ***differential form*** ω of degree k on M (or k-***form*** for short) is a section $\omega \in \Gamma^\infty(\bigwedge^k T^*M)$ of the exterior product bundle $\bigwedge^k T^*M$ of the cotangent space over M.

We write $\Omega^k(M)$ for the space of differential k-forms on M. By convention, 0-forms are identified with functions $\Omega^0(M) = C^\infty(M,\mathbb{R})$. The ***degree*** of a differential form is denoted by $\deg\omega = k$ for $\omega \in \Omega^k(M)$.

Yet another approach is to define alternating k-forms on vector spaces and then introduce differential k-forms analogous to the construction of vector fields from tangent vectors (see e. g. [Jäno1b]).

Remark 2.7.2 (Alternative definition) A differential k-form ω can alternatively be seen as a $C^\infty(M,\mathbb{R})$-*multilinear* smooth mapping:

$$\omega : \mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M) \to C^\infty(M)\,, \quad (X_1,\ldots,X_k) \mapsto \omega(X_1,\ldots,X_k)\,,$$

which is *skew symmetric*, i. e.

$$\omega(X_{\sigma(1)},\ldots,X_{\sigma(k)}) = \operatorname{sign}(\sigma)\,\omega(X_1,\ldots,X_k)\,,$$

for any permutation σ in the symmetric group $\mathbb{S}_k$.

There is a natural way to multiply differential forms:

Definition 2.7.3 (Wedge product) The ***wedge product*** (also called ***exterior product***) of two differential forms $\omega \in \Omega^k(M)$ and $\eta \in \Omega^l(M)$ is a mapping:

$$\wedge : \Omega^k(M) \times \Omega^l(M) \to \Omega^{k+l}(M)\,, \quad (\omega,\eta) \mapsto \omega\wedge\eta\,,$$

$\omega\wedge\eta$ is essentially the complete antisymmetrisation of the tensor product $\omega\otimes\eta$.

where $\omega\wedge\eta$ is defined in a coordinate-free manner as:

$$\begin{aligned}&\omega\wedge\eta(X_1,\ldots,X_k,X_{k+1},\ldots,X_{k+l})\\ &\quad := \frac{1}{k!l!}\sum_{\sigma\in\mathbb{S}_{k+l}} \operatorname{sign}(\sigma)\,\omega(X_{\sigma(1)},\ldots,X_{\sigma(k)})\eta(X_{\sigma(k+1)},\ldots,X_{\sigma(k+l)})\,,\end{aligned} \tag{2.25}$$

for vector fields $X_1,\ldots,X_{k+l} \in \mathfrak{X}(M)$.

Lemma 2.7.4 (Properties of the wedge product) The wedge product has the following properties:

(i) it is *associative*: $(\omega\wedge\eta)\wedge\zeta = \omega\wedge(\eta\wedge\zeta)$,

(ii) it is *graded commutative*: $\omega\wedge\eta = (-1)^{\deg\omega\deg\eta}\eta\wedge\omega$,

(iii) it is $C^\infty(M,\mathbb{R})$-*bilinear*: $(f\omega)\wedge g = \omega\wedge(f\eta) = f(\omega\wedge\eta)$,

(iv) it fulfils $1_M\wedge\omega = \omega$,

where ω, η and ζ are differential forms, f a smooth function on M and 1_M is the constant function $1_M : M \to \mathbb{R}\,,\ p \mapsto 1$.

Proof. The first two properties and the last one follow easily from the definition. The third property is a direct consequence of (i) and (ii), since functions are identified with 0-forms, which means we can write $f\omega = f \wedge \omega$, and since the degree of f is $\deg f = 0$. ∎

Theorem 2.7.5 (Algebra of differential forms) The space of differential forms:

$$\Omega(M) := \bigoplus_{k=0}^{\infty} \Gamma^\infty\big(\textstyle\bigwedge^k T^*M\big) \cong \Gamma^\infty\Big(\bigoplus_{k=0}^{\infty} \big(\textstyle\bigwedge^k T^*M\big)\Big) ,$$

(sections of the *Grassmann algebra bundle* over T^*M) together with the wedge product is an associative, supercommutative algebra with unity.

The direct sum in the formula is actually *finite* because differential k-forms of degree $k > \dim M$ vanish, due to the skew symmetry.

Proof. We already mentioned that sections over a vector bundle with fibrewise addition and multiplication form a $C^\infty(M, \mathbb{R})$-module. The remaining algebra properties follow from lemma 2.7.4. Supercommutative thereby simply means that the wedge product is $\mathbb{Z}_2$-graded. ∎

Lemma 2.7.6 (Basis of T_p^*M) Let (U, x) be a chart of M around the point $p \in M$. Then the differentials $(\mathrm{d}x_p^1, \ldots, \mathrm{d}x_p^n)$ of the coordinate functions $x^i : U \to \mathbb{R}$ form a basis of the cotangent space T_p^*M. This basis is dual to the basis $(\partial_1|_p, \ldots, \partial_n|_p)$ of the tangent space T_pM at this point.

Moreover, a basis of the space $\bigwedge^k T^*M$ is given by the exterior products:

$$\mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} , \quad \text{with } 1 \le i_1 < \ldots < i_k \le n .$$

The dimension of $\bigwedge^k T^*M$ is $\dim \bigwedge^k T^*M = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Differentials $\mathrm{d}\varphi_p$ were introduced in definition 2.3.1; there we used the name *tangent map* and wrote $\mathrm{d}\varphi_p = T_p\varphi$.

Proof. Given a chart $x : U \to \mathbb{R}^n$, we have the coordinate functions $x^i : U \to \mathbb{R}$, which are maps between manifolds (since $\mathbb{R}$ is a manifold over itself). The differential $\mathrm{d}x_p^i$ is then a mapping $\mathrm{d}x_p^i : T_pM \to T_{x(p)}\mathbb{R} \cong \mathbb{R}$ and thus can be seen as an element of the dual space T_p^*M. Furthermore, for tangent vectors $\partial_j|_p \in T_pM$ equation (2.7) yields:

$$\mathrm{d}x_p^i(\partial_j|_p) = \partial_j|_p(x_p^i) = \delta_j^i ,$$

hence the differentials $(\mathrm{d}x_p^1, \ldots, \mathrm{d}x^n)$ form a basis dual to the basis $(\partial_1|_p, \ldots, \partial_n|_p)$ of the tangent space T_pM.

It follows from theorem 2.7.5 that the products $\mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k}$ span $\bigwedge^k T^*M$. Two such products are linearly dependent if and only if they differ by a permutation of the factors. To obtain a basis we therefore select a *specific* permutation with the aid of the rule $1 \le i_1 < \ldots < i_k \le n$. ∎

Corollary 2.7.7 (Coordinate expression of differential form) Given a chart (U, x) of M, a differential k-form $\omega \in \Omega^k(M)$ can locally be written as:

$$\omega = \sum_{1 \le i_1 < \ldots < i_k \le n} \omega_{i_1 \ldots i_k} \, \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} , \quad \text{with } \omega_{i_1 \ldots i_k} = \omega(\partial_{i_1}, \ldots, \partial_{i_k}) ,$$

where $\omega_{i_1 \ldots i_k}$ is a function in $C^\infty(M, \mathbb{R})$ and totally antisymmetric in the indices. Alternatively, we can use the Einstein summation convention and write:

$$\omega = \frac{1}{k!} \omega_{i_1 \ldots i_k} \, \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} ,$$

where the factor $1/k!$ makes up for the $k!$ permutations of the indices $i_1, \ldots, i_k$ that appear in the sum if we don't use the restriction from above.

In contrast to the *pushforward* (definition 2.4.5), where φ had to be a diffeomorphism, the *pullback* is more well-behaved and φ needs only to be smooth.

Definition 2.7.8 (Pullback) Let $\varphi : M \to N$ be a smooth mapping and $\omega \in \Omega^k(N)$ a differential form. The ***pullback*** $\varphi^*\omega \in \Omega^k(M)$ is defined pointwise by:

$$(\varphi^*\omega)_p(v_1, \ldots, v_k) := \omega_{\varphi(p)}(\varphi_{*,p}v_1, \ldots, \varphi_{*,p}v_k) \,, \tag{2.26}$$

on tangent vectors $v_i \in T_pM$, or globally by:

$$\varphi^*\omega(X_1, \ldots, X_k) = \omega(\varphi_\# X_1, \ldots, \varphi_\# X_k) \circ \varphi \,, \tag{2.27}$$

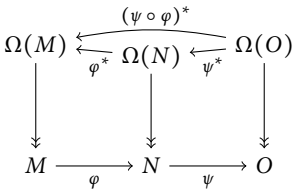
for vector fields $X_i \in \mathfrak{X}(M)$. The mapping $\varphi^* : \Omega^k(N) \to \Omega^k(M)$ goes in the direction opposite to the one of φ, so that the following diagram commutes:

$$\begin{array}{ccc} \Omega^k(M) & \xleftarrow{\varphi^*} & \Omega^k(N) \\ \Big\downarrow{\scriptstyle \pi_M} & & \Big\downarrow{\scriptstyle \pi_N} \\ M & \xrightarrow[\varphi]{} & N \end{array} \,.$$

Lemma 2.7.9 (Naturality of the pullback) The pullback is *natural* in the sense that it is an algebra homomorphism of $\Omega(M)$. In particular, it fulfils:

$$\varphi^*(\omega \wedge \eta) = (\varphi^*\omega) \wedge (\varphi^*\eta) \,. \tag{2.28}$$

Proof. By looking at the defining equation (2.25) of the wedge product it is easily seen that it doesn't matter if we first take the product and then apply the pullback or vice versa. ∎



Lemma 2.7.10 (Chain rule) Given smooth maps $\varphi : M \to N$ and $\psi : N \to O$ between manifolds, we have the *chain rule* (note the order!):

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^* \,. \tag{2.29}$$

Furthermore:

$$\mathrm{id}^*_M = \mathrm{id}_{\Omega(M)} \,. \tag{2.30}$$

Accordingly, the mapping $M \mapsto \Omega(M)$ and $\varphi \mapsto \varphi^*$ is a *contravariant functor* from the category Man^k of manifolds with differentiable maps to the category of real, graded commutative algebras with algebra homomorphisms.

Proof. The chain rule for the pullback follows from the definition by using the chain rule for the pushforward (lemma 2.3.4). ∎

Definition 2.7.11 (Insertion operator) The ***insertion operator*** i_X 'inserts' a vector field $X \in \mathfrak{X}(M)$ into a differential form $\omega \in \Omega^k(M)$ according to:

$$\mathrm{i}_X\omega(Y_1, \ldots, Y_{k-1}) := \omega(X, Y_1, \ldots, Y_{k-1}) \,.$$

We set $\mathrm{i}_X f = 0$ for any function/0-form $f \in \Omega^0(M)$.

The general definition of a graded derivation is given in definition 2.7.13, below.

Lemma 2.7.12 (Properties of the insertion operator) The insertion operator is a *graded derivation of degree* -1 of the algebra $\Omega(M)$, meaning:

$$\mathrm{i}_X(\omega \wedge \eta) = (\mathrm{i}_X\omega) \wedge \eta + (-1)^{-\deg\omega}\omega \wedge (\mathrm{i}_X\eta) \,. \tag{2.31}$$

Moreover:

$$\mathrm{i}_X \circ \mathrm{i}_Y + \mathrm{i}_Y \circ \mathrm{i}_X = 0 \,. \tag{2.32}$$

Proof. The first property can be checked by using the definition of the wedge product and some combinatorial considerations (details can be found in [Mico8, lemma 9.7]). The second property follows directly from the skew symmetry of differential forms, since $\omega(X, Y, \ldots) = -\omega(Y, X, \ldots)$. ∎

Definition 2.7.13 (Graded algebra, graded derivation) Let G be a monoid (a semigroup with identity, e. g. $(\mathbb{Z}, +)$). An algebra $\mathcal{A}$ is called a G-***graded algebra*** if there exists a decomposition into a direct sum of submodules $\mathcal{A}_g \subseteq \mathcal{A}$:

$$\mathcal{A} = \bigoplus_{g \in G} \mathcal{A}_g \,,$$

such that $\mathcal{A}_g \mathcal{A}_h \subseteq \mathcal{A}_{gh}$. A ***superalgebra*** is simply a $\mathbb{Z}_2$-graded algebra.

Given a $\mathbb{Z}$-graded algebra $\mathcal{A}$, a ***graded derivation*** of degree d is a linear map D with $D|_{\mathcal{A}_k} : \mathcal{A}_k \to \mathcal{A}_{k+d}$ such that the following graded product rule holds:

$$D(ab) = D(a)b + (-1)^{d \deg a} a D(b) \,.$$

A graded derivation is sometimes also called ***superderivation***.

2.8 Exterior Derivative and de Rham Cohomology

We already know that, given a function $f : M \to \mathbb{R}$, the differential $\mathrm{d}f_p$ at the point $p \in M$ is an element of the cotangent space $T_p^* M$. Since $\mathrm{d}f : M \to T^*M$, $p \mapsto \mathrm{d}f_p$ is smooth, the map $\mathrm{d}f$ defines a 1-form on M. Accordingly, the symbol d can be seen as an operator $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$ which takes 0-forms and produces 1-forms. Furthermore, there is a natural extension that yields an operator acting on k-forms:

Theorem 2.8.1 (Exterior derivative) There is a unique extension of the differential $\mathrm{d} : \Omega^0(M) \to \Omega^1(M)$ as a *superderivation of degree* +1 with respect to the wedge product such that the *complex property* $\mathrm{d} \circ \mathrm{d} = 0$ holds.

The name 'complex property' comes from *chain complexes*, which are defined below.

In detail, this means that there exists a family of mappings $\mathrm{d}_k : \Omega^k(M) \to \Omega^{k+1}(M)$ with the following properties:

(i) $\mathrm{d}_k : \Omega^k(M) \to \Omega^{k+1}(M)$ is $\mathbb{R}$-linear,

(ii) we have the graded product rule:

$$\mathrm{d}(\omega \wedge \eta) = (\mathrm{d}\omega) \wedge \eta + (-1)^{\deg \omega} \omega \wedge (\mathrm{d}\eta) \,, \tag{2.33}$$

(iii) the complex property $\mathrm{d}_{k+1} \circ \mathrm{d}_k = 0$ holds.

Proof. A complete proof can be found in [Jäno1b, sections 8.3–8.5]. As Waldmann points out [Walo7, Satz 2.3.14], the main problem is to show that d_k can be 'localised' on open subsets $U \subseteq M$. Afterwards, given the coordinate expression of a k-form $\omega \in \Omega^k(M)$ in some chart (U, x):

$$\omega|_U = \frac{1}{k!} \omega_{i_1 \ldots i_k} \, \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} \,,$$

the map d_k can be defined locally by the action on ω as:

$$\mathrm{d}_k \omega|_U := \frac{1}{k!} \, \mathrm{d}\omega_{i_1 \ldots i_k} \wedge \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_k} \,, \tag{2.34}$$

where $\omega_{i_1 \ldots i_k} : M \to \mathbb{R}$ is a normal function. ∎

Definition 2.8.2 (Exterior derivative) The ***exterior derivative*** $\mathrm{d} : \Omega(M) \to \Omega(M)$ is a derivation of the algebra of differential forms, where the action on k-forms is given by the map $\mathrm{d}|_{\Omega^k(M)} := \mathrm{d}_k : \Omega^k(M) \to \Omega^{k+1}(M)$ from above.

An alternative motivation for this formula can be found in [Mico8, section 9.8], where Michor argues why it is the most natural choice.

Lemma 2.8.3 (Explicit expression) If $\omega \in \Omega^k(M)$ is a differential form, the exterior derivative assumes the explicit form:

$$\begin{aligned}\mathrm{d}\omega(X_0, \ldots, X_k) = &\sum_{i=0}^{k} (-1)^k X_i\big(\omega(X_0, \ldots, \check{X}_i, \ldots, X_k)\big) \\ &+ \sum_{i<j} (-1)^{i+j} \omega\big([X_i, X_j], X_0, \ldots, \check{X}_i, \ldots, \check{X}_j, \ldots, X_k\big) ,\end{aligned} \tag{2.35}$$

for vector fields $X_0, \ldots, X_k \in \mathfrak{X}(M)$, where the notation $\check{X}_i$ means that the vector field X_i is omitted.

Proof. Because of the universal property of exterior product spaces, a multilinear map $A : \bigwedge^k V \to \mathbb{R}$ is uniquely determined by a family of k linear maps $A_i : V \to \mathbb{R}$. As a result, two multilinear maps are equal if they agree on a set of basis vectors. Thus, to check equation (2.35) we only have to calculate both sides using basis vector fields ∂_i. More details can be found in [Walo7, Satz 2.3.15]. ∎

Lemma 2.8.4 (Naturality of d) The exterior derivative is *natural* in the sense that, if $\varphi : M \to N$ is a smooth map between manifolds, the following diagram commutes:

$$\begin{array}{ccc} \Omega^k(N) & \xrightarrow{\mathrm{d}_N} & \Omega^{k+1}(N) \\ \Big\downarrow{\scriptstyle \varphi^*} & & \Big\downarrow{\scriptstyle \varphi^*} \\ \Omega^k(M) & \xrightarrow{\mathrm{d}_M} & \Omega^{k+1}(M) \end{array} , \tag{2.36}$$

for each k, that is if:

$$\varphi^* \circ \mathrm{d}_N = \mathrm{d}_M \circ \varphi^* , \tag{2.37}$$

where d_N and d_M are the exterior derivatives on N and M, respectively.

To prove this, we first cite a useful lemma [Walo7, lemma 2.3.19]:

Lemma 2.8.5 (Derivations and generating sets) Let $\mathcal{A}$ be a $\mathbb{Z}$-graded associative algebra and let D_1 and D_2 be graded derivations of degree $d_i = \deg D_i$. Then the graded commutator:

$$[\![D_1, D_2]\!] := D_1 \circ D_2 - (-1)^{d_1 d_2} D_2 \circ D_1 ,$$

is a graded derivation of degree $d_1 + d_2$.

If D is a graded derivation of $\mathcal{A}$, then it is uniquely determined by the values $D(a_i)$ evaluated on a set $\{a_i\}$ generating $\mathcal{A}$ (by addition and algebra multiplication).

Proof. We can see that $[\![D_1, D_2]\!]$ is a linear map $\mathcal{A}_k \to \mathcal{A}_{k+d_1+d_2}$ for each k. Checking the graded product rule is then done by an explicit but not very enlightening calculation.

The second part follows because any element $a \in \mathcal{A}$ can be written as a combination of sums and products of the generators a_i. By applying the product rule and linearity, $D(a)$ can be reduced to an expression which contains only sums and products of the values $D(a_i)$. These values thus completely determine D. ∎

Proof of lemma 2.8.4. It follows from lemma 2.7.6 that $\Omega(N)$ is generated by functions $f \in C^\infty(N, \mathbb{R})$ and 1-forms of the form $\mathrm{d}f \in \Omega^1(N)$. Since $\varphi^* \circ \mathrm{d}_N$ and $\mathrm{d}_M \circ \varphi^*$ are both graded derivations of degree +1 we only have to check equality on such a set of generators. For functions $f \in C^\infty(N, \mathbb{R})$ we obtain:

$$\begin{aligned}\big(\varphi^*(\mathrm{d}_N f)\big)_p(X) &= (\mathrm{d}_N f)_{\varphi(p)}(\varphi_* X_p) = (\varphi_* X_p)(f) = X_p(f \circ \varphi) = X_p(\varphi^* f) \\ &= \big(\mathrm{d}_M(\varphi^* f)\big)_p(X)\,.\end{aligned}$$

The calculation for a 1-form of the form $\mathrm{d}f = \mathrm{d}_N f$ is even easier:

$$\mathrm{d}_M(\varphi^*(\mathrm{d}_N f)) = \mathrm{d}_M \circ \mathrm{d}_M(\varphi^* f) = 0 = \varphi^*(\mathrm{d}_N \circ \mathrm{d}_N f)\,.$$

By means of the last lemma, equation (2.37) thus holds on the whole algebra. ∎

In definiton 2.4.12 we declared the Lie derivative of a vector field. The idea was that the Lie derivative $\mathcal{L}_X$ measures the *change* along the flow of a vector field X. It should be obvious how this can be translated to differential forms:

Definition 2.8.6 (Lie derivative of a differential form) Let M be a smooth manifold, let X be a vector field on M and Φ its flow. The ***Lie derivative*** $\mathcal{L}_X\omega$ ***of*** $\omega \in \Omega^k(M)$ ***along*** X is then given by:

$$(\mathcal{L}_X\omega)_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\Phi_t^*\omega)_p = \lim_{t\to 0} \frac{1}{t}\Big(\Phi_t^*(\omega_{\Phi_t(p)}) - \omega_p\Big)\,. \tag{2.38}$$

Lemma 2.8.7 (Properties of the Lie derivative) The Lie derivative of differential forms has the following properties:

(i) $\mathcal{L}_X(\omega \wedge \eta) = (\mathcal{L}_X\omega) \wedge \eta + \omega \wedge (\mathcal{L}_X\eta)$, (*derivation* of degree 0)

(ii) $\mathcal{L}_X = \mathrm{i}_X \circ \mathrm{d} + \mathrm{d} \circ \mathrm{i}_X$, (*Cartan's magic formula*)

(iii) $\mathcal{L}_X \circ \mathrm{d} = \mathrm{d} \circ \mathcal{L}_X$, (*naturalilty* with respect to d)

(iv) $\mathcal{L}_X \circ \mathrm{i}_Y - \mathrm{i}_Y \circ \mathcal{L}_X = \mathrm{i}_{[X,Y]}$.

Proof. (i) The first property follows from the fact that the pullback is natural with respect to the wedge product, as stated in lemma 2.7.9. Therefore:

$$\begin{aligned}\mathcal{L}_X(\omega \wedge \eta) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^*(\omega \wedge \eta) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\Phi_t^*\omega) \wedge (\Phi_t^*\eta) \\ &= \Big(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^*\omega\Big) \wedge \eta + \omega \wedge \Big(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^*\eta\Big) = (\mathcal{L}_X\omega) \wedge \eta + \omega \wedge (\mathcal{L}_X\eta)\,,\end{aligned}$$

where we used $\Phi_0^* = \mathrm{id}$ when applying the product rule.

(ii) Now that we know that $\mathcal{L}_X$ is a derivation of differential forms we can again use lemma 2.8.5 to reduce the proof to two simple calculations. For functions $f \in C^\infty(M, \mathbb{R})$ we have $\mathrm{i}_X f = 0$, so:

$$(\mathrm{i}_X \circ \mathrm{d} + \mathrm{d} \circ \mathrm{i}_X)f = \mathrm{i}_X(\mathrm{d}f) = \mathrm{d}f(X) = Xf = \mathcal{L}_X f\,.$$

For a differential 1-form $\mathrm{d}f \in \Omega^1(M)$ we use $\mathrm{d} \circ \mathrm{d} = 0$ to obtain:

$$\begin{aligned}(\mathrm{i}_X \circ \mathrm{d} + \mathrm{d} \circ \mathrm{i}_X)\mathrm{d}f &= \mathrm{d}(\mathrm{i}_X \mathrm{d}f) = \mathrm{d}(\mathrm{d}f(X)) = \mathrm{d}(Xf) \\ &= \mathrm{d} \circ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^* f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{d} \circ \Phi_t^* f = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \Phi_t^* \circ \mathrm{d}f = \mathcal{L}_X(\mathrm{d}f)\,,\end{aligned}$$

where the naturality of the exterior derivative (lemma 2.8.4) allowed us to interchange the order of Φ_t^* and d. Thus Cartan's magic formula holds.

(iii) This is a simple corollary of Cartan's formula:

$$\mathcal{L}_X \circ \mathrm{d} = (\mathrm{i}_X \circ \mathrm{d} + \mathrm{d} \circ \mathrm{i}_X) \circ \mathrm{d} = (\mathrm{i}_X \circ \mathrm{d}^2 + \mathrm{d} \circ \mathrm{i}_X \circ \mathrm{d}) + \mathrm{d}^2 \circ \mathrm{i}_X = \mathrm{d} \circ \mathcal{L}_X \,,$$

where we added an additional $0 = \mathrm{d}^2 \circ \mathrm{i}_X$ in the second step.

(iv) The last property is again checked in two steps. For functions both sides vanish, because i_X applied to 0-forms yields 0. For a 1-form $\mathrm{d}f \in \Omega^1(M)$ we get:

$$\begin{aligned}(\mathcal{L}_X \circ \mathrm{i}_Y - \mathrm{i}_Y \circ \mathcal{L}_X)\mathrm{d}f &= \mathcal{L}_X\big(\mathrm{d}f(Y)\big) - \mathrm{i}_Y(\mathcal{L}_X \mathrm{d}f) = \mathcal{L}_X\big(\mathrm{d}f(Y)\big) - (\mathcal{L}_X \mathrm{d}f)(Y) \\ &= \mathcal{L}_X\big(\mathrm{d}f(Y)\big) - \mathcal{L}_X\big(\mathrm{d}f(Y)\big) + \mathrm{d}f(\mathcal{L}_X Y) = \mathrm{d}f\big([X,Y]\big) = \mathrm{i}_{[X,Y]}\mathrm{d}f \,.\end{aligned}$$

Lemma 2.8.5 again does the rest for us, because each side is a derivation. ∎

An interesting question – in fact one of the foundations of *cohomology* theory and *homological algebra* – is the following: Given a differential form ω, is there a differential form η such that ω can be written as $\omega = \mathrm{d}\eta$? Before we give some answers to this question, we first need to introduce a few new words:

Definition 2.8.8 (Closed and exact forms) A differential form $\omega \in \Omega^k(M)$ is called ***closed*** if its exterior derivative vanishes, $\mathrm{d}\omega = 0$. On the other hand, ω is called ***exact*** if there exists $\eta \in \Omega^{k-1}(M)$ such that $\mathrm{d}\eta = \omega$. The set of closed forms will be denoted by $Z^k(M)$, the set of exact forms by $B^k(M)$.

Given $\mathrm{d}^2 = 0$, it is clear that an exact form fulfils $\mathrm{d}\omega = \mathrm{d}^2\eta = 0$. A necessary condition for a form to be exact is thus that it is closed. The *Poincaré lemma* tells us, if some k-form $\omega \in \Omega^k(M)$ is closed then it is at least *locally* exact: For any point p in M there is an open neighbourhood U of p and some locally defined form $\eta \in \Omega^{k-1}(U)$ such that $\omega|_U = \mathrm{d}\eta$ (see e. g. [Mico8, lemma 9.10] or [Walo7, Satz 2.3.25]).

To go further, it is useful to introduce the concepts of chain and cochain complexes as well as homology and cohomology classes. Here we take a little broader approach, since we not only need *de Rham cohomology*, which is the cohomology of differential forms, but later, in section 5.3, we will also need to talk about *Chevalley–Eilenberg cohomology*, the cohomology of Lie algebras.

Definition 2.8.9 (Chain complex) A ***chain complex*** $A_\bullet = (A_k, \partial_k)$ over some unital ring Λ is a sequence of Λ-modules A_k, connected by morphisms $\partial_k : A_k \to A_{k-1}$, called ***boundary operators***, such that the *complex property* $\partial_k \circ \partial_{k+1} = 0$ holds. They are usually written in the form:

The complex property is equivalent to $\operatorname{im}\partial_{k+1} \subseteq \ker\partial_k$. Compare this to *exact sequences*, where the stronger condition $\operatorname{im}\partial_{k+1} = \ker\partial_k$ holds.

$$\cdots \longrightarrow A_{k+1} \xrightarrow{\partial_{k+1}} A_k \xrightarrow{\partial_k} A_{k-1} \xrightarrow{\partial_{k-1}} A_{k-2} \longrightarrow \cdots \,.$$

A ***cochain complex*** $A^\bullet = (A^k, d^k)$ is defined analogously, only the direction of the morphisms $d^k : A^k \to A^{k+1}$ is reversed and the complex property changes to $d^k \circ d^{k-1} = 0$:

$$\cdots \longrightarrow A^{k-1} \xrightarrow{d^{k-1}} A^k \xrightarrow{d^k} A^{k+1} \xrightarrow{d^{k+1}} A^{k+2} \longrightarrow \cdots \,.$$

The map d^k is then called a ***coboundary operator***.

Definition 2.8.10 (de Rham complex) Let M be a manifold, $\dim M = n$. The cochain complex $(\Omega^k(M), \mathrm{d})$, consisting of the vector spaces $\Omega^k(M)$ of differential k-forms together with the exterior derivative, is called the ***de Rham complex*** of M:

$$0 \longrightarrow \Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^{n-1}(M) \xrightarrow{\mathrm{d}} \Omega^n(M) \xrightarrow{\mathrm{d}} 0 \ .$$

Remark 2.8.11 (Pullback as chain map) The naturality from lemma 2.8.4 implies that the pullback φ^* is a ***chain map*** between de Rham complexes:

$$\varphi^* : (\Omega^k(M), \mathrm{d}) \to (\Omega^k(N), \mathrm{d}) \ ,$$

i. e. a family of morphisms between the two cochain complexes such that the following diagram commutes:

$$\begin{array}{ccccccccccc}
\cdots & \xrightarrow{\mathrm{d}} & \Omega^{k-1}(N) & \xrightarrow{\mathrm{d}} & \Omega^k(N) & \xrightarrow{\mathrm{d}} & \Omega^{k+1}(N) & \xrightarrow{\mathrm{d}} & \Omega^{k+2}(N) & \xrightarrow{\mathrm{d}} & \cdots \\
 & & \downarrow{\scriptstyle \varphi^*} & & \downarrow{\scriptstyle \varphi^*} & & \downarrow{\scriptstyle \varphi^*} & & \downarrow{\scriptstyle \varphi^*} & & \\
\cdots & \xrightarrow{\mathrm{d}} & \Omega^{k-1}(M) & \xrightarrow{\mathrm{d}} & \Omega^k(M) & \xrightarrow{\mathrm{d}} & \Omega^{k+1}(M) & \xrightarrow{\mathrm{d}} & \Omega^{k+2}(M) & \xrightarrow{\mathrm{d}} & \cdots \ .
\end{array}$$

The next thing is to see how the definition of closed and exact forms translates into the language of chain complexes:

Definition 2.8.12 (Chains, cycles, boundaries, homology and cohomology classes) Let $A_\bullet = (A_k, \partial_k)$ be a chain complex with boundary operator ∂. Define:

$$C_k(A_\bullet) := A_k \ , \quad Z_k(A_\bullet) := \ker \partial_k \ , \quad B_k(A_\bullet) := \operatorname{im} \partial_{k+1} \ . \tag{2.39}$$

Elements of $C_k(A_\bullet)$ are called k-***chains***; elements of $Z_k(A_\bullet)$ are k-***cycles***; elements of $B_k(A_\bullet)$ are called k-***boundaries***.

By definition, $Z_k(A_\bullet)$ and $B_k(A_\bullet)$ are submodules of A_k with $B_k(A_\bullet) \subseteq Z_k(A_\bullet)$. Thus it makes sense to define:

$$H_k(A_\bullet) := Z_k(A_\bullet)/B_k(A_\bullet) = \ker \partial_k / \operatorname{im} \partial_{k+1} \ , \tag{2.40}$$

which is another Λ-module, called the k-***th homology class*** $H_k(A_\bullet)$ of $A_\bullet$.

The same definitions can be made for a cochain complex $A^\bullet = (A^k, d^k)$, leading to k-***cochains*** $C^k(A^\bullet)$, k-***cocycles*** $Z^k(A^\bullet)$ and k-***coboundaries*** $B^k(A^\bullet)$. The ***cohomology classes*** $H^k(A^\bullet)$ are given by:

$$H^k(A^\bullet) := Z^k(A^\bullet)/B^k(A^\bullet) = \ker d^k / \operatorname{im} d^{k-1} \ . \tag{2.41}$$

Remark 2.8.13 As elaborated on in [HS71, section IV.1], the k-th homology can be seen as a covariant functor H_k from the category of chain complexes with chain maps as morphisms, to the category of left modules with module homomorphisms. The k-th *co*homology is then a contravariant functor H^k.

Using this new language, we can see that a closed form $\omega \in \Omega^k(M)$ is a *cocycle* of the de Rham complex, because $\mathrm{d}\omega = 0$ is equivalent to $\omega \in \ker \mathrm{d}$. Moreover, ω is exact if and only if it is of the form $\omega = \mathrm{d}\eta$ for some $(k-1)$-form η, which is equivalent to $\omega \in \operatorname{im} \mathrm{d}$. Hence, exact forms are *coboundaries*.

Definition 2.8.14 (de Rham cohomology) The k-th ***de Rham cohomology class*** is the k-th cohomology class of the de Rham cochain complex:

$$H^k_{\mathrm{dR}}(M) = Z^k(M)/B^k(M) \,.$$

The elements of $H^k_{\mathrm{dR}}(M)$ are equivalence classes of closed forms, where two of them are considered equivalent if the differ only by an exact form: $\omega \sim \eta \;\Leftrightarrow\; \omega = \eta + \mathrm{d}\zeta$. All the cohomology classes together form a graded algebra, called the ***de Rham cohomology algebra*** of M:

$$H^\bullet_{\mathrm{dR}}(M) = \bigoplus_{k=0}^{\dim M} H^k_{\mathrm{dR}}(M) \,.$$

We can now given an answer to the question from above:

Lemma 2.8.15 (Exactness of closed forms) A closed k-form $\omega \in Z^k(M)$ is exact if and only if it is equivalent to $[0] \in H^k_{\mathrm{dR}}(M)$. Moreover, if $H^k_{\mathrm{dR}}(M)$ is trivial, any closed k-form is also exact.

Proof. Let $\omega \sim \eta$ be two equivalent k-forms where η is exact, i. e. there exists some $\alpha \in \Omega^{k-1}(M)$ such that $\eta = \mathrm{d}\alpha$. Then ω is also exact, since equivalency means that there exists some ζ such that $\omega = \eta + \mathrm{d}\zeta$ and thereby $\omega = \eta + \mathrm{d}\zeta = \mathrm{d}\alpha + \mathrm{d}\zeta = \mathrm{d}(\alpha + \zeta)$. Now, the space of k-forms $\Omega^k(M)$ always contains a zero-form 0_k which is closed and exact, given that $\mathrm{d}0_k = 0_{k+1}$ and $0_k = \mathrm{d}0_{k-1}$. Thus we have shown that $\omega \in Z^k(M)$ is exact if and only if it is equivalent to 0_k. ∎

Another approach to calculate cohomology classes is via the *Mayer–Vietoris sequence* (see [Mico8, section 11.10] or [HS71, section VI.14] for details).

Of course, we still have to find a way to actually *calculate* the de Rham cohomology classes. This can be done with the aid of de Rham's theorem [Nako3, theorem 6.2], which tells us that for compact manifolds the cohomology classes are isomorphic to certain homology classes. The basic observation comes from Stokes' theorem:

Theorem 2.8.16 (Stokes) Let G be an oriented, k-dimensional manifold with border ∂G and let ω be a $k-1$-form compactly supported on G. Then we have:

$$\int_G \mathrm{d}\omega = \int_{\partial G} \omega \,. \tag{2.42}$$

Proof. See for example [Jäno1b, chapter 9] or [Nako3, theorem 6.1]. ∎

Since the boundary of a boundary of some manifold always vanishes, we have the complex property $\partial \circ \partial = 0$, and, in fact, it is possible to define a chain complex based on the notion of geometric boundaries. To allow for actual computations, this has to be formalised a little bit. One possibility is to use a *simplicial complex* homeomorphic to the manifold M we are interested in, and define ∂ as an operator acting on such a simplicial complex in a purely algebraic way. Then it is possible to calculate *homology classes* $H_k(M)$ associated to the manifold M. This approach is taken for example in [Nako3, chapter 3].

Stokes' theorem then can be interpreted as representing some *duality* between the homology complex calculated using simplicial complexes, and the de Rham complex of M. More concretely, if we define a pairing $\langle G, \omega \rangle$ of manifolds and differential forms:

$$\langle G, \omega \rangle := \int_G \omega \,,$$

equation (2.42) translates to:

$$\int_G \mathrm{d}\omega = \int_{\partial G} \omega \;\Leftrightarrow\; \langle G, \mathrm{d}\omega \rangle = \langle \partial G, \omega \rangle \ .$$

Hence, the boundary operator ∂ and the exterior derivative d are kind of *adjoint* to each other. This idea then leads to de Rham's theorem [Nako3, theorem 6.2]:

Theorem 2.8.17 (de Rham) If M is a compact manifold, then the homology classes $H_k(M)$ and the cohomology classes $H^k_{\mathrm{dR}}(M)$ are both finite-dimensional. Moreover, the map $H_k(M) \times H^k_{\mathrm{dR}}(M) \to \mathbb{R}$ (induced by the pairing $\langle \cdot, \cdot \rangle$ from above) is bilinear and non-degenerate. Hence, homology and cohomology classes are dual to each other:

$$H^k_{\mathrm{dR}}(M) = \big(H_k(M)\big)^* \ .$$

As a corollary, the question if some closed form on M is exact can be answered by calculating the homology classes of an associated simplicial complex, which is sometimes much easier to do. Probably an even more astounding result is that the answer depends only on the *topological structure* of the manifold M.

3 Groups and Group Actions

3.1 Basics of Group Actions

Definition 3.1.1 (Topological group) A ***topological group*** G is simultaneously a group and a topological space such that the group multiplication $\mu : G \times G \to G$ and the inversion operation $\mathrm{inv} : G \to G$ are continuous.

A ***morphism*** between topological groups G and H is a continuous group homomorphism $G \to H$. An ***isomorphism*** is a group homomorphism $G \to H$ which is also a homeomorphism of the underlying topological spaces.

One can think of a topological group as a *group object* [Lan02, section 1.11] in the category of topological spaces.

Definition 3.1.2 (Continuous group action) Let G be a topological group and let X be a topological space. A ***left action*** of G on X is then a continuous mapping $\ell : G \times X \to X$ such that (if we write $\ell_g = \ell(g, \cdot)$):

- $\ell_e = \mathrm{id}_X$ for the neutral element $e \in G$,
- $\ell_g \circ \ell_h = \ell_{gh}$ for all $g, h \in G$.

A ***right action*** of G on X is a continuous mapping $r : X \times G \to X$ such that:

- $r_e = \mathrm{id}_X$ for the neutral element $e \in G$,
- $r_g \circ r_h = r_{hg}$ for all $g, h \in G$. (Note the reverse order in the group product!)

It is common to use $G \times X$ as the domain for a left action and $X \times G$ for a right action, yet the *important* difference is the multiplication rule $\ell_g \circ \ell_h = \ell_{gh}$ or $r_g \circ r_h = r_{hg}$.

Remark 3.1.3 Any *left* action ℓ can be turned into a *right* action by taking the inverse of the group elements. In detail, if we consider $r(g, x) := \ell(g^{-1}, x)$, then:

$$r_g \circ r_h = \ell_{g^{-1}} \circ \ell_{h^{-1}} = \ell_{g^{-1}h^{-1}} = \ell_{(hg)^{-1}} = r_{hg} \,,$$

which is the multiplication rule for a right action. The same can be done to obtain a left action from a given right action.

Remark 3.1.4 (Mappings induced by a group action) Any group action $\rho : G \times X \to X$ induces four additional mappings:

$$\begin{aligned} &\rho_g : X \to X \,,\ x \mapsto \rho_g(x) = \rho(g, x) \,, \qquad &&\hat{\rho} : G \to \mathrm{Aut}(X) \,,\ g \mapsto \rho_g \,, \\ &\rho^x : G \to X \,,\ g \mapsto \rho^x(g) = \rho(g, x) \,, \qquad &&\tilde{\rho} : X \to C(G, X) \,,\ x \mapsto \rho^x \,. \end{aligned}$$

The mapping $\hat{\rho} : G \to \mathrm{Aut}(X)$ is called a ***realisation*** of the group G. The mappings $\rho_g : X \to X$ are all *homeomorphisms*, since they are continuous and invertible with continuous inverse. The homeomorphisms are the *automorphisms* in the category of topological spaces. This explains why we write $\mathrm{Aut}(X)$ for the codomain of $\hat{\rho}$.

Using the realisation, there is an alternative way how to characterise group actions. The idea is that, due to the category laws, the automorphisms $\mathrm{Aut}(X)$ of an object X in some category $\mathcal{C}$ form a group. The definition from above can then be rephrased – and generalised – in the following way:

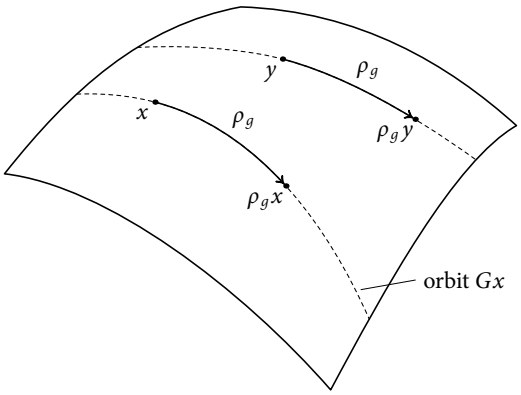


Figure 3.1 Group G acting on some space X. The *orbit* Gx is the set of all points that can be reached via the group action when starting from x.

Definition 3.1.5 (Group action, alternative version) Let G be a group and X be an object in some category $\mathcal{C}$. A ***left action*** of G on X is then a mapping $\ell : G \times X \to X$ such that the realisation $\hat{\ell} : G \to \mathrm{Aut}_{\mathcal{C}}(X)$ is a *group homomorphism*. This implies:

$\mathrm{Aut}_{\mathcal{C}}(X)$ here denotes the automorphism group of X, where *automorphism* is meant in the sense of the category $\mathcal{C}$.

$$\hat{\ell}(gh) = \hat{\ell}(g) \circ \hat{\ell}(h) \,, \quad \hat{\ell}(e) = \mathrm{id}_X \,.$$

A ***right action*** of G on X is a mapping $r : X \times G \to X$ such that the realisation is a *group* anti*homomorphism*, i. e.

$$\hat{r}(gh) = \hat{r}(h) \circ \hat{r}(g) \,, \quad \hat{r}(e) = \mathrm{id}_X \,.$$

Remark 3.1.6 (Group representation) Let V be a vector space. Then the automorphism group $\mathrm{Aut}(V)$ – in the category of vector spaces – is the set of invertible linear maps from V to V, better known as the *general linear group* $\mathrm{GL}(V) = \mathrm{Aut}(V)$. In this case group elements $g \in G$ are represented by linear transformations $\rho_g \in \mathrm{GL}(V)$ and $\hat{\rho}$ is called (*linear*) ***representation*** of the group G.

Remark 3.1.7 (Group action as functor) Yet another way to look at a group actions is as functors. In example 2.3.6 we already mentioned that a group can be seen as a category with a single object and where all the morphisms are isomorphisms. A *left action* is then a *covariant functor* from the group G seen as category, into another category $\mathcal{C}$. The single object in G is mapped to an object X in $\mathcal{C}$, and the morphisms in G (the group elements) are mapped to automorphisms of X. The functor properties ensure that we get a left group action. Similarly, a *right action* can be seen as a *contravariant functor*.

Definition 3.1.8 (Orbit, stabiliser) Let X be a topological space with group action of G on X denoted by ρ. The ***orbit*** Gx of a point $x \in X$ is the set:

$$Gx := \{\rho_g x : g \in G\} \subseteq X \,, \tag{3.1}$$

in other words: the orbit Gx consists of all points in X that can be reached via the group action when starting from x.

The ***stabiliser*** G_x of the point $x \in X$ (also called ***stabiliser subgroup***, ***isotropy group*** or ***little group*** of x) is the subgroup:

$$G_x := \{g \in G : \rho_g x = x\} \leq G \,, \tag{3.2}$$

i. e. the stabiliser G_x is the subgroup of G that leaves x invariant (so x is a *fixed point*).

Definition 3.1.9 (Orbit space) Orbits define an *equivalence relation* on X, where two points $x, y \in X$ are considered equivalent if they belong to the same orbit:

$$x \sim_G y :\Leftrightarrow \exists g \in G : x = \rho_g y \Leftrightarrow Gx = Gy .$$

The set of all orbits (the equivalence classes), endowed with the quotient topology, is called ***orbit space*** X/G.

Even if X is a manifold, the orbit space X/G in general doesn't have to be one. It is, however, the case if G is a Lie group with a *free* and *proper* action on X [Walo7, Satz 3.3.18].

Group actions are usually classified according to the following criteria:

Definition 3.1.10 (Types of group actions) A group action $\rho : G \times X \to X$ is called:

- ***transitive*** if there is only one orbit $Gx = X$ (for any choice of $x \in X$). In other words: for any two points $x, y \in X$ there exists some $g \in G$ such that $y = \rho_g x$.
- ***effective*** (or ***faithful***) if $\rho_g x = x$ for *all* $x \in X$ (this is $\rho_g = \mathrm{id}_X$) implies $g = e$. We can also say that $\hat{\rho}$ is injective (a *monomorphism*), i. e. $\ker \hat{\rho} = \{e\}$. Informally speaking, different elements of G then induce different transformations of X.
- ***almost effective*** if there exists a small neighbourhood U of the neutral element $e \in G$ such that the action restricted to U is effective, meaning such that $\hat{\rho}|_U$ is injective.
- ***free*** (or ***semiregular***) if $\rho_g x = x$ for any choice of $x \in X$ implies $g = e$. In this case, each orbit Gx is isomorphic to the group $Gx \cong G$. By definition, a *free* action is also *effective*, but the reverse is not true.

  The isomorphism $Gx \cong G$ for a free action is a result of the *orbit–stabiliser theorem*, presented below.
- ***regular*** (or ***simply/sharply transitive***) if the action is *free* and *transitive*. This means that for any two points $x, y \in X$ there exists precisely one $g \in G$ such that $y = \rho_g x$. For a regular action we get $X = Gx \cong G$, and the space X equipped with the group action is then known as a ***principal homogeneous space*** or as a G-***torsor***.

3.2 G-spaces and Equivariant Maps

Definition 3.2.1 (G-space, equivariant map) A **left/right** G-***space*** is a topological space X together with a continuous left/right action $\rho : G \times X \to X$.

A mapping $F : X \to Y$ between two G-spaces X and Y is called ***equivariant*** if it respects the group actions, i. e. if the following diagram commutes for all $g \in G$:

An equivariant mapping $F : X \to Y$ maps all points of the orbit $Gx \subseteq X$ to the orbit $GF(x) \subseteq Y$.

$$\begin{array}{ccc} X & \xrightarrow{F} & Y \\ {\scriptstyle \rho_g}\downarrow & & \downarrow{\scriptstyle \varphi_g} \\ X & \xrightarrow{F} & Y \end{array} ,$$

where ρ and φ are the group actions on X and Y, respectively. This is equivalent to the following condition:

$$F \circ \rho_g = \varphi_g \circ F \quad \forall g \in G .$$

Equivariant maps are the *morphisms* in the *category of G-spaces*. Accordingly, an ***isomorphism of*** G-***spaces*** is an equivariant *homeomorphism* between G-spaces – not just an equivariant, continuous bijection.

An important result is the following theorem:

Note that G/G_x is not a group unless G_x is a *normal* subgroup. This is why the isomorphism in question can not be a group homomorphism.

Theorem 3.2.2 (Orbit–stabiliser theorem) Let X be a left G-space and $x \in X$, then:

$$G/G_x \cong_G Gx\,, \tag{3.3}$$

where the isomorphism is an equivariant continuous bijection.

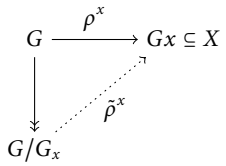
Intuitively, the stabiliser subgroup G_x can be thought of as the set of group elements that 'don't move' x. The theorem then says that the orbit of a point $x \in X$ is isomorphic to the set G/G_x of group elements that actually 'do something' to the point x.

Proof. Let X be a left G-space with action denoted by ρ. For fixed $x \in X$ we have the induced mapping $\rho^x : G \to X$, $\rho^x(g) := \rho_g x$ which can be used to define an equivalence relation of group elements: $g \sim h :\Leftrightarrow \rho^x(g) = \rho^x(h)$. The equivalence classes defined this way are the same as the equivalence classes in G/G_x, since:

$$g \sim h \;\Leftrightarrow\; \rho^x(g) = \rho^x(h) \;\Leftrightarrow\; \rho_g x = \rho_h x \;\Leftrightarrow\; \rho_{h^{-1}g} x = x \;\Leftrightarrow\; h^{-1}g \in G_x\,.$$

Thus we obtain $G/\!\sim\; = G/G_x$.

The following argument is essentially the *first isomorphism theorem* of groups translated to the category of sets, where the kernel of $\tilde\rho^x$ can be defined as $\ker \tilde\rho^x = [0] = G_x$.



Given that the image of ρ^x is $\operatorname{im}\rho^x = Gx$, the next step is to show $G/\!\sim\; \cong \operatorname{im}\rho^x$. To this end, consider the map:

$$\tilde\rho^x : G/\!\sim\; \to X\,, \quad \tilde\rho^x([g]) := \rho^x(g) = \rho_g x\,,$$

which is *well-defined*, since $[g] = [g']$ is equivalent to $\rho_x(g) = \rho_x(g')$ and therefore also $\tilde\rho^x([g]) = \tilde\rho^x([g'])$. Furthermore, the map $\tilde\rho^x$ is *injective*, because $\tilde\rho^x([g]) = \tilde\rho^x([g'])$ is defined as $\rho^x(g) = \rho^x(g')$ which implies $[g] = [g']$. Finally, the image of $\tilde\rho^x$ is the same as the image of ρ^x, hence we have our bijection:

$$G/G_x = G/\!\sim\; \cong \operatorname{im}\tilde\rho^x = \operatorname{im}\rho^x = Gx\,.$$

By the universal property of the quotient space G/G_x the bijection is also continuous.

It remains to show that the bijection is *equivariant*. First note that G/G_x naturally inherits a group action $\varphi : G \times G/G_x \to G/G_x$, $\varphi_g[h] = [gh]$, whereas on Gx we have the group action of X restricted to the orbit. Using this we can check the equivariance property:

$$\tilde\rho^x \circ \varphi_g([h]) = \tilde\rho^x([gh]) = \rho_{gh}x = \rho_g(\rho_h x) = \rho_g(\rho^x(h)) = \rho_g \circ \tilde\rho^x([h])\,,$$

where we used that ρ is a *left* action. Accordingly, the bijection $\tilde\rho^x$ is equivariant and thus we have shown $G/G_x \cong_G Gx$. ∎

Remark 3.2.3 If G is a compact group and the space X is Hausdorff it can be shown that the mapping $\tilde\rho^x$ is not just a continuous bijection but also that the inverse of $\tilde\rho^x$ is continuous. In this case, the map $\tilde\rho^x$ is an equivariant homeomorphism, i. e. an *isomorphism of G-spaces*. A proof is given in [Jän01a, section 3.5].

Groupoids were introduced in example 2.3.6. A nice survey is given by Brown [Bro87].

Before we continue, we just want to say a few words about *action groupoids* associated to G-spaces. Although we don't need action groupoids in the following, they provide an interesting view on G-spaces from a very different perspective.

Remark 3.2.4 (Action groupoid) A given G-space X can be represented by its associated *action groupoid* $G \ltimes X$. The idea is the following:

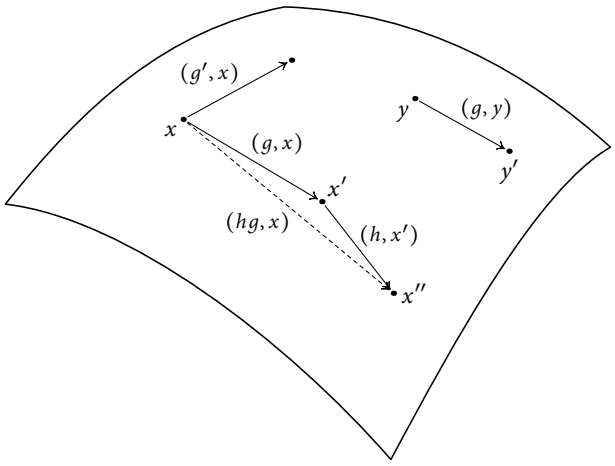


Figure 3.2 Group action pictured as a collection of arrows on X. Two arrows (g, x) and (h, x') can be composed if $\rho_g x = x'$.

Picture a group action ρ of G on X as a collection of arrows, where an arrow taking some point $x \in X$ to the point $\rho_g x$ will be denoted by $(g, x) : x \mapsto \rho_g x$, as in figure 3.2. We will call the point x the *source* of the arrow (g, x), while the point $\rho_g x$ will be its *target*. The group action then induces a composition of arrows – but only if the target of the first arrow coincides with the source of the second:

$$(h, x') \circ (g, x) := (hg, x) \,, \quad \text{if } \rho_g x = x' \,.$$

This is obviously just placing arrows tail to head after one another. Since arrows have inverses and since there is an 'identity arrow' $\mathrm{id}_x = (e, x)$ for each $x \in X$, it should be clear that we just obtained a groupoid.

To summarise, the *action groupoid* $G \ltimes X$ associated to some G-space X is the category where the set of objects is taken to be X, and the morphisms are the arrows indexed by the set $G \times X$ (this means we identify the set of arrows with the set $G \times X$; an arrow hence is exactly a tuple $(g, x) \in G \times X$). The composition of the morphisms/arrows is already defined above.

In applications, action groupoids can be used, for example, to better describe the symmetry properties of a system when different types of orbits have different symmetry groups. Moreover, the groupoid view on G-spaces yields some interesting calculational methods, where techniques from group theory can be generalised to work on groupoids. Such an example where the concept of a groupoid leads to a deeper understanding of the underlying structure is the *fundamental groupoid* $\Pi_1(X)$ of a space X, which generalises the *fundamental group* $\pi_1(X, x_0)$ in such a way that $\Pi_1(X)$ no longer depends on a specific base point $x_0 \in X$.

3.3 Lie Groups and Lie Algebras

Like topological groups were group objects in the category of topological spaces, Lie groups are group objects in the category of smooth manifolds. There is an extensive literature on the topic of Lie groups and Lie algebras. Here, we just want to mention the

book of Marsden and Ratiu [MR99], which focuses on applications in classical mechanics, as well as Fulton and Harris [FH91] and Simon [Sim96], where the representation theory of Lie groups is explained.

Definition 3.3.1 (Lie group) A ***Lie group*** G is simultaneously a group and a smooth manifold such that the group multiplication $\mu : G \times G \to G$ and the inversion operation $\mathrm{inv} : G \to G$ are smooth.

The group actions of ***left translation*** $L_g : G \to G$ and ***right translation*** $R_g : G \to G$ by an element $g \in G$ are defined as $L_g h := gh$ and $R_g h := hg$, respectively. Both are regular actions of the Lie group G on itself.

Interesting properties of Lie groups arise from the interplay of group structure and differential structure. For example it is possible to define vector fields on a Lie group G. An important case are the *left invariant vector fields* which are in some way compatible with the group structure of G.

Definition 3.3.2 (Left invariant vector field) A vector field $X \in \mathfrak{X}(G)$ on a Lie group G is called ***left invariant*** if the following diagram commutes for all $g \in G$:

$$\begin{array}{ccc} TG & \xrightarrow{TL_g} & TG \\ {\scriptstyle X}\uparrow & & \uparrow{\scriptstyle X} \\ G & \xrightarrow{L_g} & G \end{array} .$$

Pointwise we thus have the following relation for the tangent vectors:

$$(T_h L_g) X_h = X_{gh} \quad \forall g, h \in G . \tag{3.4}$$

Equation (3.5) says that a left invariant vector field is L_g-related to itself (see definition 2.4.5).

Using the pullback of vector fields this equation can be written as:

$$L_g{}^{\#} X = X , \tag{3.5}$$

with $(L_g{}^{\#} X)_q = (L_{g\,\#}^{-1} X)_q = TL_{g^{-1}}(X_{L_g q})$, according to definition 2.4.6.

Similarly, one defines ***right invariant*** vector fields by the property $(T_h R_g) X_h = X_{hg}$. We will denote the set of left invariant vector fields on G by $\mathfrak{X}_{\mathrm{L}}(G)$, the set of right invariant vector fields by $\mathfrak{X}_{\mathrm{R}}(G)$.

Since the tangent map TL_g is linear, the left invariant vector fields form a real vector space. Furthermore, due to lemma 2.4.17, the Jacobi–Lie bracket $[X, Y]$ of two left invariant vector fields $X, Y \in \mathfrak{X}_{\mathrm{L}}(G)$ is again left invariant. The space $\mathfrak{X}_{\mathrm{L}}(G)$ thus possesses the structure of a Lie algebra; the Lie algebra associated to G:

Definition 3.3.3 (Lie algebra of G) Let G be a Lie group. The ***Lie algebra*** $\mathfrak{L}G$ of G is the vector space of left invariant vector fields $\mathfrak{X}_{\mathrm{L}}(G)$ together with the Jacobi–Lie bracket (definition 2.4.13):

$$[X, Y](f) := X\big(Y(f)\big) - Y\big(X(f)\big) ,$$

for all $X, Y \in \mathfrak{X}_{\mathrm{L}}(G)$.

Alternatively, we could have defined the Lie algebra $\mathfrak{L}G$ via *right* invariant vector fields. Although both possibilities lead to Lie algebras, the results are not identical. The relation is given by the following lemma:

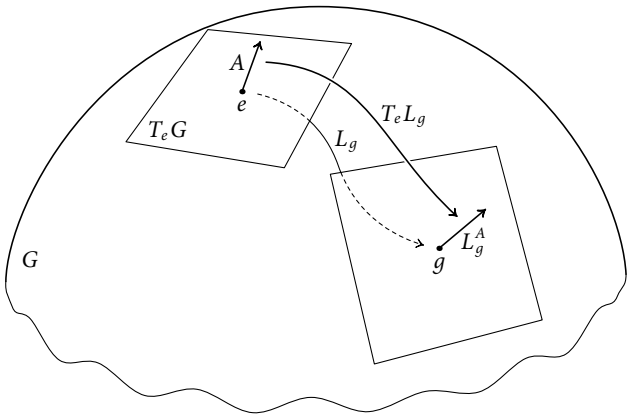


Figure 3.3 A left invariant vector field L^A is generated by a tangent vector $A \in T_eG$ via the tangent map of the left translation.

Lemma 3.3.4 (Relation between left and right invariant vector fields) Let G be a Lie group and write $\mathrm{inv} : G \to G$ for the inversion operation. Then:

$$\mathrm{inv}^{\#} : \mathfrak{X}_{\mathrm{L}}(M) \to \mathfrak{X}_{\mathrm{R}}(M)$$

is an isomorphism of Lie algebras.

Proof. Given a left invariant vector field $X \in \mathfrak{X}_{\mathrm{L}}(G)$, we first show that $Y := \mathrm{inv}^{\#} X$ is right invariant. This is due to:

$$R_g{}^{\#} Y = R_g{}^{\#} \mathrm{inv}^{\#} X = (\mathrm{inv} \circ R_g)^{\#} X = (L_{g^{-1}} \circ \mathrm{inv})^{\#} X = \mathrm{inv}^{\#} L_{g^{-1}}{}^{\#} X = \mathrm{inv}^{\#} X = Y \,,$$

where we used the relation $\mathrm{inv} \circ R_g(h) = (hg)^{-1} = g^{-1}h^{-1} = L_{g^{-1}} \circ \mathrm{inv}(h)$ for $h \in G$, and the fact that X is left invariant. Given that inv is an *involution* of G, i. e. inv is its own inverse, it immediately follows that $\mathrm{inv}^{\#}$ is an isomorphism of vector spaces.

That $\mathrm{inv}^{\#}$ maps the Jacobi–Lie bracket of $\mathfrak{X}_{\mathrm{L}}(G)$ to the one of $\mathfrak{X}_{\mathrm{R}}(G)$ then follows from lemma 2.4.17, since inv is by definition a diffeomorphism. ∎

A closer look at equation (3.4) reveals that a left invariant vector field X is completely determined by its tangent vector at the identity $X_e \in T_eG$, as we can obtain other tangent vectors according to $X_g = X_{ge} = (T_eL_g)X_e$. There actually is an isomorphism between the space of left invariant vector fields $\mathfrak{X}_{\mathrm{L}}(G)$ and the tangent space T_eG (see figure 3.3):

$$T_eG \to \mathfrak{X}_{\mathrm{L}}(G) \,, \quad A \mapsto L^A \quad \text{where } L^A_g := T_eL_g(A) \,. \tag{3.6}$$

This gives us an alternative way to describe the Lie algebra $\mathfrak{L}G$:

Theorem 3.3.5 ($\mathfrak{L}G$ and T_eG are isomorphic) Define a Lie bracket of tangent vectors $A, B \in T_eG$ as the Jacobi–Lie bracket of the generated left invariant vector fields, evaluated at the point $e \in G$:

Due to this theorem we will identify $\mathfrak{L}G$ with T_eG in the following text.

$$[A, B]_{T_eG} := [L^A, L^B](e) \,, \quad A, B \in T_eG \,. \tag{3.7}$$

Then T_eG together with the bracket is isomorphic to the Lie algebra $\mathfrak{L}G$.

Proof. From the discussion above it is clear that T_eG and $\mathcal{L}G$ are isomorphic as vector spaces; the isomorphism is given in equation (3.6). The definition of the Lie bracket on T_eG is chosen such that:

$$L_g^{[A,B]_{T_eG}} = T_eL_g([A,B]_{T_eG}) = T_eL_g\big([L^A,L^B](e)\big) = [L^A,L^B](g) \quad \forall g \in G\,,$$

where we used that the Jacobi–Lie bracket of two left invariant vector fields is again left invariant. Thus, $A \mapsto L^A$ is an isomorphism of Lie algebras. ∎

Lemma 3.3.6 (Left and right bracket of T_eG) Denote by $[\,\cdot\,,\,\cdot\,]^R_{T_eG}$ the alternative Lie bracket defined via *right* invariant vector fields:

$$[A,B]^R_{T_eG} := [R^A,R^B](e)\,.$$

Then we have:

$$[A,B]_{T_eG} = [L^A,L^B](e) = -[R^A,R^B](e) = -[A,B]^R_{T_eG}\,, \tag{3.8}$$

so $-\mathrm{id} : T_eG \to T_eG$ is a Lie algebra isomorphism between T_eG equipped with the Lie bracket defined via left and right invariant vector fields, respectively.

Proof. This is a simple consequence of lemma 3.3.4, since $L^A_e = A = R^A_e$ and hence $\mathrm{inv}^{\#}\,L^A = R^A$. Furthermore, the tangent map of the inversion is given by $T_e\,\mathrm{inv} = -\mathrm{id}$, which follows from:

$$(\mathrm{inv}^{\#}X)_e = T_e\,\mathrm{inv}(X_e) = T_e\,\mathrm{inv}([\alpha]) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big(\alpha(t)\big)^{-1} = -\dot\alpha(0) = -[\alpha] = -X_e\,,$$

for any vector field $X \in \mathfrak{X}_{\mathrm{L}}(G)$, where α is a curve that represents the tangent vector X_e of X at the identity. ∎

If we forget the group structure of G, the Lie functor $\mathcal{L}$ is just the tangent functor T_e at the point $e \in G$.

Definition 3.3.7 (Lie functor) The so-called ***Lie functor*** $\mathcal{L}$ is the covariant functor from the category of Lie groups to the category of Lie algebras that maps a Lie group G to its Lie algebra $\mathcal{L}G$ and a morphism $\varphi : G \to H$ between Lie groups to a Lie algebra morphism $\mathcal{L}\varphi := T_e\varphi : \mathcal{L}G \to \mathcal{L}H$.

Although it is now possible to calculate the Lie bracket using left invariant vector fields, we will usually take another approach via the *adjoint action* $\mathrm{ad} : T_eG \times T_eG \to T_eG$ of the Lie algebra on itself. As explained in [FH91, section 8.1], the basic observation is that the inner automorphism $I_g : G \to G\,,\ h \mapsto ghg^{-1}$ of conjugation by $g \in G$ is a left action of G on itself, and that it has a fixed point at $e \in G$. It is therefore possible to induce an action of G on its Lie algebra with the help of the tangent functor:

Note that I_g is a group *automorphism* for each $g \in G$. The induced left action therefore 'respects' the group structure, whereas the left translation L_g doesn't have this property.

Definition 3.3.8 (Adjoint action of G on T_eG) Let I_g be the inner automorphism of conjugation by $g \in G$:

$$I_g : G \to G\,, \quad h \mapsto I_g(h) = ghg^{-1}\,. \tag{3.9}$$

Since $I_g = L_g \circ R_{g^{-1}} = R_{g^{-1}} \circ L_g$ we can write the adjoint action as $\mathrm{Ad}_g = T_eI_g = T_eL_g \circ T_eR_{g^{-1}}$.

The ***adjoint action*** (or ***adjoint representation***) Ad of the group G on its Lie algebra T_eG is given by:

$$\mathrm{Ad}_g := T_eI_g : T_eG \to T_eG\,. \tag{3.10}$$

To induce an action of the Lie algebra T_eG on itself we consider Ad as a mapping $\mathrm{Ad} : G \to \mathrm{Aut}(T_eG)$, $g \mapsto \mathrm{Ad}_g$. Using that $T_{\mathrm{id}}\mathrm{Aut}(T_eG)$ is isomorphic to the space of endomorphisms $\mathrm{End}(T_eG)$ we arrive at the following definition:

Definition 3.3.9 (Adjoint action of T_eG on T_eG) Let $\mathrm{Ad} : G \to \mathrm{Aut}(T_eG)$ be the adjoint action of G on T_eG. The ***adjoint action*** (or ***adjoint representation***) ad of the Lie algebra T_eG on itself is:

$$\mathrm{ad} := T_e\mathrm{Ad} : T_eG \to \mathrm{End}(T_eG) \,. \tag{3.11}$$

The surprising result is that the adjoint action $\mathrm{ad}(A)(B)$ is nothing other than the Lie bracket $[A, B]_{T_eG}$ calculated using the left invariant vector fields:

Theorem 3.3.10 (Adjoint action and Lie bracket) The Lie bracket on T_eG is:

$$[A, B]_{T_eG} = \mathrm{ad}(A)(B) \,, \quad \text{for } A, B \in T_eG \,. \tag{3.12}$$

Proof. Proofs are given in [MR99, section 9.1] and [Mico8, lemma 4.24]. ∎

If we want to calculate the Lie bracket via the adjoint action it is easiest if we express the tangent vectors $A, B \in T_eG$ as equivalence classes of smooth curves $A = [\alpha]$ and $B = [\beta]$, that is $\alpha(0) = \beta(0) = e$ and $\dot\alpha(0) = A$ as well as $\dot\beta(0) = B$. The tangent map of I_g is then given by:

$$\mathrm{Ad}_g([\beta]) = T_eI_g([\beta]) = [I_g \circ \beta] = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} I_g\big(\beta(t)\big) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} g\beta(t)g^{-1} \,.$$

This yields the explicit expression for the Lie bracket:

$$[A, B]_{T_eG} = \mathrm{ad}(A)(B) = T_e\mathrm{Ad}([\alpha])([\beta]) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \alpha(s)\beta(t)\alpha(s)^{-1} \,. \tag{3.13}$$

It follows immediately that the Lie bracket of $\mathfrak{L}G$ vanishes if the Lie group G is Abelian.

3.4 The Exponential Map and One-Parameter Subgroups

In the last section we found a way how to calculate the Lie algebra $\mathfrak{L}G$ of a given Lie group G. The exponential map, in turn, will show that all continuous one-parameter subgroups of G are generated by elements of the Lie algebra. This way it is possible to *reconstruct* the connected component of e in G; actually, the Lie algebra $\mathfrak{L}G$ *uniquely determines* the connected component of the identity up to an isomorphism (see [FH91, chapter 8] or [Mico8, remark 4.19]). This gives us a hint why the Lie algebra is such an important tool in the study of Lie groups.

Definition 3.4.1 (One-parameter subgroup) Let G be a Lie group. A ***one-parameter subgroup*** of G is a Lie group morphism $\alpha : (\mathbb{R}, +) \to G$, i. e. a smooth curve α on G with $\alpha(t + s) = \alpha(t)\alpha(s)$ for all $t, s \in \mathbb{R}$.

It is already sufficient that α is a *continuous* curve on G, since any continuous morphism of finite-dimensional Lie groups is automatically smooth [MR99, theorem 9.1.9].

The next lemma establishes a correspondence between left invariant vector fields on G and one-parameter subgroups:

Lemma 3.4.2 (Left invariant vector fields and subgroups of G) Let G be a Lie group. Then $\alpha : \mathbb{R} \to G$ is a one-parameter subgroup of G if and only if α is the integral curve through $e \in G$ of the left invariant vector field L^A generated by $A = \dot{\alpha}(0) \in T_eG$.

Proof. Our proof follows the one of [Mico8, lemma 4.17]:

'$\Rightarrow$': For the forward direction let α be a one-parameter subgroup of G and let L^A be the left invariant vector field generated by $A = \dot{\alpha}(0)$. We first show $g\alpha(t) = \Phi(t, g)$, where Φ is the flow of L^A. Since α is a one-parameter subgroup we get:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} g\alpha(t) &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} g\alpha(t+s) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} g\alpha(t)\alpha(s) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} L_{g\alpha(t)}\alpha(s) \\ &= T_eL_{g\alpha(t)}\Big(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\alpha(s)\Big) = T_eL_{g\alpha(t)}(A) = L^A_{g\alpha(t)}\,,\end{aligned}$$

where the last equality is just the definition of L^A from equation (3.6). On the other hand, the flow Φ of L^A fulfils the equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, g) = L^A_{\Phi(t,g)}\,, \quad \Phi(0, g) = g\,.$$

Since $\Phi(t, g)$ and $g\alpha(t)$ satisfy the same differential equation and since they are subject to the same initial condition $g\alpha(0) = ge = g = \Phi(0, g)$ for all $g \in G$, we obtain $\Phi(t, g) = g\alpha(t)$ for all $t \in \mathbb{R}$ by the uniqueness of the solution. It follows immediately that $\alpha(t) = \Phi(t, e)$ is the integral curve of L^A through the identity $e \in G$.

'$\Leftarrow$': For the reverse direction let α be the integral curve through $e \in G$ of the left invariant vector field L^A generated by $A \in T_eG$. We have:

$$\frac{\mathrm{d}}{\mathrm{d}s}\alpha(t)\alpha(s) = \frac{\mathrm{d}}{\mathrm{d}s}L_{\alpha(t)}\alpha(s) = TL_{\alpha(t)}\Big(\frac{\mathrm{d}}{\mathrm{d}s}\alpha(s)\Big) = TL_{\alpha(t)}(L^A_{\alpha(s)}) = L^A_{\alpha(t)\alpha(s)}\,,$$

where the third equality uses the integral curve property. Since the initial condition $\alpha(t)\alpha(0) = \alpha(t)$ holds we obtain $\alpha(t)\alpha(s) = \Phi(s, \alpha(t))$, where Φ is the flow of L^A, and further:

$$\alpha(t)\alpha(s) = \Phi\big(s, \alpha(t)\big) = \Phi_s \circ \Phi_t(e) = \Phi_{t+s}(e) = \alpha(t+s)\,. \qquad (*)$$

Another way to prove that $(*)$ holds globally is by noticing that $TL_{\alpha(t)}$ has full rank and thus $\mathrm{d}\alpha/\mathrm{d}t$ never vanishes (see [CDD77, section III.C.3]).

At first equation $(*)$ is valid only for small values of t, s because lemma 2.4.9 says that $\Phi_t \circ \Phi_s = \Phi_{t+s}$ only holds for values t, s where both sides are defined. But then we can successively enlarge this domain by multiplying up the values on the left-hand side to define α for bigger values of t. Thus, the integral curve $\alpha : \mathbb{R} \to G$ is complete and defines a subgroup of G. ∎

Remark 3.4.3 (Left/right invariant vector fields are complete) Alongside the last proof we have just shown that the flow of left invariant vector fields is always complete. This is because the flow of L^A is given by $\Phi(t, g) = g\alpha(t)$ and $\alpha(t)$ is defined for all $t \in \mathbb{R}$ (by the definition of a one-parameter subgroup). Also there is a one-parameter subgroup for any $A \in T_eG$. The same holds for right invariant vector fields.

It is an interesting fact that the flow of the *left* invariant vector field L^A is given by $\Phi(t, g) = g\alpha(t) = R_{\alpha(t)}g$ via the *right* action. Using the concept of a *generator* (see definition 2.4.10 on page 49) this can be formulated in the following way:

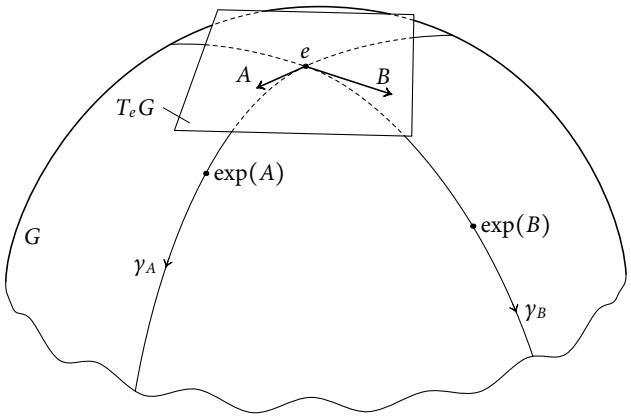


Figure 3.4 Exponential map $\exp : T_eG \to G$. The curve γ_A is the integral curve through $e \in G$ of the left invariant vector field generated by A.

Lemma 3.4.4 The *left* invariant vector field L^A is the generator of the one-parameter group of *right* translations $\{R_{\alpha(t)}\}$ of G, where α is the integral curve of L^A through $e \in G$. The *right* invariant vector field R^A generates the one-parameter group of *left* translations $\{L_{\alpha(t)}\}$ of G (here, α is the integral curve of R^A through e).

Proof. If Φ denotes the (global) flow of L^A this is just the simple observation that the diffeomorphism Φ_t acts on G by right translation $\Phi_t = R_{\alpha(t)}$. Analogously for R^A. ■

Now it is time to define the exponential map:

Definition 3.4.5 (Exponential map) The ***exponential map*** is given by:

$$\exp : \mathcal{L}G \to G\,, \quad A \mapsto \gamma_A(1)\,, \tag{3.14}$$

where γ_A is the integral curve through $e \in G$ of the left invariant vector field L^A generated by $A \in T_eG \cong \mathcal{L}G$.

Theorem 3.4.6 (Properties of exp) The exponential map $\exp : \mathcal{L}G \to G$ has the following properties:

(i) $\exp(tA) = \gamma_A(t)$ for all $t \in \mathbb{R}$, with γ_A from the definition,

(ii) $\exp(tA)\exp(sA) = \exp\big((t+s)A\big)$ for all $t, s \in \mathbb{R}$,

(iii) $\exp(0) = e$ and $T_0 \exp = \mathrm{id}_{\mathcal{L}G}$; hence there exist open neighbourhoods $0 \in U \subseteq \mathcal{L}G$ and $e \in V \subseteq G$ such that $\exp|_U : U \to V$ is a diffeomorphism.

Proof. For the first property we have to show $\gamma_{tA}(1) = \gamma_A(t)$. This follows from the uniqueness of integral curves and:

$$tA = t\,\frac{\mathrm{d}\alpha(s)}{\mathrm{d}s}\Big|_{s=0} = t\,\frac{\mathrm{d}\alpha(ts)}{\mathrm{d}(ts)}\Big|_{s=0} = \frac{\mathrm{d}\alpha(ts)}{\mathrm{d}s}\Big|_{s=0}\,.$$

The second property is then a simple consequence, making use of lemma 3.4.2:

$$\exp(tA)\exp(sA) = \gamma_{tA}(1)\gamma_{sA}(1) = \gamma_A(t)\gamma_A(s) = \gamma_A(t+s) = \exp\big((t+s)A\big)\,.$$

For the third property we have $\exp(0) = \exp(0A) = \gamma_A(0) = e$ for any choice of $A \in T_eG$. Furthermore:

$$T_0 \exp(A) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(tA) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma_{tA}(1) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma_A(t) = \dot{\gamma}_A(0) = A\,,$$

for each $A \in T_eG$, and hence $T_0 \exp = \mathrm{id}_{\mathfrak{L}G}$. The rest follows from the constant rank theorem, since $T_0 \exp = \mathrm{id}_{\mathfrak{L}G}$ is locally a linear isomorphism (see [Mico8, theorem 1.13]; the constant rank theorem is a generalisation of the inverse function theorem). ∎

There is an important characterisation of the exponential map which we want to cite from [FH91, Proposition 8.33]:

Theorem 3.4.7 (Uniqueness of exp) The exponential map is the unique map from $\mathfrak{L}G$ to G taking 0 to e whose differential at the origin $T_e \exp$ is the identity, and whose restrictions to the lines through the origin in $\mathfrak{L}G$ are one-parameter subgroups of G.

As consequence of the uniqueness theorem we can immediately conclude that the exponential function for *Matrix Lie groups* – Lie subgroups of $\mathrm{GL}(n)$ – is the well-known matrix exponential:

$$\exp : \mathfrak{L}\mathrm{GL}(n) \to \mathrm{GL}(n)\,, \quad A \mapsto e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k\,,$$

simply because the mapping defined this way has the correct properties.

Lemma 3.4.8 (Naturality of exp) If $\varphi : G \to H$ is a smooth morphism between Lie groups G and H, then the following diagram commutes:

$$\begin{array}{ccc} \mathfrak{L}G & \xrightarrow{T_e\varphi} & \mathfrak{L}H \\ {\scriptstyle\exp}\downarrow & & \downarrow{\scriptstyle\exp} \\ G & \xrightarrow{\varphi} & H \end{array}$$

Proof. We see that the curve $\alpha : \mathbb{R} \to H\,,\ t \mapsto \varphi(\exp(tA))$ defines a one-parameter subgroup of H, and thus is an integral curve of a left invariant vector field. To find the element of $\mathfrak{L}H$ that generates this vector field we calculate the tangent vector of α at the identity:

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \alpha(t) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \varphi\big(\exp(tA)\big) = T_e\varphi\big(T_0 \exp(A)\big) = T_e\varphi(A)\,.$$

Now $\exp(t\, T_e\varphi(A))$ is also an integral curve of the left invariant vector field generated by $T_e\varphi(A)$ and by uniqueness of the solutions (or with the help of theorem 3.4.7) we conclude that $\exp\big(t\, T_e\varphi(A)\big)$ has to be identical to $\varphi\big(\exp(tA)\big)$ for all $t \in \mathbb{R}$. Hence the diagram commutes. ∎

Corollary 3.4.9 (Some formulas) A consequence of lemma 3.4.8 is the formula:

$$g \exp(A) g^{-1} = I_g \circ \exp(A) = \exp(\mathrm{Ad}_g A)\,, \tag{3.15}$$

since $I_g : G \to G$ is a Lie group automorphism and $\mathrm{Ad}_g = T_e I_g$. Also:

$$\mathrm{Ad} \circ \exp_G = \exp_{\mathrm{GL}(\mathfrak{L}G)} \circ \mathrm{ad}\,, \tag{3.16}$$

since $\mathrm{Ad} : G \to \mathrm{Aut}(T_e G) \cong \mathrm{GL}(\mathfrak{L}G)$ is a Lie group morphism and $\mathrm{ad} = T_e \mathrm{Ad}$.

Although we have the simple multiplication rule $\exp(tA)\exp(sA) = \exp\big((t+s)A\big)$ from theorem 3.4.6, the multiplication rule for general exponentials is a bit more involved:

Theorem 3.4.10 (Baker–Campbell–Hausdorff (BCH) formula) Let G be a Lie group. Then there exists a neighbourhood $U \subseteq \mathfrak{L}G$ of the origin on which exp is invertible, and for $A, B \in U$ we have the multiplication formula:

$$\exp(A)\exp(B) = \exp\Bigg(A + B + \frac{1}{2}[A,B] + \frac{1}{12}\Big(\big[A,[A,B]\big] - \big[B,[B,A]\big]\Big) + \cdots\Bigg)\,. \tag{3.17}$$

Proof. See [Mico8, theorem 4.29] or [FH91, section 8.3]. ∎

If G is a simply connected, finite dimensional Lie group, the neighbourhood U in theorem 3.4.10 extends to all of G and the BCH formula holds globally. It is easy to see, on the other hand, that the neighbourhood U for the rotation group $\mathrm{SO}(3)$ is really a proper subset of G, because two rotations through π and through $-\pi$ are the same and hence the exponential map cannot be invertible unless we restrict ourselves to 'small' rotations.

The multiply connected group $G = \mathrm{SL}_2\mathbb{R}$ is an example where any $g \in \mathrm{SL}_2\mathbb{R}$ can be written as a *product* of exponentials, yet there exist $g \in \mathrm{SL}_2\mathbb{R}$ which can *not* be written as a single exponential (see remark 6.4.3).

Corollary 3.4.11 A variation of the BCH formula is:

$$\exp(tA)\exp(sB)\exp(-tA)\exp(-sB) = \exp\big(ts\,[A,B] + \mathcal{O}(t^2) + \mathcal{O}(s^2)\big)\,, \tag{3.18}$$

valid for small values of $t, s \in \mathbb{R}$. If, for example, the left-hand side is known explicitly, the formula can be used to calculate the Lie bracket of A and B.

3.5 Group Actions on Manifolds

In physical applications, Lie groups usually appear as groups of continuous transformations on manifolds. A lot of the concepts developed for Lie groups in the last sections immediately translate to transformation groups if we replace the canonical left or right translation by a general group action.

Definition 3.5.1 (Fundamental vector field) Let ρ be a smooth action of a Lie group G on some manifold M. Then for any element $A \in T_e G = \mathfrak{L}G$ of the associated Lie algebra we define the ***fundamental vector field*** $\zeta_\rho(A) \in \mathfrak{X}(M)$ with respect to ρ via:

$$\zeta_\rho(A)_x := T_e \rho^x(A) = T_{(e,x)}\rho(A, 0_x)\,, \tag{3.19}$$

where $\rho^x : G \to M$ is the partial mapping $\rho^x = \rho(\,\cdot\,, x)$ for fixed $x \in M$ (this is shown in figure 3.5).

There is a different convention where the definition depends on whether ρ is a *left* or *right* action (see remark 3.5.4).

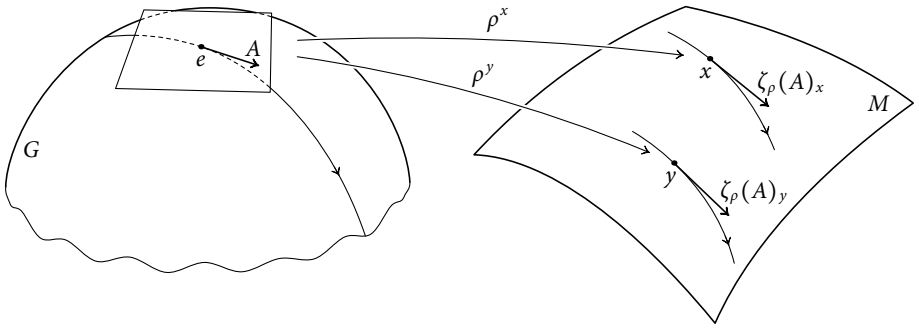


Figure 3.5 The partial mapping ρ^x in the definition of the fundamental vector field $\zeta_\rho(A)$ maps the curve on G with tangent vector A to a curve on M such that $\rho^x(e) = x$.

If we need to calculate the fundamental vector fields explicitly, we can use the exponential map to express the curve with tangent vector $A \in \mathfrak{L}G$ as $t \mapsto \exp(tA)$ and obtain the formula:

$$\zeta_\rho(A)_x = T_e\rho^x(A) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{\exp(tA)}x \,. \tag{3.20}$$

Lemma 3.5.2 Given a *left* action ℓ of a Lie group G on a manifold M we have:

(i) $T_x\ell_g(\zeta_\ell(A)_x) = \zeta_\ell(\mathrm{Ad}_g(A))_{\ell_g x}$ for all $x \in M$ and for all $g \in G$, that is, the following diagram commutes for each $g \in G$:

$$\begin{array}{ccc} \mathfrak{L}G & \xrightarrow{\zeta_\ell} & \mathfrak{X}(M) \\ {\scriptstyle \mathrm{Ad}_g}\downarrow & & \downarrow{\scriptstyle \ell_{g\#}} \\ \mathfrak{L}G & \xrightarrow{\zeta_\ell} & \mathfrak{X}(M) \end{array} \,. \tag{3.21}$$

Hence ζ_ℓ is *equivariant* with respect to the G-actions given in the diagram.

(ii) The vector field $R^A \times 0_M \in \mathfrak{X}(G \times M)$ is ℓ-related to $\zeta_\ell(A) \in \mathfrak{X}(M)$, that is, the following diagram commutes:

$$\begin{array}{ccc} T(G \times M) & \xrightarrow{T\ell} & TM \\ {\scriptstyle R^A \times 0_M}\uparrow & & \uparrow{\scriptstyle \zeta_\ell(A)} \\ G \times M & \xrightarrow{\ell} & M \end{array} \,. \tag{3.22}$$

Proof. (i) For the first property we write out the left-hand side and insert the definition of the fundamental vector field $\zeta_\ell(A)$:

$$T_x\ell_g\big(\zeta_\ell(A)_x\big) = T_x\ell_g \circ T_e\ell^x(A) = T_e(\ell_g \circ \ell^x)(A) \,.$$

Now remember that the adjoint action on the right-hand side is defined as $\mathrm{Ad}_g = T_eI_g$, with $I_g(h) = ghg^{-1}$. We therefore rewrite $\ell_g \circ \ell^x$ to contain I_g:

$$\ell_g \circ \ell^x(h) = \ell(gh, x) = \ell(ghg^{-1}g, x) = \ell(I_g(h), \ell_g x) = \ell^{\ell_g x} \circ I_g(h) \quad \forall h \in G \,.$$

Inserting this into the equation above we get the desired result:

$$T_x\ell_g\big(\zeta_\ell(A)_x\big) = T_e(\ell_g \circ \ell^x)(A) = T_e(\ell^{\ell_x}) \circ \mathrm{Ad}_g(A) = \zeta_\ell(\mathrm{Ad}_g(A))_{\ell_g x} .$$

(ii) The second property states that diagram (3.22) is commutative. To prove this, take $g \in G$ and $x \in M$, and use the definition of $\zeta_\ell(A)$ to obtain:

$$\zeta_\ell(A) \circ \ell(g,x) = \zeta_\ell(A)_{\ell_g x} = T_{e,\ell_g x}\ell(A, 0_{\ell_g x}) = T_{e,\ell_g x}\ell\big((\mathrm{id}_{T_eG} \times T\ell_g)(A, 0_x)\big) ,$$

where we have $T\ell_g(0_x) = 0_{\ell_g x}$ because the tangent map is linear. Due to the chain rule this is equivalent to:

$$\zeta_\ell(A)_{\ell_g x} = T\ell \circ (\mathrm{id}_{T_eG} \times T\ell_g)(A, 0_x) = T\big(\ell \circ (\mathrm{id}_G \times \ell_g)\big)(A, 0_x) .$$

For the mapping inside the brackets we have:

$$\ell \circ (\mathrm{id}_G \times \ell_g)(h,x) = \ell(h, \ell_g x) \overset{(*)}{=} \ell(hg, x) = \ell(R_g h, x) = \ell \circ (R_g \times \mathrm{id}_M)(h,x) ,$$

In the equality marked $(*)$ we used that ℓ_g is a *left* action. For a *right* action r we would have gotten $r(h, r_g x) = r(gh, x)$ instead, and hence the *left* invariant vector field L^A would appear in the final result.

for all $h \in G$, $x \in M$, and thus:

$$\begin{aligned}\zeta_\ell(A)_{\ell_g x} &= T\big(\ell \circ (R_g \times \mathrm{id}_M)\big)(A, 0_x) = T\ell \circ (TR_g \times \mathrm{id}_{TM})(A, 0_x) \\ &= T\ell\big(T_eR_g(A), 0_x\big) = T\ell \circ (R^A \times 0_M)(g,x) .\end{aligned}$$

Hence $R^A \times 0_M$ is ℓ-related to $\zeta_\ell(A)$. ∎

Theorem 3.5.3 (Fundamental vector fields for left and right actions) Given a *left* action ℓ of a Lie group G on a manifold M, the map $\zeta_\ell : \mathfrak{L}G \to \mathfrak{X}(M)$, $A \mapsto \zeta_\ell(A)$ is a Lie algebra *anti*morphism, meaning ζ_ℓ is $\mathbb{R}$-linear and:

$$\big[\zeta_\ell(A), \zeta_\ell(B)\big] = -\zeta_\ell\big([A,B]\big) , \tag{3.23}$$

The 'anti' in antimorphism is responsible for the minus sign in the formula.

for all $A, B \in \mathfrak{L}G$. In turn, if $r : G \times M \to M$ is a *right* action, the map $\zeta_r : \mathfrak{L}G \to \mathfrak{X}(M)$ is a Lie algebra *morphism*, i. e.

$$\big[\zeta_r(A), \zeta_r(B)\big] = \zeta_r\big([A,B]\big) , \tag{3.24}$$

for all $A, B \in \mathfrak{L}G$.

Proof. The $\mathbb{R}$-linearity of ζ_ℓ follows immediately from the definition.

Then, according to part (ii) of the last lemma, the vector field $R^A \times 0_M$ is ℓ-related to $\zeta_\ell(A)$ and thus the Lie bracket $[R^A \times 0_M, R^B \times 0_M]$ is ℓ-related to $[\zeta_\ell(A), \zeta_\ell(B)]$, due to lemma 2.4.17 (the proof uses φ-related vector fields).

On the other hand, given $A, B \in \mathfrak{L}G$, we have:

$$\big[R^A \times 0_M, R^B \times 0_M\big] = \big[R^A, R^B\big] \times 0_M = R^{-[A,B]} \times 0_M ,$$

where we first used the definition of the Jacobi–Lie bracket and then lemma 3.3.6. Again by the last lemma we see that $-R^{[A,B]} \times 0_M$ is ℓ-related to $\zeta_\ell(-[A,B]) = -\zeta_\ell([A,B])$, and hence equation (3.23) is proven.

For the second part use the remark in the proof of the last lemma to show that the vector field $L^A \times 0_M$ is r-related to $\zeta_r(A)$. Then equation (3.24) follows analogously to the argument above, only that $[L^A, L^B] = L^{[A,B]}$ doesn't produce a minus sign. ∎

Remark 3.5.4 In the light of theorem 3.5.3, for a left action ℓ it is common to define:

$$\gamma_\ell(A)_x := \zeta_\ell(-A)_x = T_e\ell^x(-A) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \ell_{\exp(-tA)}x\,, \tag{3.25}$$

If you prefer, you can also turn ℓ into a right action r, via $r_g := \ell_{g^{-1}}$. The right-hand side of equation (3.25) is then the *fundamental* vector field $\zeta_r(A)$ with respect to r.

such that $\gamma_\ell : \mathfrak{L}G \to \mathfrak{X}(M)$ is a Lie algebra *morphism* – not an *anti*morphism. We will call $\gamma_\ell(A)$ the ***antifundamental vector field*** associated to $A \in \mathfrak{L}G$. In case of a left action, however, some authors prefer to call $\gamma(A)$ a *fundamental* vector field and, in turn, $\zeta(A)$ an *antifundamental* vector field, which is the opposite of our convention. Irrespective of terminology, the mapping γ_ℓ is the morphism γ that is later used by the quantization scheme via the Canonical Group.

3.6 Semidirect Product and Extensions of Groups

The semidirect product can be found in books on *Lie groups*, in particular its relation to Mackey's theory of *induced representations* (see e. g. [Sim96, chapter V]). Group extensions in general are often discussed in works on *Homological algebra*, for example [HS71].

For the quantization approach via the Canonical Group we will need a *transitive* group action on phase space, corresponding to a *complete* set of classical observables. A typical situation, however, is that we have found by physical considerations a group that *doesn't* act transitively, and hence we seek a way to somehow 'extend' it.

Let's say the group G acts on a given space X by a non-transitive action. Since points of a fixed orbit are already connected by transformations out of G, we look for another transformation group, say N, that maps elements *between* these orbits. The question is how both groups can be combined to give another group that acts on X. The direct product $N \times G$ springs to mind, but it turns out this only works if the group actions of N and G commute with each other. Instead, we will have to use the *semidirect product*, which respects also non-commuting group actions.

Observe, $\ell_g \in \mathrm{Aut}(N)$ implies that ℓ_g is a *group homomorphism* of N for each $g \in G$.

Definition 3.6.1 (External semidirect product) Let N, G be two groups together with a left action $\ell : G \to \mathrm{Aut}(N)$ (for topological groups ℓ has to be continuous; for Lie groups ℓ has to be smooth). Then the (***external***) ***semidirect product*** $N \rtimes_\ell G$ with respect to ℓ is the set $N \times G$ together with the multiplication law:

$$(n_2, g_2) \cdot (n_1, g_1) := \big(n_2\, \ell_{g_2}(n_1), g_2 g_1\big)\,. \tag{3.26}$$

The triangle in the symbol '$\rtimes$' can remind us of the fact that the natural embedding of N in the semidirect product is a *normal subgroup*, often written as $N \lhd N \rtimes G$.

It is common to write $N \rtimes G$ instead of $N \rtimes_\ell G$ if the group action of N on G is apparent from the context.

The neutral element of $N \rtimes G$ is (e_N, e_G); the inverse of $(n, g) \in N \rtimes G$ is given by $(n, g)^{-1} = (\ell_{g^{-1}}(n^{-1}), g^{-1})$. The semidirect product reduces to the *direct* product of groups if ℓ is taken to be the trivial action $\ell : G \to \mathrm{Aut}(N)\,,\ g \mapsto \mathrm{id}_N$.

Example 3.6.2

- The *Euclidean group* $\mathrm{E}(n)$ is isomorphic to $\mathrm{T}(n) \rtimes \mathrm{O}(n)$, the semidirect product of translations $T(n)$ and orthogonal transformations $\mathrm{O}(n)$. If ρ denotes the action of $\mathrm{T}(n)$ on $\mathbb{R}^n$ and σ the action of $\mathrm{O}(n)$, the combined action action of $\mathrm{T}(n) \rtimes \mathrm{O}(n)$ on $\mathbb{R}^n$ is given by:

  $$\tau_{(\boldsymbol{a},R)}(\boldsymbol{x}) := \rho_{\boldsymbol{a}} \circ \sigma_R(\boldsymbol{x}) = R\boldsymbol{x} + \boldsymbol{a}\,,$$

  with group multiplication according to (3.26):

  $$(\boldsymbol{a}_2, R_2) \cdot (\boldsymbol{a}_1, R_1) = (\boldsymbol{a}_2 + R_2\boldsymbol{a}_1, R_2R_1)\,.$$

  The action ℓ is implicit, since $\mathrm{O}(n)$ acts on $\mathrm{T}(n) \cong \mathbb{R}^n$ in a canonical way.

- The *Poincaré group* is isomorphic to the semidirect product $T(1,3) \rtimes \mathrm{O}(1,3)$ of space-time translations and Lorentz transformations.
- The orthogonal group $\mathrm{O}(n)$ is isomorphic to $\mathrm{SO}(n) \rtimes \mathbb{Z}_2$, where $\mathbb{Z}_2$ acts on $\mathbb{R}^n$ by reflections.

Definition 3.6.3 (Internal semidirect product) Let H be a group with subgroups N and G. Then H is called an ***internal semidirect product*** of N and G if:

The semidirect product turns into a *direct product* $N \times G$ if and only if *both* N and G are normal subgroups of H.

- $N \lhd H$ is a *normal* subgroup,
- N and G generate H, i. e. each $h \in H$ can be written as $h = ng$ for $n \in N$ and $g \in G$,
- $N \cap G = \{e_H\}$.

Theorem 3.6.4 Each *external* semidirect product $N \rtimes G$ is isomorphic to an *internal* semidirect product of N and G, and vice versa.

Proof. '$\Rightarrow$': Let $H := N \rtimes G$ be an external semidirect product. Let $\pi_G : N \rtimes G \to G$ be the natural projection $\pi_G(n,g) := g$ and let $\iota_N : N \to N \rtimes G$ be the natural embedding $\iota_N(n) := (n, e_G)$. Both are group homomorphisms and we have $\operatorname{im} \iota_N = \ker \pi_G$. Since the kernel of a group homomorphism is a normal subgroup, the image $\operatorname{im} \iota_N$ is thus a normal subgroup of H.

Note, however, that π_N is *not* in general a homomorphism, so the embedding of G doesn't have to be a normal subgroup.

Let $\iota_G : N \to N \rtimes G$, $\iota_G(g) := (e_N, g)$ be the natural embedding of G, then $N \rtimes G$ is generated by $\operatorname{im} \iota_N$ and $\operatorname{im} \iota_G$, since for each $(n,g) \in N \rtimes G$ we have a decomposition $(n,g) = (n, e_G) \cdot (e_N, g) = \iota_N(n) \cdot \iota_G(g)$ according to the multiplication law (3.26) of the external semidirect product, with $n \in N$ and $g \in G$.

Furthermore, $\operatorname{im} \iota_N \cap \operatorname{im} \iota_G = \{(n, e_G)\} \cap \{(e_N, g)\} = \{(e_N, e_G)\} = e_H$ and hence the external semidirect product $N \rtimes G$ is isomorphic to the internal semidirect product of $\operatorname{im} \iota_N \cong N$ and $\operatorname{im} \iota_G \cong G$.

'$\Leftarrow$': Let H be an internal semidirect product of N and G. We show that H is isomorphic to the external semidirect product $N \rtimes_\ell G$ with $\ell_g(n) := gng^{-1}$.

Given that $N \cap G = \{e_H\}$, each element $h \in H$ has a unique decomposition $h = ng$ for $n \in N$ and $g \in G$ (because $\tilde{n}\tilde{g} = ng$ implies $n^{-1}\tilde{n} = g\tilde{g}^{-1} \in N \cap G = \{e\}$ and hence $\tilde{n}n^{-1} = e = \tilde{g}^{-1}g$, so $\tilde{n} = n$ and $\tilde{g} = g$). Thus $\varphi : N \times G \to H$, $(n,g) \mapsto ng$ is a bijection of the underlying sets.

To show that φ is, additionally, a group homomorphism, take two group elements (n,g) and $(\tilde{n},\tilde{g})$ in $N \rtimes_\ell G$ and compute:

$$\varphi\big((n,g)\cdot(\tilde{n},\tilde{g})\big) = \big(n\,\ell_g(\tilde{n})\big)(g\tilde{g}) = ng\tilde{n}g^{-1}g\tilde{g} = ng\tilde{n}\tilde{g} = \varphi(n,g)\cdot\varphi(\tilde{n},\tilde{g})\,.$$

Consequently, the mapping $\varphi : N \rtimes_\ell G \to H$ is an isomorphism of groups. ∎

There is an interesting relation between semidirect products and split short exact sequences of groups. This will help us understand why the semidirect product is in some cases more natural than the direct product of groups.

Definition 3.6.5 (Split short exact sequence) Consider the following short exact sequence (with additional maps t and u that may not exist, drawn for reference):

$$0 \longrightarrow A \underset{t}{\overset{j}{\hookrightarrow}} B \underset{u}{\overset{p}{\twoheadrightarrow}} C \longrightarrow 0\ ,$$

We call the sequence:

If t exists it is surjective.
If u exists it is injective.

- ***left split*** if j has a left inverse $t : B \to A$, that is $t \circ j = \mathrm{id}_A$.
- ***right split*** (or just ***split***) if p has a right inverse $u : C \to B$, that is $p \circ u = \mathrm{id}_C$.

Lemma 3.6.6 (Splitting lemma for groups) Consider groups H, N and G. Then H is isomorphic to the semidirect product of N and G:

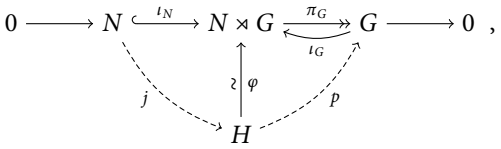
$$H \cong N \rtimes G \,,$$

if and only if there exists a split short exact sequence:

$$0 \longrightarrow N \overset{j}{\hookrightarrow} H \underset{u}{\overset{p}{\rightleftarrows}} G \longrightarrow 0 \quad . \tag{3.27}$$

Proof. '$\Rightarrow$': Let $H \cong N \rtimes G$ and denote the isomorphism by $\varphi : H \to N \rtimes G$. It is an easy and straight-forward calculation to check that the upper row in the following diagram is a short exact sequence that splits by the map ι_G:

Note that the natural projection π_N in general is *not* a group homomorphism, so there is no canonical left split. This is why the lemma explicitly requires a *right split* sequence.



where ι_N and ι_G are the natural embeddings, π_G the natural projection. From this we obtain a split short exact sequence like (3.27) if the maps are defined by:

$$j := \varphi^{-1} \circ \iota_N \,, \quad p := \pi_G \circ \varphi \,, \quad u := \varphi^{-1} \circ \iota_G \,.$$

'$\Leftarrow$': For the reverse direction suppose there exists a split short exact sequence like the one in (3.27). We show that H is the internal semidirect product of $\mathrm{im}\, j$ and $\mathrm{im}\, u$.

First, note that $\mathrm{im}\, j$ is a *normal subgroup* of H due to $\mathrm{im}\, j = \ker p$, which follows from the exactness of the sequence.

To show that H is generated by $\mathrm{im}\, j$ and $\mathrm{im}\, u$, take $h \in H$ and define $\tilde{h} := u \circ p(h)$, so that $\tilde{h}$ lies in the image $\mathrm{im}\, u$ (remember that $p \circ u = \mathrm{id}_G$, but $u \circ p \neq \mathrm{id}_H$, at least in general). We then have:

$$p(h\tilde{h}^{-1}) = p(h)p(\tilde{h})^{-1} = p(h)\big(p \circ u \circ p(h)\big)^{-1} = p(h)p(h)^{-1} = e_G \,,$$

which tells us that $h\tilde{h}^{-1} \in \ker p$. Let us now write $x := h\tilde{h}^{-1}$ and remember $\ker p = \mathrm{im}\, j$, then $h = x\tilde{h}$ is a decomposition of h with $x \in \mathrm{im}\, j$ and $\tilde{h} \in \mathrm{im}\, u$. Consequently, the images of j and u generate H.

Finally, let us show that the intersection of the images of j and u is trivial. To this end, take $h \in \mathrm{im}\, j \cap \mathrm{im}\, u$. Accordingly, there exists $g \in G$ such that $h = u(g)$ and $u(g) \in \mathrm{im}\, j$. Using $\mathrm{im}\, j = \ker p$ we obtain $g = \mathrm{id}_G(g) = p \circ u(g) = e_G$ and further $h = u(g) = u(e_G) = e_H$. Hence $\mathrm{im}\, j \cap \mathrm{im}\, u = \{e_H\}$.

At this point we know that H is an internal semidirect product of $\mathrm{im}\, j$ and $\mathrm{im}\, u$. By theorem 3.6.4 we can therefore conclude that H is isomorphic to the external semidirect product $N \rtimes G$, which concludes the proof. ∎

Before we finish this chapter, we need to say a few words about group extensions in general (a so-called *central extension*, in particular, plays an important role in section 5.3, where we discuss the relation between Canonical Group $\mathcal{C}$ and Geometric Group $\mathcal{G}$).

The last lemma already showed that semidirect products correspond exactly to *split* short exact sequences. On the other hand, given a general short exact sequence:

$$0 \longrightarrow N \overset{j}{\hookrightarrow} G \overset{p}{\twoheadrightarrow} Q \longrightarrow 0 \ ,$$

we have $N \cong \operatorname{im} j = \ker p$, so that N is isomorphic to the kernel of p. Hence, by the first isomorphism theorem of groups ($G/\ker p \cong \operatorname{im} p$) we get $Q = \operatorname{im} p \cong G/\ker p \cong G/N$, and the above sequence is thus equivalent (up to isomorphism) to the sequence:

$$0 \longrightarrow N \overset{j}{\hookrightarrow} G \overset{\pi}{\twoheadrightarrow} G/N \longrightarrow 0 \ ,$$

where π is the natural projection onto the quotient group G/N. Accordingly, we find that a short exact sequence therefore captures *quotient relations* between groups. This leads to the following definition:

Definition 3.6.7 (Group extension) A ***group extension*** is a short exact sequence:

$$0 \longrightarrow N \overset{j}{\hookrightarrow} G \overset{p}{\twoheadrightarrow} Q \longrightarrow 0 \ .$$

One then calls G the ***extension of*** Q ***by*** N.

If the sequence splits, we say that G is a ***split extension*** of Q. If the image of j lies in the centre $\mathcal{Z}(G) = \{z \in G : zg = gz \ \forall g \in G\}$, we call G a ***central extension*** of Q.

Note that the group Q is *no subgroup* of its extension G! Sometimes (if and only if the sequence splits, i. e. iff G is a semidirect product) Q can be embedded isomorphically, but this is not the case in general.

An extension not only depends on N and Q but also on the morphisms j and p. It is therefore useful to define equivalence classes, where two group extensions G and $\tilde{G}$ of Q by N are called ***equivalent*** if there exists a group homomorphism $\varphi : G \to \tilde{G}$ such that the following diagram commutes:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & N & \overset{j}{\hookrightarrow} & G & \overset{p}{\twoheadrightarrow} & Q & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \varphi} & & \| & & \\ 0 & \longrightarrow & N & \overset{\tilde{j}}{\hookrightarrow} & \tilde{G} & \overset{\tilde{p}}{\twoheadrightarrow} & Q & \longrightarrow & 0 \ . \end{array}$$

By the *short five lemma*, φ is automatically an isomorphism (see e. g. [HS71, lemma 1.1] or [Mac98, section VIII.4, lemma 1]; although usually proven for Abelian categories, the short five lemma also holds in the full category of groups).

It turns out that the equivalence classes of group extensions are deeply related to the second cohomology groups $H^2(G, A)$ of a group G. In particular, central extensions of G correspond to cohomology groups with A a trivial G-module (see [HS71, chapter VI, specifically section 10]).

The same concepts apply to Lie algebras if we use the following definition:

Definition 3.6.8 (Algebra extension) An ***algebra extension*** is a short exact sequence of algebras with algebra morphisms:

$$0 \longrightarrow \mathcal{N} \overset{j}{\hookrightarrow} \mathcal{E} \overset{p}{\twoheadrightarrow} \mathcal{Q} \longrightarrow 0 \ .$$

Like for groups, the algebra $\mathcal{Q}$ is *no subalgebra* of its extension $\mathcal{E}$. If and only if $\mathcal{E}$ is a split extension, $\mathcal{Q}$ can be embedded isomorphically.

We then say that $\mathcal{E}$ is an ***extension of*** $\mathcal{Q}$ ***by*** $\mathcal{N}$. Analogously to above, we speak of a ***split extension*** if the given sequence splits. If the image $\operatorname{im} j$ lies in the centre $\mathcal{Z}(\mathcal{E})$ we will call $\mathcal{E}$ a ***central extension*** of $\mathcal{Q}$.

A central extension of the Lie algebra $\mathfrak{LG}$, where $\mathcal{G}$ is the so-called *Geometric Group*, will appear later when we discuss the quantization scheme. If you want to learn more about group and algebra extensions, and their relation to homology, good starting points are Rotman [Rot02, chapter 10], the classic book of Mac Lane [Mac95] and, of course, the frequently cited [HS71].

4 Symplectic Geometry and Classical Mechanics

Can a symplectic camel go through the eye of a needle?
Vladimir I. Arnol'd [Arn86]

Symplectic geometry is the natural mathematical framework for the Hamiltonian formulation of classical mechanics. It is thereby mainly a matter of convention that most mathematicians prefer the phrase 'symplectic' for what physicists call 'canonical' (as in 'canonical transformation'). Nevertheless, we will stick to the mathematicians' choice in this case, because the word 'canonical' is already much too overused.

The first section of this chapter is meant to summarise the basic definitions of symplectic geometry. Afterwards, we introduce Hamiltonian vector fields and show that they constitute a Lie algebra. In the third section we look at the relation between Hamiltonian vector fields and classical observables. Finally, we show that there is a canonical symplectic structure on the cotangent bundle and discuss some implications.

4.1 Symplectic Manifolds

Definition 4.1.1 (Symplectic manifold) A ***symplectic manifold*** (M, ω) is a manifold M together with a ***symplectic form*** $\omega \in \Omega^2(M)$, i. e. a 2-form ω on M which is closed and pointwise nondegenerate. As usual, *closed* means $\mathrm{d}\omega = 0$, and *pointwise nondegenerate* is the property that for each point $x \in M$ the mapping $T_xM \to T_x^*M$, $v \mapsto \omega_x(v, \cdot)$ is an isomorphism.

The closedness $\mathrm{d}\omega = 0$ ensures that the Lie bracket of Hamiltonian vector fields is again Hamiltonian (corollary 4.2.5). This is necessary in order for the Poisson bracket to exist.

Symplectic manifolds are always *even-dimensional* because ω nondegenerate implies that the matrix representation (ω_{ij}) in any chart has to be *invertible*. On the other hand, we know that (ω_{ij}) has to be *antisymmetric*, because ω is skew-symmetric. Given that antisymmetric matrices are invertible only in even dimensions (their determinant vanishes in odd dimensions) it follows that M has to be even-dimensional.

Definition 4.1.2 (Symplectomorphism) Given two symplectic manifolds (M, ω) and (N, η), a ***symplectomorphism*** (also ***symplectic***, or ***canonical transformation***) between M and N is a *diffeomorphism* $\varphi : M \to N$ that respects the symplectic forms, i. e.

$$\varphi^*\eta = \omega \,, \tag{4.1}$$

See also definition 4.2.6 for the special case of *Hamiltonian symplectomorphisms*.

or, pointwise:

$$(\varphi^*\eta)_x(v, w) = \omega_x(v, w) \quad \forall v, w \in T_xM \,. \tag{4.2}$$

The symplectomorphisms are the *isomorphisms* in the category of symplectic manifolds. For a fixed manifold (M, ω), the symplectomorphisms from M to M (that is, symplectic automorphisms) constitute a subgroup $\mathrm{Sp}(M) \leq \mathrm{Diff}(M)$.

Definition 4.1.3 (Symplectic vector field) A vector field $X \in \mathfrak{X}(M)$ is called ***symplectic*** if the Lie derivative of the symplectic form ω along X vanishes:

$$\mathcal{L}_X\omega = 0 \,. \tag{4.3}$$

This is analogous to *Killing vector fields* on a Riemannian manifold, for which the Lie derivative $\mathcal{L}_X g$ of the metric tensor g vanishes.

The symplectic vector fields form a sub Lie algebra $\mathfrak{X}_{\mathrm{Sp}}(M) \subseteq \mathfrak{X}(M)$.

Actually, this says symplectic vector fields are *tangent vectors* to symplectomorphisms if we interpret a symplectomorphism of M as a *point* in the infinite-dimensional manifold $\mathrm{Sp}(M)$.

Lemma 4.1.4 (Flow of symplectic vector fields) A vector field X is symplectic if and only if its flow Φ_t is a symplectomorphism for each t.

Proof. '⇒': Let X be symplectic and let Φ be the flow of X, then $\mathcal{L}_X\omega = 0$ and thus:

$$0 = \Phi_t^*(\mathcal{L}_X\omega) = \mathcal{L}_X(\Phi_t^*\omega) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \Phi_s^*\Phi_t^*\omega = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \Phi_{s+t}^*\omega = \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t^*\omega\ .$$

Accordingly, we get $\Phi_t^*\omega = \omega$ for all t where Φ_t is defined.

'⇐': Let the flow Φ_t be a symplectomorphism, that is $\Phi_t^*\omega = \omega$, for each t where Φ_t is defined. Using the definition of the Lie derivative we obtain:

$$\mathcal{L}_X\omega = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Phi_t^*\omega = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \omega = 0\ ,$$

and thus X is a symplectic vector field. ∎

Lemma 4.1.5 A vector field X is symplectic if and only if $\mathrm{d}(\mathrm{i}_X\omega) = 0$, i. e.

$$\mathcal{L}_X\omega = 0 \ \Leftrightarrow\ \mathrm{d}(\mathrm{i}_X\omega) = 0\ . \tag{4.4}$$

Proof. By Cartan's magic formula 2.8.7 we have $\mathcal{L}_X\omega = \mathrm{d}(\mathrm{i}_X\omega) + \mathrm{i}_X(\mathrm{d}\omega)$. Since ω is a symplectic form, it is closed, i. e. $\mathrm{d}\omega = 0$, and thus (4.4) follows. ∎

Definition 4.1.6 (Symplectic group action) Let G be a Lie group and let (M, ω) be a symplectic manifold. A ***symplectic group action*** ρ is a group action $\rho : G \to \mathrm{Sp}(M)$ that acts on M by symplectomorphisms, i. e. $\rho_g^*\omega = \omega$ for each $g \in G$.

Lemma 4.1.7 (Fundamental vector fields of a symplectic action) Given a symplectic action ρ of G on M, the fundamental vector fields $\zeta_\rho(A)$ are symplectic.

Proof. This is an immediate consequence of lemma 4.1.4 because the flow of a fundamental vector field $\zeta_\rho(A)$ is given by $\Phi_t = \rho_{\exp(tA)}$. Moreover, $\rho_{\exp(tA)}$ is a symplectomorphism for each $t \in \mathbb{R}$ because ρ is a symplectic action. ∎

The prototype for a symplectic space is the vector space $\mathbb{R}^{2n}$, which has a *canonical* symplectic form ω_0, given by:

$$\omega_0(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^{\mathrm{T}} J \boldsymbol{y} \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{2n}\ , \quad J = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}\ , \tag{4.5}$$

where J is the matrix representation of ω_0 in the canonical basis. An important theorem in symplectic geometry is *Darboux's theorem* which states that symplectic manifolds are locally symplectomorphic to a subset of $(\mathbb{R}^{2n}, \omega_0)$:

An analogous theorem does not exist for Riemannian manifolds since they have *local invariants* (e. g. curvature and torsion). The proof of Darboux's theorem in [AM78, theorem 3.2.2] reveals that this is a direct consequence of $\mathrm{d}\omega = 0$.

Theorem 4.1.8 (Darboux) Let (M, ω) be a $2n$-dimensional symplectic manifold. Then for each point in M there exists an open neighbourhood U together with a *symplectomorphism* $\xi : U \to \tilde{U} \subseteq \mathbb{R}^{2n}$ (the pair (U, ξ) is called a ***Darboux chart***) to an open subset $\tilde{U}$ of $\mathbb{R}^{2n}$ endowed with the canonical symplectic form ω_0 as in (4.5).

Proof. See for example [Wal07, Satz 3.1.24] or [MR99, section 5.1]. ∎

For our purposes the most important implication of Darboux's theorem is the fact that around each point in M there exists a Darboux charts (U, ξ) such that the symplectic form ω locally takes on the particularly simple form:

$$\omega|_U = \sum_{i=1}^{n} \mathrm{d}\xi^i \wedge \mathrm{d}\xi^{i+n} = \sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p_i \,, \tag{4.6}$$

where we write $(q^1, \dots, q^n, p_1, \dots, p_n) := (\xi^1, \dots, \xi^{2n})$. In physical applications, ξ is the coordinate representation of a point in phase space, where the q^i are the *generalised coordinates*, and the p_i are the *canonically conjugate momenta*.

4.2 Hamiltonian Vector Fields

Before we define Hamiltonian vector fields, we first want to introduce the so-called *musical isomorphisms*. The basic observation is that, although any finite dimensional vector space V is isomorphic to its dual V^*, there is in general no *canonical* choice for this isomorphism. On symplectic manifolds, however, we have the symplectic form ω which is pointwise nondegenerate and thus singles out an isomorphism $T_xM \to T_x^*M$ for each $x \in M$. Given that ω is a smooth 2-form, this induces a smooth vector bundle isomorphism $TM \to T^*M$ and thereby also a $C^\infty(M, \mathbb{R})$-linear isomorphism of sections $\Gamma^\infty(TM) \to \Gamma^\infty(T^*M)$, i. e. an isomorphism between vector fields $\mathfrak{X}(M) = \Gamma^\infty(TM)$ and 1-forms $\Omega^1(M) = \Gamma^\infty(T^*M)$ on M.

Definition 4.2.1 (Musical isomorphisms) Let (M, ω) be a symplectic manifold. The ***musical isomorphisms*** are:

$$\begin{array}{ccc} TM & \overset{\flat}{\underset{\sharp}{\rightleftarrows}} & T^*M \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ M & \xrightarrow[\mathrm{id}_M]{} & M \end{array}$$

$$\flat : \ \mathfrak{X}(M) \to \Omega^1(M) \,, \quad X \mapsto X^\flat \,, \quad X^\flat := \mathrm{i}_X\omega = \omega(X, \cdot) \,, \tag{4.7}$$

$$\sharp : \Omega^1(M) \to \mathfrak{X}(M) \,, \quad \alpha \mapsto \alpha^\sharp \,, \quad \sharp := \flat^{-1} \,. \tag{4.8}$$

They are named ***flat*** and ***sharp***, after the musical notation, since $\flat$ takes vector fields and 'flattens' them to their corresponding 1-forms, whereas $\sharp$ takes 1-forms and 'sharpens' them to give the associated vector fields.

In a Darboux chart (U, ξ), if we write $\xi = (q, p)$ as before, we get a decomposition $(\partial_i) = (\partial_i^q, \partial_p^i)$ of the basis vector fields of TU. Moreover, according to equation (4.6), we have $\omega|_U = \mathrm{d}q^i \wedge \mathrm{d}p_i$ for the symplectic form ω. This way we obtain a coordinate expression for $X^\flat$:

$$\begin{aligned} \left(X^i\partial_i\right)^\flat &= \left(X_q^i\partial_i^q + X_j^p\partial_p^j\right)^\flat = \mathrm{d}q^j \wedge \mathrm{d}p_j(X^i\partial_i, \cdot) \\ &= \mathrm{d}q^j(X^i\partial_i)\mathrm{d}p_j - \mathrm{d}p_j(X^i\partial_i)\mathrm{d}q^j = X_q^j\,\mathrm{d}p_j - X_j^p\,\mathrm{d}q^j \,. \end{aligned}$$

The isomorphism between TM and T^*M is thus given locally by:

$$(\partial_i) = \begin{pmatrix} \partial_i^q \\ \partial_p^i \end{pmatrix} \overset{\flat}{\underset{\sharp}{\rightleftarrows}} \begin{pmatrix} \mathrm{d}p_i \\ -\mathrm{d}q^i \end{pmatrix} = (J_{ik}\mathrm{d}x^k) \quad , \tag{4.9}$$

where J is the matrix representation of ω_0 from equation (4.5).

Definition 4.2.2 (Hamiltonian vector field) Let (M, ω) be a symplectic manifold. The ***Hamiltonian vector field*** X_f ***associated to*** $f \in C^\infty(M, \mathbb{R})$ is the vector field obtained by 'sharpening' the differential:

$$X_f := \mathrm{d}f^\sharp \;\Leftrightarrow\; X_f{}^\flat = \mathrm{d}f \,. \tag{4.10}$$

A vector field X is thus (***globally***) ***Hamiltonian*** if $X^\flat = \mathrm{i}_X\omega$ is an *exact* 1-form. On the other hand, a vector field X is called ***locally Hamiltonian*** if $X^\flat$ is a *closed* form.

It is common to drop the induced bundle chart of TM from the notation and write X_f also for the coordinate representation of X_f, instead of the somewhat clumsy $Tx(X_f)$.

In a Darboux chart (U, ξ) the local expression for X_f is:

$$X_f = \frac{\partial f}{\partial p_i}\,\partial_q^i - \frac{\partial f}{\partial q^i}\,\partial_i^p \,. \tag{4.11}$$

We denote the set of globally Hamiltonian vector fields by $\mathfrak{X}_{\mathrm{Ham}}(M)$.

We will see in a second that Hamiltonian vector fields form a Lie algebra – this will also lead us to define the Poisson bracket – and that a vector field is *locally Hamiltonian* if and only if it is *symplectic*. Still, before we continue we want to take a short break and gain a better understanding of the relation between a Hamiltonian vector field X_f and the function f it stems from.

The given example translates easily to general symplectic manifolds, because Darboux's theorem tells us that *all* symplectic manifolds are locally symplectomorphic to open subsets of $\mathbb{R}^{2n}$.

Remark 4.2.3 (Relation between f and X_f) Let us look at the flat, two-dimensional phase space $(\mathbb{R}^2, \omega_0)$ with the canonical symplectic form ω_0 from (4.5). We will also identify the tangent space $T_{\boldsymbol{x}}\mathbb{R}^2$ and the cotangent space $T^*_{\boldsymbol{x}}R^2$ with $\mathbb{R}^2$. If we now take a function $f \in C^\infty(\mathbb{R}^2, \mathbb{R})$, the differential $\mathrm{d}f_{\boldsymbol{x}}$ at the point $\boldsymbol{x} \in \mathbb{R}^2$ is essentially given by the *gradient* of f (for this reason we will call $\mathrm{d}f^\sharp$ the ***symplectic gradient*** of f):

$$\mathrm{d}f_{\boldsymbol{x}} = \frac{\partial f}{\partial x^i}(\boldsymbol{x})\,\mathrm{d}x^i|_{\boldsymbol{x}} = \big(\nabla f(\boldsymbol{x})\big)_i\,\mathrm{d}x^i|_{\boldsymbol{x}} \,,$$

only the basis vectors are here the basis 1-forms $\mathrm{d}x^i$ of the cotangent bundle $T^*\mathbb{R}^2$.

We know that the gradient points in the direction where f has the greatest increase rate and hence at each point is *orthogonal* to the level surfaces of f. The 'sharpening' of

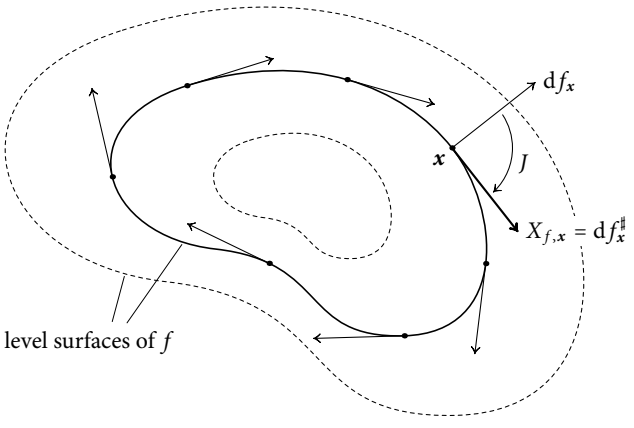


Figure 4.1 Hamiltonian vector field X_f associated to the function f. The matrix J is the matrix representation of the canonical symplectic form ω_0 and acts on $\mathrm{d}f_x$ like a rotation by an angle of $-\pi/2$. (In the figure, both $T_{\boldsymbol{x}}\mathbb{R}^2$ and $T^*_{\boldsymbol{x}}\mathbb{R}^2$ are identified with $\mathbb{R}^2$).

$\mathrm{d}f$ can then be thought of as a *rotation* by an angle of $-\pi/2$ (best seen by a close look at the matrix representation J of ω_0) followed by replacing the basis 1-forms of $T^*\mathbb{R}^2$ by the basis vector fields of $T\mathbb{R}^2$, such that the Hamiltonian vector field X_f afterwards is at each point *parallel* to the level surface of f (see figure 4.1). So, as a result, the value of the function f is *conserved* along the flow curves of the Hamiltonian vector field X_f.

Now remember that Hamilton's *canonical equations*, which describe physical motion in Hamiltonian mechanics, are given by:

$$\dot{q}^i(t) = +\frac{\partial H}{\partial p_i}(q,p,t)\,, \quad \dot{p}_i(t) = -\frac{\partial H}{\partial q^i}(q,p,t)\,, \tag{4.12}$$

where H is the *Hamiltonian function* of the dynamical system in question. If we write $\xi(t) = (q(t), p(t))$, the two equations (4.12) can be combined to give:

$$\dot{\xi}(t) = X_H|_{\xi(t)}\,. \tag{4.13}$$

Hence, the canonical equations say that physical motion happens along integral curves ξ of X_H. In the light of the last remark we can immediately conclude that, because the physical motion is generated by X_H, the Hamiltonian function H is always conserved in Hamiltonian systems.

Lemma 4.2.4 (Properties of Hamiltonian vector fields)

(i) Given any vector field $Y \in \mathfrak{X}(M)$, we have:

$$\omega(X_f, Y) = \mathrm{d}f(Y) = Y(f) = \mathcal{L}_Y f\,. \tag{4.14}$$

(ii) A vector field is *locally Hamiltonian* if and only if it is *symplectic* (therefore any globally Hamiltonian vector field is also symplectic, i. e. $\mathfrak{X}_{\mathrm{Ham}}(M) \subseteq \mathfrak{X}_{\mathrm{Sp}}(M)$).

(iii) A diffeomorphism $\varphi : M \to N$ between symplectic manifolds is *symplectic* if and only if:

$$\varphi^{\#} X_f = X_{\varphi^* f}\,, \tag{4.15}$$

for all functions $f \in C^\infty(N, \mathbb{R})$.

(iv) The Lie bracket of two symplectic vector fields $X, Y \in \mathfrak{X}_{\mathrm{Sp}}(M)$ is a globally Hamiltonian vector field:

$$[X,Y] = \mathrm{d}\big(-\omega(X,Y)\big)^\sharp\,, \tag{4.16}$$

associated to the function $-\omega(X,Y) \in C^\infty(M, \mathbb{R})$.

Proof. (i) Equation (4.14) is a simple consequence of the definition, since $X_f = \mathrm{d}f^\sharp$ is equivalent to ${X_f}^\flat = \mathrm{d}f$ and thus $\omega(X_f, Y) = {X_f}^\flat(Y) = \mathrm{d}f(Y) = Y(f) = \mathcal{L}_Y f$ for every vector field $Y \in \mathfrak{X}(M)$.

(ii) Lemma 4.1.5 shows that $\mathcal{L}_X \omega = 0$ is equivalent to $\mathrm{d}(\mathrm{i}_X \omega) = 0$. This is exactly what we need, because X locally Hamiltonian is defined as $X^\flat = \mathrm{i}_X \omega$ being closed.

(iii) By using the naturality of the exterior derivative $\mathrm{d} \circ \varphi^* = \varphi^* \circ \mathrm{d}$ (lemma 2.8.4) and the global definition of the pullback according to equation (2.27) one obtains:

$$\omega_M(X_{\varphi^* f}, Y) = \mathrm{d}_M(\varphi^* f)(Y) = (\varphi^* \mathrm{d}_N f)(Y) = \mathrm{d}_N f(\varphi_\# Y) \circ \varphi = \omega_N(X_f, \varphi_\# Y) \circ \varphi\,,$$

for every vector field $Y \in \mathfrak{X}(M)$. Now the diffeomorphism φ is symplectic if and only if $\varphi^*\omega_N = \omega_M$, or – to express it via the inverse φ^{-1} – if $\varphi_*\omega_M = \omega_N$. In this case we can continue according to:

$$\omega_N(X_f, \varphi_\# Y) \circ \varphi = \varphi_*\omega_M(X_f, \varphi_\# Y) \circ \varphi = \omega_M(\varphi^\# X_f, \varphi^\#\varphi_\# Y) = \omega_M(\varphi^\# X_f, Y) \,,$$

where we used $\varphi^\# = (\varphi^{-1})_\#$, so $\varphi^\#\varphi_\# = \mathrm{id}$. We thus have shown that φ is symplectic if and only if $\omega(X_{\varphi^* f}, Y) = \omega(\varphi^\# X_f, Y)$ for all $Y \in \mathfrak{X}(M)$, which is equivalent to equation (4.15) because the symplectic form ω is nondegenerate.

(iv) Consider symplectic vector fields $X, Y \in \mathfrak{X}_{\mathrm{Sp}}(M)$, that is $\mathcal{L}_X\omega = \mathcal{L}_Y\omega = 0$. According to lemma 2.8.7 we have:

$$\begin{aligned}[X,Y]^\flat = \mathrm{i}_{[X,Y]}\omega &= (\mathcal{L}_X \circ \mathrm{i}_Y - \mathrm{i}_Y \circ \mathcal{L}_X)\omega = \mathcal{L}_X(\mathrm{i}_Y\omega) - \mathrm{i}_Y(\mathcal{L}_X\omega) = \mathcal{L}_X(\mathrm{i}_Y\omega) \\ &= (\mathrm{d} \circ \mathrm{i}_X - \mathrm{i}_X \circ \mathrm{d})(\mathrm{i}_Y\omega) = \mathrm{d}\big(\omega(Y,X)\big) - \mathrm{i}_X \circ \mathrm{d} \circ \mathrm{i}_Y(\omega) = \mathrm{d}\big(-\omega(X,Y)\big) \,,\end{aligned}$$

where in the last step the second term vanishes because the formula $\mathcal{L}_Y = \mathrm{d} \circ \mathrm{i}_Y + \mathrm{i}_Y \circ \mathrm{d}$ implies $\mathrm{d} \circ \mathrm{i}_Y = \mathcal{L}_Y - \mathrm{i}_Y \circ \mathrm{d}$ and thus:

$$\mathrm{i}_X \circ \mathrm{d} \circ \mathrm{i}_Y(\omega) = \mathrm{i}_X \circ \mathcal{L}_Y\omega - \mathrm{i}_X(\mathrm{d}\omega) = \mathrm{i}_X(\mathcal{L}_Y\omega) = 0 \,,$$

given that ω is closed and Y is symplectic. Equation (4.16) follows by applying $\sharp$. ∎

Corollary 4.2.5 The Hamiltonian vector fields $\mathfrak{X}_{\mathrm{Ham}}(M)$ constitute a Lie algebra.

Proof. We have to show that $\mathfrak{X}_{\mathrm{Ham}}(M)$ is closed under the Jacobi–Lie bracket. This follows immediately from the last lemma, which states that Hamiltonian vector fields X, Y are symplectic, and hence $[X, Y]$ is Hamiltonian according to property (iv). ∎

We know from lemma 4.1.4 that a vector field is symplectic if and only if its flow Φ_t is a symplectomorphism for each t. Since Hamiltonian vector fields constitute a sub Lie algebra of the symplectic vector fields, this leads to the following definition:

Definition 4.2.6 (Hamiltonian symplectomorphism) Consider a symplectomorphism $\varphi \in \mathrm{Sp}(M)$ of the symplectic manifold (M, ω). The symplectomorphism is called a ***Hamiltonian symplectomorphism*** if it can be connected to the identity id_M via the flow Φ of a Hamiltonian vector field. The Hamiltonian symplectomorphisms constitute a subgroup $\mathrm{Ham}(M) \leq \mathrm{Sp}(M)$ of the symplectomorphisms of M.

Corollary 4.2.5 shows that the Lie bracket of Hamiltonian vector fields X_f and X_g is again Hamiltonian. Consequently, there has to exist a function $h \in C^\infty(M, \mathbb{R})$ such that $[X_f, X_g] = X_h$. Property (iv) of lemma 4.2.4 tells us how to determine this function: it is essentially the *Poisson bracket* of f and g, up to a minus sign.

Definition 4.2.7 (Poisson bracket) The ***Poisson bracket*** of two functions f and g on a symplectic manifold (M, ω) is given by:

$$\{f, g\} := \omega(X_f, X_g) = \mathrm{d}f(X_g) = -\mathrm{d}g(X_f) = \mathcal{L}_{X_g}(f) \,. \tag{4.17}$$

Note that some authors prefer the opposite convention, i. e. $\{f, g\} = \omega(X_g, X_f)$, which is the negative of our Poisson bracket.

The local coordinate expression in a Darboux chart is:

$$\{f, g\} = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q^i}\frac{\partial f}{\partial p_i} \,. \tag{4.18}$$

The smooth functions $C^\infty(M, \mathbb{R})$ compose a Lie algebra under the Poisson bracket. A subalgebra of them is used to represent the observables in classical mechanics.

Lemma 4.2.8 (Properties of the Poisson bracket) The Poisson bracket has the following properties:

(i) The relation between Jacobi–Lie bracket and Poisson bracket is:

$$[X_f, X_g] = -X_{\{f,g\}} \ . \tag{4.19}$$

(ii) $\{f,g\} = -\{g,f\}$, (*antisymmetry*)

(iii) $\big\{f,\{g,h\}\big\} = \big\{\{f,g\},h\big\} + \big\{g,\{f,h\}\big\}$, (*Jacobi identity*)

(iv) The Lie derivative along Hamiltonian vector fields is a *derivation* of the Poisson bracket:

$$\mathcal{L}_X\{f,g\} = \{\mathcal{L}_X f, g\} + \{f, \mathcal{L}_X g\} \ . \tag{4.20}$$

(v) A diffeomorphism $\varphi : M \to N$ between symplectic manifolds is *symplectic* iff:

$$\varphi^*\{f,g\} = \{\varphi^* f, \varphi^* g\} \ , \tag{4.21}$$

for all functions $f, g \in C^\infty(M,\mathbb{R})$.

Proof. (i) Equation (4.19) was actually the *motivation* to define the Poisson bracket, because it tells us to which function h the Hamiltonian vector field $[X_f, X_g] = X_h$ is associated. As such it is a consequence of equations (4.16) and (4.17).

(ii) The antisymmetry follows immediately from the definition.

(iii) For the third property we have:

$$\begin{aligned}\big\{f,\{g,h\}\big\} + \big\{g,\{h,f\}\big\} + \big\{h,\{f,g\}\big\} &= X_{\{g,h\}}(f) + X_g\big(X_h(f)\big) + X_h\big(X_g(f)\big) \\ &= X_{\{g,h\}}(f) + [X_g, X_h](f) = X_{\{g,h\}}(f) + X_{-\{g,h\}}(f) = 0 \ ,\end{aligned}$$

where we used (4.19). The Jacobi identity follows.

(iv) This follows from the Jacobi identity if we remember $\{f,g\} = \mathcal{L}_{X_f}(g)$ from the definition in equation (4.17).

(v) The last property follows from:

$$\begin{aligned}\varphi^*\{f,g\} = \varphi^*\big(\omega(X_f,X_g)\big) &= (\varphi^*\omega)(\varphi^\# X_f, \varphi^\# X_g) \overset{(*)}{=} \omega(\varphi^\# X_f, \varphi^\# X_g) \\ &= \omega(X_{\varphi^* f}, X_{\varphi^* g}) = \{\varphi^* f, \varphi^* g\} \ ,\end{aligned}$$

where the equality marked $(*)$ holds if and only if φ is symplectic. ∎

4.3 The Algebra of Classical Observables

The definition of the Poisson bracket in definition 4.2.7 implies that the mapping $f \mapsto X_f$ is an *anti*morphism between the Lie algebra of smooth functions on phase space $C^\infty(M,\mathbb{R})$ and the Lie algebra of Hamiltonian vector fields $\mathfrak{X}_{\mathrm{Ham}}(M)$, i. e.

$$[X_f, X_g] = -X_{\{f,g\}} \ .$$

Accordingly, we may define a new mapping j:

$$j : C^\infty(M,\mathbb{R}) \to \mathfrak{X}_{\mathrm{Ham}}(M) \ , \quad f \mapsto -X_f \ ,$$

to get a Lie algebra *morphism*. Since any Hamiltonian vector field stems from a function this morphism is clearly *surjective*. On the other hand, the mapping j is *not injective* because the formula $X_f = \mathrm{d}f^\sharp$ tells us that two functions that differ by a constant function produce the same Hamiltonian vector field. Identifying constant functions with $\mathbb{R}$ we thus obtain an isomorphism of Lie algebras:

$$\mathfrak{X}_{\mathrm{Ham}}(M) \cong C^\infty(M,\mathbb{R})/\mathbb{R}\,.$$

Regarding the discussion in section 3.6, the quotient relation can also be written as a short exact sequence of Lie algebras:

$$0 \longrightarrow \mathbb{R} \hookrightarrow C^\infty(M,\mathbb{R}) \xrightarrow{\ j\ } \mathfrak{X}_{\mathrm{Ham}}(M) \longrightarrow 0\ . \tag{4.22}$$

This sequence plays a fundamental role in the quantization programme via the Canonical Group.

One more thing about classical observables is noteworthy. Although we already know the smooth functions $C^\infty(M,\mathbb{R})$ on phase space together with the Poisson bracket form a Lie algebra, they have some additional structure: two smooth functions f and g can be multiplied with each other to give a new function $fg \in C^\infty(M,\mathbb{R})$. The multiplication of functions is *commutative*, and the combination of a Lie algebra with a commutative algebra – subject to a little compatibility condition – is known as a *Poisson algebra*.

Definition 4.3.1 (Poisson algebra) A ***Poisson algebra*** $\mathcal{A}$ is a Λ-module $\mathcal{A}$ over a commutative ring Λ with identity together with two Λ-bilinear products $\cdot : \mathcal{A}\times\mathcal{A}\to\mathcal{A}$ (multiplication) and $[\,\cdot\,,\,\cdot\,] : \mathcal{A}\times\mathcal{A}\to\mathcal{A}$ (bracket), such that:

- $(\mathcal{A},\cdot)$ is a commutative algebra,
- $(\mathcal{A},[\,\cdot\,,\,\cdot\,])$ is a Lie algebra,
- The bracket is a derivation of the commutative algebra $(\mathcal{A},\cdot)$, i. e.

$$[A, B\cdot C] = [A,B]\cdot C + B\cdot[A,C]\ , \tag{4.23}$$

for all $A, B, C \in \mathcal{A}$.

Lemma 4.3.2 The smooth functions $C^\infty(M,\mathbb{R})$ endowed with Poisson bracket and the common multiplication of functions constitute a real Poisson algebra.

Proof. Since we already know that the Poisson bracket gives $C^\infty(M,\mathbb{R})$ the structure of a Lie algebra we only have to check that (4.23) holds. This follows from:

$$\{f, gh\} = -X_f(gh) = -(X_f g)\,h - g\,(X_f h) = \{f,g\}\,h + g\,\{f,h\}\ ,$$

where we used that a vector field acts on functions like a derivation. ∎

It is this *commutative multiplication* of classical observables, compared to the noncommutative multiplication of operators, that is the most visible difference between the mathematical structures underlying classical mechanics and quantum mechanics. The Lie bracket, on the other hand, is *preserved* by most quantization methods (a famous exception being *deformation quantization*, where the Lie algebra structures have to match only in the limit $\hbar \to 0$).

4.4 Symplectic Structure of the Cotangent Bundle

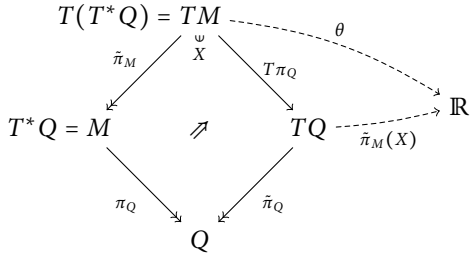
MOST SYSTEMS usually discussed in classical mechanics use as phase space the cotangent bundle T^*Q over the configuration space Q. The surprising fact is that, unlike generic fibre bundles, the cotangent bundle comes with a *canonical* symplectic structure. Moreover, the symplectic form ω of the cotangent bundle is not only closed but *exact*. It can thus be derived from a 1-form which is known as *Liouville form*.

Definition 4.4.1 (Liouville 1-form) Given a smooth manifold Q, the cotangent bundle $M = T^*Q$ supports a canonical 1-form $\theta \in \Omega^1(M)$, called the ***Liouville form***:

$$\theta(X) := \tilde{\pi}_M(X)\big(T\pi_Q(X)\big) \quad \forall X \in TM\,, \tag{4.24}$$

Note that $\theta \in \Omega^1(M)$ means it is a section $T^*Q \to T^*(T^*Q)$. The *double* cotangent structure can be confusing at first sight, yet it is this feature which gives us a *canonical* 1-form on M.

where the projections, as well as the induced mapping $T\pi_Q$, are given in the following diagram:



The projection $\tilde{\pi}_M$ assigns the basepoint $m \in M$ to a given point $X = (m, X_m) \in TM$, that is $\tilde{\pi}_M : X = (m, X_m) \mapsto m$. Because M is the cotangent bundle T^*Q, this *point* $m = \tilde{\pi}_M(X) \in M$ can be reinterpreted as a *cotangent vector* in T^*_qQ (with $q = \pi_Q(m)$), which then acts on the tangent vector $T\pi_Q(X) \in T_{\pi_Q(m)}Q$ via the *natural pairing* between T^*_qQ and T_qQ (i. e. $(\alpha, Y) \mapsto \alpha(Y)$ for $\alpha \in T^*_qQ$ and $Y \in T_qQ$).

Mind the subtlety of notation: a point $X \in TM$ is split (by a local trivialisation of the bundle) into its *basepoint* $m \in M$ and its *value* $X_m \in T_mM$. The same is done below for $m = (q, m_q)$.

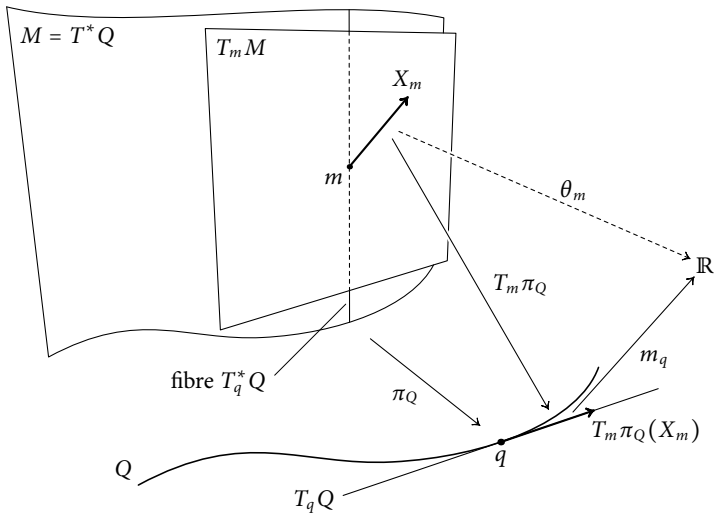


Figure 4.2 Pointwise definition of the Liouville 1-form θ. For convenience we use the abbreviation $q = \pi_Q(m) \in Q$.

The situation becomes clearer if we evaluate equation (4.24) at a given point $m \in M$. In this case, the projection $\tilde{\pi}_M$ can be eliminated from the formula and we obtain the following expression (which can also serve as the definition of θ):

$$\theta_m(X_m) = m_q\big(T_m\pi_Q(X_m)\big) \quad \forall X_m \in T_mM\,, \tag{4.25}$$

with $q = \pi_Q(m) \in Q$. Note how $m = (q, m_q)$ on the left-hand side is again seen as a *point* in $M = T^*Q$, whereas on the right-hand side its 'value' m_q is a *cotangent vector* acting on the tangent vector $T_m\pi_Q(X_m) \in T_qQ$ at the point q.

Instead of using a local trivialisation to write $m = (q, m_q)$, we could also represent m by the value $m = \alpha(q) \in T^*Q$ of a 1-form $\alpha : Q \to T^*Q$. In principle we then would have to show that the calculation is independent of the choice of α.

After this invariant definition of the Liouville form we want to derive a local coordinate expression. Before we can do this, we first need to construct a Darboux chart on the cotangent bundle. Incidentally, a Darboux chart of T^*Q can be induced by a given chart of the configuration space Q, where we again use the 'schizophrenic nature' of points in T^*Q/cotangent vectors on Q.

Remark 4.4.2 (Canonical Darboux charts of the cotangent bundle) Any chart (U, x) of the configuration space manifold Q induces a chart $\big(T^*U, (q, p)\big)$ of the cotangent bundle $M = T^*Q$ due to:

$$q^i(m) := x^i \circ \pi_Q(m)\,, \quad p_i(m) := m_{\pi_Q(m)}\Big(\frac{\partial}{\partial x^i}\Big|_{\pi_Q(m)}\Big)\,, \tag{4.26}$$

where $\pi_Q : T^*Q \to Q$ is the bundle projection and the tangent vectors $\partial/\partial x^i|_{\pi_Q(m)}$ form a basis of $T_{\pi_Q(m)}Q$. It is clear that (4.26) defines a *chart* of T^*Q. We will see that it is a *Darboux chart* when we calculate the coordinate expression of the symplectic 2-form ω associated to θ.

A basis of the tangent space T_mM (tangent to M) is given by the $2n$ tangent vectors:

$$\frac{\partial}{\partial q^i}\Big|_m\,,\ \frac{\partial}{\partial p_j}\Big|_m \in T_mM\,.$$

It turns out that the tangent map $T_m\pi_Q$ maps the first n basis vectors $\partial/\partial q^i|_m$ of T_mM to the basis vectors $\partial/\partial x^i|_{\pi_Q(m)}$ of $T_{\pi_Q(m)}Q$, whereas the remaining ones are mapped to zero:

$$T_m\pi_Q\Big(\frac{\partial}{\partial q^i}\Big|_m\Big) = \frac{\partial}{\partial x^i}\Big|_{\pi_Q(m)}\,, \quad T_m\pi_Q\Big(\frac{\partial}{\partial p_i}\Big|_m\Big) = 0\,. \tag{4.27}$$

This follows from $\partial q^j/\partial q^i = \delta^{ij}$ and $\partial q^j/\partial p_i = 0$, respectively, if you look closely at the coordinate representation of the tangent map, given in equation (2.9), where we remember $q^i = x^i \circ \pi_Q$. A proof by explicit calculation can be found in [Walo7, lemma 3.2.2].

Lemma 4.4.3 (Coordinate representation of θ) Given a chart $\big(T^*U, (q, p)\big)$ as in equation (4.26), the local coordinate representation of θ is:

$$\theta|_{T^*U} = p_i\,\mathrm{d}q^i + 0\,. \tag{4.28}$$

All the terms proportional to $\mathrm{d}p^i$ vanish.

Proof. A general 1-form $\theta \in \Omega^1(M)$ locally has the form:

$$\theta|_{T*U} = \theta^q_i\,\mathrm{d}q^i + \theta^i_p\,\mathrm{d}p_i\,,$$

with coefficient functions $\theta_i^q = \theta(\partial/\partial q^i)$ and $\theta_p^i = \theta(\partial/\partial p_i)$. Using equation (4.27) we obtain:

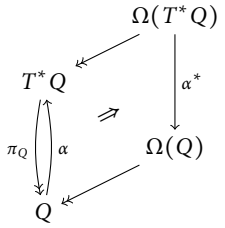
$$\theta_i^q(m) = \theta_m\Big(\frac{\partial}{\partial q^i}\Big|_m\Big) = m\Big(\frac{\partial}{\partial x^i}\Big|_{\pi_Q(m)}\Big) = p_i(m)\,, \quad \theta_p^i(m) = \theta_m\Big(\frac{\partial}{\partial p_i}\Big|_m\Big) = 0\,.$$

This gives the coordinate form stated in (4.28). ∎

Lemma 4.4.4 (Universal property of θ) The Liouville form $\theta \in \Omega^1(T^*Q)$ is uniquely determined by the universal property that for any 1-form $\alpha \in \Omega^1(Q)$ it fulfils:

$$\alpha^*\theta = \alpha\,, \tag{4.29}$$

where $\alpha^* : \Omega(T^*Q) \to \Omega(Q)$ is the pullback along $\alpha : Q \to T^*Q$ seen as a section.



Proof. Suppose α is a 1-form in $\Omega^1(Q)$. We then have:

$$(\alpha^*\theta)_q(X_q) = \theta_{\alpha(q)}(T_q\alpha(X_q)) = \alpha_q\big(T_{\alpha(q)}\pi_Q \circ T_q\alpha(X_q)\big) = \alpha_q\big(T_q(\pi_Q \circ \alpha)(X_q)\big)\,,$$

for any tangent vector $X_q \in T_qQ$. Further, since α is a section, we obtain $\pi_Q \circ \alpha = \mathrm{id}_Q$ and thus $T_q(\pi_Q \circ \alpha)(X_q) = X_q$. This proves that θ as defined above fulfils equation (4.29) for every $\alpha \in \Omega^1(Q)$.

Although uniqueness of θ can now be shown by direct calculation, there is a much simpler way (due to [AM78, proposition 3.2.11]). First observe that equation (4.29) is equivalent to:

$$\theta_{\alpha(q)}\big(T_q\alpha(X_q)\big) \stackrel{!}{=} \alpha_q(X_q)\,,$$

for all $\alpha \in \Omega^1(Q)$ and all $X_q \in T_qQ$ for each $q \in Q$. It is then easy to see this equation fixes θ uniquely, because $\alpha(q)$ spans the whole cotangent space T_q^*Q for variable α, and $T_q\alpha(X_q)$ spans the whole tangent space $T_{\alpha(q)}(T^*Q)$ for variable α and variable X_q. ∎

For the direct calculation note that (4.26) implies $q^i \circ \alpha = x^i$ and $p_i \circ \alpha = \alpha_i$; afterwards use the coordinate expression for $T_q\alpha$ from equation (2.9).

Definition 4.4.5 (Canonical 2-form on T^*Q) The *canonical symplectic structure* of the cotangent bundle T^*Q is given by the ***canonical* 2-*form*** $\omega \in \Omega^2(T^*Q)$:

$$\omega := -\mathrm{d}\theta\,, \tag{4.30}$$

where θ is the Liouville 1-form from above.

Because of lemma 4.4.3, the coordinate expression for ω is:

$$\omega|_{T^*U} = -\mathrm{d}(p_i\,\mathrm{d}q^i) = \mathrm{d}q^i \wedge \mathrm{d}p_i\,, \tag{4.31}$$

which finally proves that equation (4.26) defines a Darboux chart on T^*Q.

Furthermore, the definition of ω in equation (4.30) together with the explicit expression (2.35) for the exterior derivative gives:

$$\begin{aligned}\omega(X,Y) &= -\mathrm{d}\theta(X,Y) = -X\big(\theta(Y)\big) + Y\big(\theta(X)\big) - \theta\big([X,Y]\big)\\ &= Y\big(\theta(X)\big) - \mathcal{L}_X\theta(Y) = \mathrm{d}\big(\theta(X)\big)(Y) - \mathcal{L}_X\theta(Y)\,,\end{aligned}$$

so we obtain the useful formula:

$$X^\flat = \mathrm{i}_X\omega = \mathrm{d}(\mathrm{i}_X\theta) - \mathcal{L}_X\theta\,. \tag{4.32}$$

4.5 Symplectic Group Actions on the Cotangent Bundle

In this section we present two important constructions for symplectomorphisms of cotangent bundles: the *cotangent lift* of diffeomorphisms and *fibre translations* using 1-forms. Together, they provide a natural source for symplectic group actions on T^*Q. The resulting group $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ features prominently in the quantization programme via the Canonical Group.

Definition 4.5.1 (Cotangent lift) Suppose Q and S are manifolds and let $\varphi : Q \to S$ be a diffeomorphism. The ***cotangent lift*** $T^*\varphi : T^*S \to T^*Q$ of φ is then given by:

$$\big(T^*_s\varphi(\alpha_s)\big)(v_q) := \alpha_s\big(T_q\varphi(v_q)\big) \quad \text{and} \quad q := \varphi^{-1}(s)\,, \tag{4.33}$$

for all points $\alpha = (s,\alpha_s) \in T^*S$, and $T^*\varphi|_{T_sS} := T^*_s\varphi$. Moreover:

$$T_*\varphi := T^*\varphi^{-1} : T^*Q \to T^*S\,.$$

At first, equation (4.33) looks suspiciously like the pullback φ^* of differential forms (as in definition 2.7.8). However, where the pullback acts on *sections* $\varphi^* : \Omega(S) \to \Omega(Q)$, the cotangent lift $T^*\varphi : T^*S \to T^*Q$ is a ***point transformation*** between cotangent bundles. The relation between them is analogous to the relation between pushforward of vector fields $\varphi_\# : \mathfrak{X}(Q) \to \mathfrak{X}(S)$ and tangent map $T\varphi = \varphi_* : TQ \to TS$. We will explain this in a second. Before, we need to clarify some properties:

Lemma 4.5.2 (Properties of the cotangent lift)

(i) The cotangent lift $T^*\varphi$ of $\varphi : Q \to S$ is a *vector bundle isomorphism* covering the map φ^{-1}, i. e. the following diagram commutes:

$$\begin{array}{ccc} T^*S & \xrightarrow{T^*\varphi} & T^*Q \\ {\scriptstyle\pi_S}\downarrow & & \downarrow{\scriptstyle\pi_Q} \\ S & \xrightarrow{\varphi^{-1}} & Q \end{array}\,. \tag{4.34}$$

(ii) For $\varphi : Q \to S$ and $\psi : S \to R$ we have the *chain rule*:

$$T^*(\psi\circ\varphi) = T^*\varphi \circ T^*\psi\,. \tag{4.35}$$

Moreover:

$$T^*\,\mathrm{id}_Q = \mathrm{id}_{T^*Q}\,. \tag{4.36}$$

Accordingly, the mapping $T^* : \mathrm{Man}^k \to \mathrm{VB}$ is a *contravariant functor* from the category Man^k of C^k-manifolds to the category VB of vector bundles.

(iii) The cotangent lift is a *symplectomorphism* of cotangent bundles. If ω_Q and ω_S are the canonical symplectic forms on T^*Q and T^*S, respectively, this means:

$$\begin{array}{ccc} \Omega(T^*S) & \xleftarrow{(T^*\varphi)^*} & \Omega(T^*Q) \\ \downarrow & \Uparrow & \downarrow \\ T^*S & \xrightarrow{T^*\varphi} & T^*Q \\ {\scriptstyle\pi_S}\downarrow & \Uparrow & \downarrow{\scriptstyle\pi_Q} \\ S & \xleftarrow{\varphi} & Q \end{array}$$

$$(T^*\varphi)^*\omega_Q = \omega_S\,.$$

Even stronger, the cotangent lift preserves the Liouville 1-forms:

$$(T^*\theta)^*\theta_Q = \theta_S\,.$$

(Analogous properties hold for $T_*\varphi$, only the map goes in the opposite direction.)

Proof. (i) That $T^*\varphi$ is a vector bundle morphism covering φ^{-1}, in particular that it is fibre-preserving, follows from the definition, because $T^*\varphi$ is already *given* by its restriction to fibres: $T^*_s\varphi : T_sS \to T_{\varphi^{-1}(s)}Q$. Moreover, we can see from the chain rule (which we will prove next) that the inverse to $T^*\varphi$ is the map $T_*\varphi = T^*\varphi^{-1}$. Hence $T^*\varphi$ is a vector bundle *iso*morphism.

(ii) The chain rule as well as $T^*\,\mathrm{id}_Q = \mathrm{id}_{T^*Q}$ come from the corresponding properties of the tangent map $T\varphi$ (see lemma 2.3.4).

(iii) We first show that the cotangent lift preserves the Liouville 1-forms. To prove this, let us write $\psi := T^*\varphi$ and denote the Liouville form on T^*Q by $\theta = \theta_Q$. Then, for any point $\alpha = (s, \alpha_s) \in T^*S$ and any $X_\alpha \in T_\alpha(T^*S)$:

$$(\psi^*\theta)_\alpha(X_\alpha) = \theta_{\psi(\alpha)}\big(T_\alpha\psi(X_\alpha)\big) = \big(\psi(\alpha)\big)_{\pi_Q\circ\psi(\alpha)}\big(T_{\psi(\alpha)}\pi_Q \circ T_\alpha\psi(X_\alpha)\big)\,,$$

where we first used the definition 2.7.8 of the pullback, then the definition of θ. The next step is to reinsert $\psi(\alpha) = T^*\varphi(\alpha_s)$ into the equation and apply the formula (4.33) for the cotangent lift, so:

$$\begin{aligned}\big(\psi(\alpha)\big)_{\pi_Q\circ\psi(\alpha)}\big(T_{\psi(\alpha)}\pi_Q \circ T_\alpha\psi(X_\alpha)\big) &= \big(T^*_s\varphi(\alpha_s)\big)_{\pi_Q\circ\psi(\alpha)}\big(T_\alpha(\pi_Q\circ\psi)(X_\alpha)\big)\\ &= \alpha_{\varphi\circ\pi_Q\circ\psi(\alpha)}\big(T(\varphi\circ\pi_Q\circ\psi)(X_\alpha)\big)\,.\end{aligned}$$

Now, observe that $\varphi\circ\pi_Q\circ\psi = \pi_S$ holds because of the commutativity of diagram (4.34). This finally leaves us with:

$$(\psi^*\theta)_\alpha(X_\alpha) = \alpha_{\varphi\circ\pi_Q\circ\psi(\alpha)}\big(T(\varphi\circ\pi_Q\circ\psi)(X_\alpha)\big) = \alpha_{\pi_S(\alpha)}\big(T_\alpha\pi_S(X_\alpha)\big) = \theta_{S,\alpha}(X_\alpha)\,,$$

where, in the last equality, we used the definition of the Liouville form θ_S on T^*S. Hence $\psi^*\theta_Q = \theta_S$, so the Liouville 1-forms are preserved.

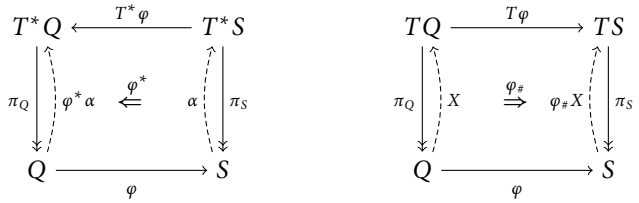
Given $\omega_Q = -\mathrm{d}\theta_Q$ and $\omega_S = -\mathrm{d}\theta_S$ this further implies:

$$\omega_S = -\mathrm{d}\theta_S = -\mathrm{d}(\psi^*\theta_Q) = \psi^*(-\mathrm{d}\theta_Q) = \psi^*\omega_Q\,,$$

because pullback and exterior derivatives commute with each other. The cotangent lift $\psi = T^*\varphi$ is thus a symplectomorphism from T^*S to T^*Q. ∎

Waldmann [Wal07, Satz 3.2.11] also shows the reverse: Any diffeomorphism $\psi : T^*S \to T^*Q$ that preserves the Liouville 1(!)-forms is the cotangent lift of some diffeomorphism $\varphi : Q \to S$ (however, this is not the case if ψ only preserves the symplectic 2-forms).

Now back to the promised relation between pullback and cotangent lift. Compare diagram (4.34) for the cotangent lift $T^*\varphi$ with the corresponding diagram for the tangent map $T\varphi$ from definition 2.3.1:



As already mentioned, both tangent map $T\varphi$ and cotangent lift $T^*\varphi$ are vector bundle morphisms, covering φ and φ^{-1}, respectively. In contrast, the pullback φ^* and the push-forward of vector fields $\varphi_\#$ act on sections. The big difference when a section is relocated is that not only the *values* have to be transformed but also the *basepoints*. Hence, a 1-form $\alpha \in \Omega^1(S)$ is pulled back to a form in $\Omega^1(Q)$ via:

$$\varphi^*\alpha = T^*\varphi \circ \alpha \circ \varphi\,, \quad \text{or pointwise:} \quad (\varphi^*\alpha)_q = T^*\varphi(\alpha_{\varphi(q)})\,.$$

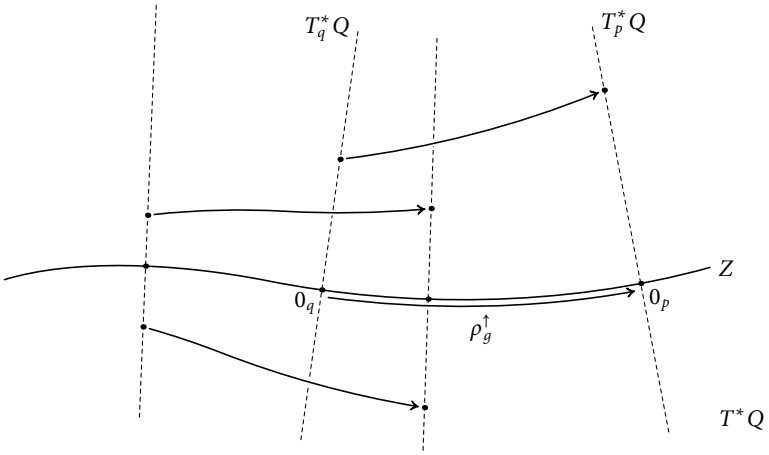


Figure 4.3 Cotangent lift $\rho^\uparrow$ of a group action ρ. Note that Z (the embedding of Q via the zero section $Q \to T^*Q\,,\ q \mapsto (q, 0_q)$) is necessarily an orbit of the lifted group action $\rho^\uparrow$. Hence the action cannot be transitive.

A vector field is pushed forward according to (this is exactly definition 2.4.6):

$$\varphi_\# X = T\varphi \circ X \circ \varphi^{-1}\,, \quad \text{or pointwise:} \quad (\varphi_\# X)_s = T\varphi(X^{-1}_\varphi(s))\,.$$

This explains again why the *pushforward of vector fields* can only be defined if the map φ is *invertible*, whereas the *pullback* of differential forms doesn't require this condition.

The contangent lift can be used to *lift* a given group action on the base manifold Q to a group action on the contangent bundle T^*Q.

$$\begin{array}{ccc} T^*Q & \xrightarrow{T_*\rho_g} & T^*Q \\ {\scriptstyle \pi_Q}\downarrow & \Uparrow & \downarrow{\scriptstyle \pi_Q} \\ Q & \xrightarrow{\rho_g} & Q \end{array}$$

Definition 4.5.3 (Cotangent lift of a group action) Let $\rho : G \to \mathrm{Diff}(Q)$ be a smooth action of some Lie group G acting on the manifold Q. The ***cotangent lift of*** ρ is a group action $\rho^\uparrow$ on the cotangent bundle, given by:

$$\rho^\uparrow : G \to \mathrm{Sp}(T^*Q)\,, \quad g \mapsto \rho^\uparrow_g := T_*\rho_g = T^*\rho_{g^{-1}}\,. \tag{4.37}$$

Lemma 4.5.4 Given a smooth left (right) action ρ of G on Q, the cotangent lift $\rho^\uparrow$ is a *symplectic* left (right) action of G on the cotangent bundle T^*Q.

Proof. The 'symplectic' part is clear because the cotangent lift of a diffeomorphism of Q is a symplectomorphism of T^*Q.

Furthermore, if ρ is a *left* action then $\rho^\uparrow$ is a *left* action as well, due to:

$$\rho^\uparrow_{gh} = T_*\rho_{gh} = T_*(\rho_g \circ \rho_h) = T_*\rho_g \circ T_*\rho_h = \rho^\uparrow_g \circ \rho^\uparrow_h\,,$$

where we used the chain rule from lemma 4.5.2 (ii), only the additional inverse φ^{-1} in $T_*\varphi = T^*\varphi^{-1}$ reverses the order of terms. Analogously, the lifted group action $\rho^\uparrow$ is a *right* action whenever ρ is a *right* action. ∎

Unfortunately, the cotangent lift of a group action is *never transitive*. The reason is simple: the cotangent lift of a diffeomorphism $\varphi \in \mathrm{Diff}\, Q$ is a *linear* map between the fibres. It always maps the zero in one cotangent space to the zero in another cotangent

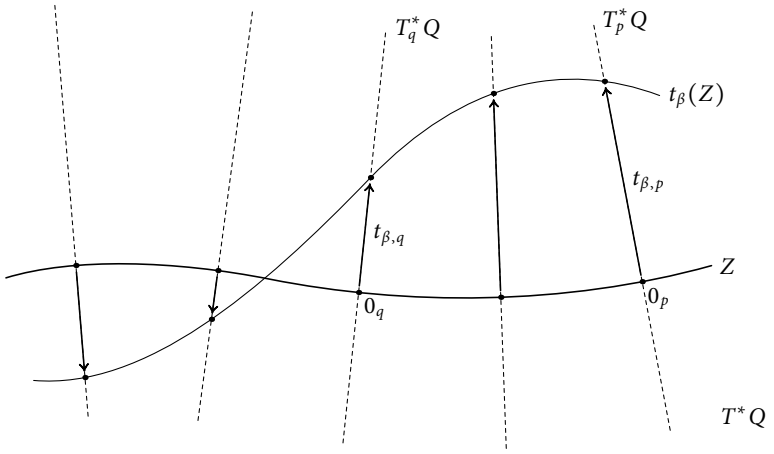


Figure 4.4 Fibre translation by a 1-form β. All points of the zero section Z are shifted to points in $\tau_\beta(Z)$ by the fibre translation τ_β. The orbits of τ_β are fibres of T^*Q.

space (see figure 4.3). The image of the zero section, $Z = \{(q, 0_q) : q \in Q\}$, is thus an *orbit* of $\rho^\uparrow$. Since $Z \subset T^*Q$ is a proper subset, the group action is not transitive.

Accordingly, to obtain a *transitive* group action on T^*Q we have to augment the transformations *between* fibres by some transformations *along* the fibres.

Definition 4.5.5 (Fibre translation) Let $\beta \in \Omega^1(Q)$ be a 1-form. The ***fibre translation*** (or ***momentum shift***) ***by*** β is the diffeomorphism:

$$\tau_\beta : T^*Q \to T^*Q\,, \quad \alpha \mapsto \alpha + \beta\big(\pi_Q(\alpha)\big)\,, \tag{4.38}$$

The fibre translation τ_β is a *symplectomorphism* if and only if β is *closed* (see lemma 4.5.8).

where α is a point in T^*Q. If we write $q = \pi_Q(\alpha)$ this yields the fibrewise expression:

$$\tau_{\beta,q} := \tau_\beta|_{T^*_q Q} : T^*_q Q \to T^*_q Q\,, \quad \alpha_q \mapsto \alpha_q + \beta_q\,,$$

where $\alpha_q \in T^*_q Q$ is a cotangent vector and $\beta_q \in T^*_q Q$ is the value of the 1-form β at the point $q \in Q$ (see also figure 4.4). The fibre translation is well-defined because T^*Q is a vector bundle, and hence elements in the same fibre can be added to each other.

Before we come to the properties of τ_β we want to define a related concept – the *vertical lift* of β to a vector field on T^*Q.

Definition 4.5.6 (Vertical lift) Let $\beta \in \Omega^1(Q)$ be a 1-form. The ***vertical lift*** of β is the vector field $V^\beta \in \mathfrak{X}(T^*Q)$ that generates the fibre translation τ_β, i. e. pointwise:

$$(V^\beta)_{\alpha_q} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \tau_{t\beta,q}(\alpha_q) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\alpha_q + t\beta_q)\,, \tag{4.39}$$

for each $\alpha_q \in T^*Q$. The map $\Phi(m, t) := \tau_{t\beta}(m)$ is thus the (complete) *flow* of V^β.

The name 'vertical lift' comes from the fact that V^β is a vector field with values in the *vertical bundle* $V(T^*Q)$, which is a subbundle of the tangent bundle $T(T^*Q)$. The definition is as follows:

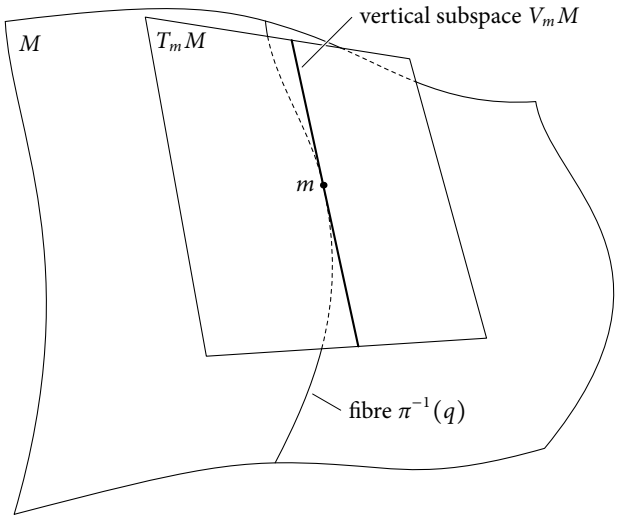


Figure 4.5 The *vertical subspace* V_mM is the subspace of T_mM which is *tangent* to the fibre containing m (i. e. the fibre $\pi^{-1}(q)$ for $q = \pi(m)$).

This definition works for *any* fibre bundle M, not just for vector bundles.

Note that there is, in general, no canonical choice for a complementary *horizontal* subspace.

Definition 4.5.7 (Vertical bundle) Given a fibre bundle $\pi : M \to Q$, the ***vertical bundle*** VM over M is the subbundle of the tangent bundle TM that consists of vectors *tangent to the fibres* of M (see figure 4.5). More precisely, for each $m \in M$ the ***vertical subspace*** $V_mM \subseteq T_mM$ is given by:

$$V_mM := \{X_m \in T_mM : T_m\pi(X_m) = 0\}\,, \tag{4.40}$$

i. e. V_mM is the kernel of the map $T_m\pi : T_mM \to T_{\pi(m)}Q$. The vertical bundle is then the disjoint union of all vertical subspaces.

Lemma 4.5.8 (Properties of fibre translations) Let $\pi : T^*Q \to Q$ denote the cotangent bundle $M = T^*Q$ over Q.

(i) The fibre translations are *fibre preserving*, i. e. $\pi \circ \tau_\beta = \pi$ for each $\beta \in \Omega^1(Q)$, and satisfy:

$$\tau_{r\beta} \circ \tau_{s\gamma} = \tau_{r\beta+s\gamma}\,,$$

for all $r, s \in \mathbb{R}$ and $\beta, \gamma \in \Omega^1(Q)$.

(ii) The fibre translation τ_β (the vertical lift V^β) is *symplectic* if and only if β is *closed*.

(iii) The vertical lift V^β is *Hamiltonian* if and only if β is *exact*. In this case, that is if $\beta = \mathrm{d}f$ for some function $f \in C^\infty(Q, \mathbb{R})$, we have:

$$V^\beta = -X_{\pi^* f}\,. \tag{4.41}$$

Proof. (i) Both properties are immediate consequences of the definition.

(ii) First, it follows from lemma 4.1.5 that the fibre translation τ_β is a symplectic transformation if and only if the vertical lift V^β is a symplectic vector field because $\tau_{t\beta}(m)$ is the flow of V^β.

Next, we want to check under which conditions τ_β is a symplectomorphism. Let us calculate the pullback of the Liouville form θ. At the point $\alpha = (q, \alpha_q) \in T^*Q$ and for

any tangent vector $X_\alpha \in T_\alpha M$ we have:

$$\begin{aligned}(\tau_\beta{}^*\theta)_\alpha(X_\alpha) &= \theta_{\tau_\beta(\alpha)}\big(T_\alpha\tau_\beta(X_\alpha)\big) = \tau_\beta(\alpha)\big(T_{\tau_\beta(\alpha)}\pi \circ T_\alpha\tau_\beta(X_\alpha)\big)\\ &= \tau_\beta(\alpha)\big(T_\alpha(\pi\circ\tau_\beta)(X_\alpha)\big) = \tau_\beta(\alpha)\big(T_\alpha\pi(X_\alpha)\big)\,,\end{aligned}$$

where we used the property $\pi \circ \tau_\beta = \pi$ from part (i). When inserting the definition of τ_β one obtains:

$$\tau_\beta(\alpha)\big(T_\alpha\pi(X_\alpha)\big) = \alpha\big(T_\alpha\pi(X_\alpha)\big) + \beta\big(T_\alpha\pi(X_\alpha)\big) = \theta_\alpha(X_\alpha) + \pi^*\beta(X_\alpha)\,,$$

so $\tau_\beta{}^*\theta = \theta + \pi^*\beta$. Applying the exterior derivative (which commutes with pullbacks) yields the pullback of the symplectic form $\omega = -\mathrm{d}\theta$:

$$\tau_\beta{}^*\omega = \omega - \pi^*(\mathrm{d}\beta)\,.$$

The fibre translation τ_β is thus symplectic if and only if the obstructing term on the right-hand side vanishes. This is equivalent to β being closed, $\mathrm{d}\beta = 0$.

(iii) We already remarked that the vertical lift V^β produces a vertical vector field, that is $T\pi(V^\beta) = 0$. Given how θ is defined, this implies $\mathrm{i}_{V^\beta}\theta = \theta(V^\beta) = 0$. Further, we can use equation (4.32) to obtain:

$$(V^\beta)^\flat = \mathrm{i}_{V^\beta}\omega = -\mathcal{L}_{V^\beta}\theta + \mathrm{d}(\mathrm{i}_{V^\beta}\theta) = -\mathcal{L}_{V^\beta}\theta = -\pi^*\beta\,,$$

where the last equality follows if we remember that $\tau_{t\beta}(m)$ is the flow of V^β, and hence:

$$\mathcal{L}_{V^\beta}\theta = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\tau_\beta{}^*\theta = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big(\theta + t(\pi^*\beta)\big) = \pi^*\beta\,,$$

where we reused $\tau_\beta{}^*\theta = \theta + \pi^*\beta$ from part (ii). Thus V^β is Hamiltonian if and only if $\pi^*\beta$ is *exact*, which is equivalent to β being exact because pullback π^* and exterior derivative commute. ∎

Remark 4.5.9 (Fundamental vector fields and vertical lift) It turns out that the vertical lift V^β is the *fundamental vector field* associated to β.

First, note that for a 'flat' Abelian Lie group G – an Abelian Lie group where the underlying manifold is a vector space V – the associated Lie algebra $\mathcal{L}G$ is isomorphic to $T_0V \cong V$ endowed with the trivial Lie bracket $[\,\cdot\,,\,\cdot\,] = 0$. Accordingly, the exponential map $\exp : \mathcal{L}G \to G$ is just the identity $\exp = \mathrm{id}_V$.

In case of $\Omega^1(Q)$, what was written above as $t\beta$ is thus actually $t\beta = \exp(t\beta) \in G$, for β seen as an element in the Lie algebra $\mathcal{L}G = \Omega^1(Q)$. The fibre translations then define a *group action* of G on M by:

$$\tau : G \to \mathrm{Diff}(M)\,, \quad t\beta = \exp(t\beta) \mapsto \tau_{t\beta} = \tau_{\exp(t\beta)}\,.$$

where the group homomorphism property of τ is part (i) of the last lemma. With this in mind, if you now compare definition 4.5.6 of the vertical lift V^β to the definition 3.5.1 on page 81 of a fundamental vector fields, it follows that the vertical lift V^β is also the *fundamental vector field* $V^\beta = \zeta_\tau(\beta)$ with respect to τ.

If we want the fundamental vector fields V^β to be Hamiltonian – as we do in the quantization scheme via the Canonical Group – lemma 4.5.8 tells us that we have to use

exact 1-forms. Since β exact means that $\beta = \mathrm{d}f$ for some function $f \in C^\infty(Q, \mathbb{R})$, we can use the functions $C^\infty(Q, \mathbb{R})$ to describe exact 1-forms. There is only one complication: two functions f, g that differ only by a constant have the same exterior derivative and so produce the same 1-form $\beta = \mathrm{d}f = \mathrm{d}g$. This minor annoyance is easily resolved by passing over to the quotient $C^\infty(Q, \mathbb{R})/\mathbb{R}$.

Finally, the groups $\mathrm{Diff}(Q)$, acting via the cotangent lift, and $C^\infty(Q, \mathbb{R})/\mathbb{R}$, acting via fibre translations, can be combined to obtain a *transitive* group action on T^*Q. This is how the semidirect product $C^\infty(Q, \mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ enters the arena. The action of $C^\infty(Q, \mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ on the cotangent bundle T^*Q is given by:

The semidirect product was introduced in definition 3.6.1 on page 84.

$$\rho_{(f,\varphi)}(m) := T_*\varphi(m) - (\mathrm{d}f)_{\varphi\circ\pi(m)}\,, \tag{4.42}$$

for $f \in C^\infty(Q, \mathbb{R})/\mathbb{R}$, $\varphi \in \mathrm{Diff}(Q)$ and $m \in M$, resulting in the group law:

$$(f_2, \varphi_2)\cdot(f_1, \varphi_1) = \left(f_2 + f_1 \circ \varphi_2^{-1}, \varphi_2 \circ \varphi_1\right). \tag{4.43}$$

Properties of this group will be further discussed when we talk about the quantization on cotangent bundles in section 5.5.

5 Canonical Group Quantization

1. Write down the problem.
2. Think very hard.
3. Write down the answer.

Feynman Problem-Solving Algorithm, folklore

We already gave an outline of Canonical Group Quantization in section 1.8. The goal of this chapter is to put the still rather vague description on solid foundations and formulate the quantization method using the mathematical terms developed during the last few chapters. To this end, we will loosely follow the original presentation of Isham given in [Ish83], yet with an important difference:

A short overview of the full quantization procedure can also be found in [BRL10].

Unlike Isham, we make a sharp distinction between what we called the Geometric Group $\mathcal{G}$ and the Canonical Group $\mathcal{C}$. We will see that a Geometric Group equips the classical phase space with the structure of a $\mathcal{G}$-space. The Canonical Group $\mathcal{C}$, on the other hand, does not act on phase space but is related to the fundamental observables. Taking the difference seriously, we properly establish the relation between the groups $\mathcal{C}$ and $\mathcal{G}$ in section 5.3. The result is a conceptually much clearer construction of the Canonical Group. In particular, we will shed some light on the special case that requires a central Lie algebra extension, the one that Isham calls 'pathological'.

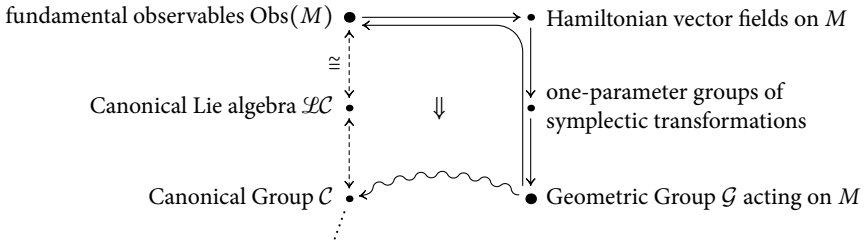
5.1 Mathematical Formulation of the Quantization Scheme

According to our discussion in section 1.8 the general quantization scheme consists of two main steps:

1. Find a Canonical Group $\mathcal{C}$ such that the associated set of fundamental quantizable observables is 'big enough',
2. Study strongly continuous unitary irreducible representations of $\mathcal{C}$; they give rise to self-adjoint operator representations of the algebra of fundamental observables.

The second step is primarily a mathematical problem in representation theory. For the examples we want to discuss in this chapter, the representation theory is well-known, so we will just point to the literature where necessary. The first step, on the other hand, is where physical considerations play a significant part, and it is this question of how to select a suitable Canonical Group $\mathcal{C}$ that lies at the heart of the quantization scheme.

In section 1.8 we already gave an outline of the idea behind the group-theoretical quantization method. Schematically, we had the following picture:



Roughly speaking, given a set of fundamental observables $\mathrm{Obs}(M)$ we can calculate the associated Hamiltonian vector fields and via their flow they serve as generators of one-parameter groups of symplectic transformations of M. All these one-parameter groups taken together constitute a Lie group, the Geometric Group $\mathcal{G}$, which canonically acts on classical phase space M.

On the other hand, if we are able to reverse the procedure it is possible to start with a Geometric Group $\mathcal{G}$ to obtain a set of fundamental quantizable observables $\mathrm{Obs}(M)$. Since the *geometry* and the global *topological* structure of the classical phase space largely restricts possible (and reasonable) symplectic transformations, we thereby obtain sets of fundamental observables reflecting the characteristics of the phase space M.

Before we explain how to select an appropriate Geometric Group $\mathcal{G}$, however, we need to show how Isham's idea is realised mathematically. The 'arena' is the following diagram:

In the following the Lie groups $\mathcal{C}$ and $\mathcal{G}$ are usually omitted. Here, the groups are included to help relate the diagram to the overview shown above.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \xrightarrow{\;j\;} & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0 \\ & & & & \uparrow & & \uparrow{\scriptstyle \gamma} & & \\ & & & & \mathcal{LC} & \rightsquigarrow & \mathcal{LG} & & \\ & & & & \uparrow{\scriptstyle \mathcal{L}} & & \uparrow{\scriptstyle \mathcal{L}} & & \\ & & & & \mathcal{C} & \rightsquigarrow & \mathcal{G} & & \end{array} \tag{5.1}$$

Almost all spaces in this diagram are Lie algebras, only $\mathcal{G}$ and $\mathcal{C}$ are Lie groups. The upper row is the short exact sequence (4.22) known from our discussion of the algebra of classical observables in section 4.3, where the Lie algebra morphism j assigns to each function the negative of its associated Hamiltonian vector field:

Remember that $f \mapsto X_f$ without the sign is a Lie algebra *anti*morphism (lemma 4.2.8).

$$j : C^\infty(M,\mathbb{R}) \to \mathfrak{X}_{\mathrm{Ham}}(M)\,, \quad f \mapsto -X_f\,. \tag{5.2}$$

It is a common convention to assume a left action of $\mathcal{G}$ on M, which makes $\gamma(A)$ an *antifundamental* vector field in the terminology of section 3.5. Nevertheless, remark 3.5.4 shows that $\gamma(A)$ is at the same time a *fundamental* vector field with respect to a right action. In the end, the important point is that γ in diagram (5.1) comes out as Lie algebra *morphism*, not as an antimorphism.

Furthermore, if we assume $\mathcal{G}$ acts on M via a *left* action $\rho : \mathcal{G} \times M \to M$ the map γ is a realisation of Lie algebra elements $A \in \mathcal{LG}$ in terms of their associated antifundamental vector fields $\gamma(A) \in \mathfrak{X}(M)$, and hence γ is a Lie algebra morphism as well (remark 3.5.4). Note that γ implicitly depends on the group action ρ; the definition of the antifundamental vector field $\gamma(A)$ from equation (3.25) shows this more clearly:

$$\gamma(A)_x = T_e\rho^x(A) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{\exp(-tA)}x \quad \forall x \in M\,. \tag{5.3}$$

The step from Geometric to Canonical Group is implemented in this diagram on the level of the corresponding Lie algebras instead of on the level of groups. The explicit construction of $\mathcal{LC}$ is a bit more involved since it depends on the existence of a specific right inverse. We will discuss the details in section 5.3.

It is important to understand how exactly diagram (5.1) reverses the construction of the Geometric Group in order to obtain the Canonical Group $\mathcal{C}$. To this end, remember that the idea in the forward direction was to begin with a complete set of fundamental observables $\mathrm{Obs}(M) \subseteq C^\infty(M,\mathbb{R})$ and assign Hamiltonian vector fields to them. This step is now accomplished by the Lie algebra morphism j, and we will denote the resulting subalgebra of the Hamiltonian vector fields by $\mathfrak{X}_{\mathrm{Obs}}(M) \subseteq \mathfrak{X}_{\mathrm{Ham}}(M)$:

$$\mathfrak{X}_{\mathrm{Obs}}(M) := j\big(\mathrm{Obs}(M)\big) = \mathrm{im}\, j|_{\mathrm{Obs}(M)}\,. \tag{5.4}$$

As lemma 2.4.9 guarantees, each Hamiltonian vector field in $\mathfrak{X}_{\mathrm{Obs}}(M)$ if it is complete will generate a one-parameter group of symplectic transformations. Moreover, all these individual one-parameter groups can be combined to give an abstract Lie group $\mathcal{G}$, the Geometric Group, that acts on the classical phase space M by means of a group action ρ, thanks to the theorem of Palais (see [Mico8, theorem 6.5]):

Theorem 5.1.1 (Palais) Let M be a smooth manifold and let $\varphi : \mathcal{L}G \to \mathfrak{X}(M)$ be a Lie algebra morphism from a finite dimensional Lie algebra $\mathcal{L}G$ into the Lie algebra $\mathfrak{X}(M)$ of vector fields on M such that each element in the image of φ is a *complete* vector field. Moreover, let G be a simply connected Lie group with Lie algebra $\mathcal{L}G$.

Then there exists a left action ρ of G on M whose associated antifundamental vector field mapping γ, given by equation (5.3), equals φ.

Remark 5.1.2 (Completeness of vector fields) From the perspective of Palais' theorem the reason why vector fields in the image of φ have to be complete is that a Lie group action ρ realises group elements $g \in \mathcal{G}$ in terms of *global* diffeomorphisms $\rho_g \in \mathrm{Diff}(M)$. Global diffeomorphisms, however, are generated only by vector fields with a global flow, that is, by *complete* vector fields.

Complete vector fields can be interpreted as *tangent vectors* to global diffeomorphisms if $\mathrm{Diff}(M)$ is seen as a Lie group (see also section 5.5).

Of course, the demand for complete Hamiltonian vector fields will in general not be fulfilled by arbitrary observables. Nevertheless, section 1.6 showed that self-adjoint operators are in one-to-one correspondence with strongly continuous one-parameter *groups* of unitary transformations, as a consequence of Stone's theorem 1.6.15. Therefore, only classical observables that generate *groups* of transformations are of interest to the quantization procedure, and for these classical observables the completeness assumption is satisfied.

Using Palais' theorem, the construction of the Geometric Group $\mathcal{G}$ is summarised by the following diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \xrightarrow{j} & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0 \\ & & & & \cup\mathrm{I} & & \cup\mathrm{I} & & \\ & & & & \mathrm{Obs}(M) & \xrightarrow[\tilde{j}]{} & \mathfrak{X}_{\mathrm{Obs}}(M) & & \\ & & & & \wr\uparrow & & \downarrow & & \\ & & & & \mathcal{LC} & & \text{theorem of Palais} & & \\ & & & & & & \mathcal{G} & & \end{array} \qquad (5.5)$$

Moreover, the argument in the forward direction shows that the choice of an algebra of fundamental observables $\mathrm{Obs}(M)$ renders classical phase space into a $\mathcal{G}$-space in the sense of definition 3.2.1.

Now, to reverse the steps leading from $\mathfrak{X}_{\mathrm{Obs}}(M)$ to $\mathcal{G}$, first recall from lemma 3.4.2 that the one-parameter subgroups of $\mathcal{G}$ are exactly the integral curves through $e \in \mathcal{G}$ of left invariant vector fields. Accordingly, we can express any one-parameter subgroup α of $\mathcal{G}$ via the exponential map $\exp : \mathcal{L}\mathcal{G} \to \mathcal{G}$ as $\alpha(t) = \exp(tA)$ using a tangent vector $A \in T_e\mathcal{G} \cong \mathcal{L}\mathcal{G}$, where A generates the one-parameter group α.

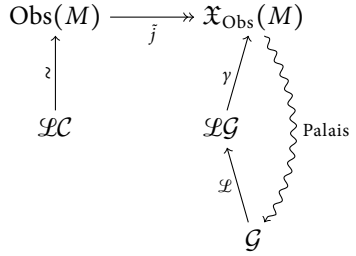
The group action ρ realises this 'abstract' one-parameter group $t \mapsto \alpha(t)$ in terms of the 'concrete' one-parameter group $t \mapsto \rho_{\alpha(t)}$ of symplectic transformations on phase

space (see also figure 3.5 on page 82). The generator of this one-parameter group of symplectomorphisms is the fundamental vector field $\zeta(A)$ as its flow Φ is given by $\Phi(t,x) = \rho_{\alpha(t)}x$ for all $t \in \mathbb{R}$ and all $x \in M$.

Another convention which can be found in the literature is to use ζ and to replace j by the antimorphism $f \mapsto X_f$ (or use the opposite sign convention for the Poisson bracket). Using morphisms, however, is more convenient for our purposes.

In the end, diagram (5.1) simply introduces an additional minus sign to replace the Lie algebra antimorphism ζ by the more convenient Lie algebra morphism γ. The following diagram displays both directions combined:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \xrightarrow{j} & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0 \\ & & & & \cup\mathrm{I} & & \cup\mathrm{I} & & \\ & & & & \mathrm{Obs}(M) & \xrightarrow[\tilde{j}]{} & \mathfrak{X}_{\mathrm{Obs}}(M) & & \end{array} \quad . \tag{5.6}$$



So far, diagram (5.6) shows that the choice of an algebra of Hamiltonian vector fields $\mathfrak{X}_{\mathrm{Obs}}(M)$, generated by an algebra $\mathrm{Obs}(M)$ of fundamental observables, corresponds to singling out a specific $\mathcal{G}$-space structure on phase space.

5.2 Properties of the Geometric Group

Although the quantization method starts with a Geometric Group $\mathcal{G}$ in mind and determines a set of fundamental classical observables $\mathrm{Obs}(M)$ based on this group, the properties of $\mathcal{G}$ must be obtained by thinking in the opposite direction. To this end, let us consider an abridged version of diagram (5.1):

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \xrightarrow{j} & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0 \\ & & & & & & \uparrow{\scriptstyle \gamma} & & \\ & & & & & & \mathcal{LG} & & \end{array} \quad . \tag{5.7}$$

The first observation is that we used Hamiltonian vector fields associated to the fundamental observables to construct the Geometric Group $\mathcal{G}$ in the forward direction. Thus, by choice of an appropriate group action ρ we will have to guarantee that γ produces only **Hamiltonian** vector fields. Such group actions play an important role in the following, so they deserve a proper name:

One-parameter subgroups of G are realised via a Hamiltonian action by one-parameter groups of *Hamiltonian symplectomorphisms* (see definition 4.2.6).

Definition 5.2.1 (Hamiltonian group action) A group action $\rho : G \times M \to M$ of a Lie group G on a symplectic manifold (M,ω) is called ***Hamiltonian group action*** if the associated fundamental vector fields are Hamiltonian.

In general, we will have to check this property 'by hand'. Nevertheless, a necessary precondition is that $\mathcal{G}$ acts on M via a **symplectic** group action. In fact, lemma 4.1.7 showed that the fundamental vector fields associated to a symplectic group action are *symplectic* vector fields and the same holds for antifundamental vector fields.

Moreover, part (ii) of lemma 4.2.4 revealed that vector fields are symplectic if and only if they are *locally* Hamiltonian. While this local property is usually not enough to guarantee *globally* Hamiltonian vector fields its practical value lies in the fact that the local property, in contrast to the global property, is easily verified by calculations. With a little bit of intuition it is possible to guess reasonable, at least symplectic group actions if we simply remember that the symplectic volume element has to be preserved.

Remark 5.2.2 There are two special cases where the property of *locally* Hamiltonian is enough to imply *globally* Hamiltonian vector fields:

- If the first de Rham cohomology group of M is trivial, i. e. $H^1_{\mathrm{dR}}(M) = 0$, every *closed* 1-form is automatically *exact* according to lemma 2.8.15. As a vector field X is locally Hamiltonian if and only if the associated 1-form $X^\flat$ is closed (definition 4.2.2), this implies that $X^\flat$ will also be exact. Hence a locally Hamiltonian vector field X will automatically be globally Hamiltonian in this case.
- The second possibility comes from part (iv) of lemma 4.2.4, which states that the Lie bracket $[X, Y]$ of two locally Hamiltonian vector fields $X, Y \in \mathfrak{X}_{\mathrm{Sp}}(M)$ is globally Hamiltonian. Accordingly, if each element $A \in \mathcal{LG}$ can be written as a commutator of two other elements in $\mathcal{LG}$ the associated fundamental vector fields will be globally Hamiltonian. This is the case if $\mathcal{G}$ is a *semisimple* Lie group (see [FH91, section 9.1]; if $\mathcal{G}$ is a semisimple Lie group, the associated Lie algebra $\mathcal{LG}$ – which is then, too, called semisimple – can be written as $\mathcal{LG} = [\mathcal{LG}, \mathcal{LG}]$).

In addition to the Hamiltonian group action of $\mathcal{G}$, Isham argues in his notes that there is no particular benefit in considering disconnected Geometric Groups unless one wants to allow anti-symplectic transformations, corresponding to anti-unitary quantum operators [Ish83, section 4.3.1]. The reason is simply that the classical phase space M can usually be assumed to be a *connected* manifold and thus it makes sense for the Geometric Group $\mathcal{G}$ – which acts on M – to share this property. Unless stated otherwise, we will hence assume the Geometric Group $\mathcal{G}$ to be a **connected** Lie group.

There are two additional properties. First, the group action of $\mathcal{G}$ on M will have to be **transitive**. One reason to require a transitive action is that only in this case the Geometric Group is 'fully aware' of the global structure of phase space – that it 'feels out' the whole space M [Ish83, section 4.3.2]. In contrast, if the group action were not transitive, the phase space would decompose into different $\mathcal{G}$-orbits, but then the physical system would be restricted to a single orbit and we could have used this orbit as our phase space in the first place.

In particular, transitivity requires $\dim \mathcal{G} \geq \dim M$.

A related point is that a transitive group action is necessary to generate a *complete* set of fundamental quantizable observables. Given the intuition we have gained from Noether's theorem, a complete set of fundamental observables should correspond to a transformation group that covers all 'directions' in phase space. Isham discusses this more thoroughly in his notes and explains that a transitive group action is sufficient to guarantee a *local generating principle*, that is, there is always a neighbourhood in phase space such that any function $f \in C^\infty(M, \mathbb{R})$ can be written locally as a function of the fundamental observables (see [Ish83, section 4.3.4] for the details).

Remember that a complete set of quantizable observables is necessary to formulate an irreducibility condition for the quantization map (see also section 1.7).

Finally, the morphism γ must establish a one-to-one correspondence between $\mathcal{LG}$ and $\mathfrak{X}_{\mathrm{Obs}}(M)$. The reason is that the Geometric Group $\mathcal{G}$ was originally constructed in the opposite direction and hence it must be possible to *re-obtain* the algebra $\mathcal{LG}$ from the set of Hamiltonian vector fields in the image of γ (this also requires that $\operatorname{im}\gamma$ coincides

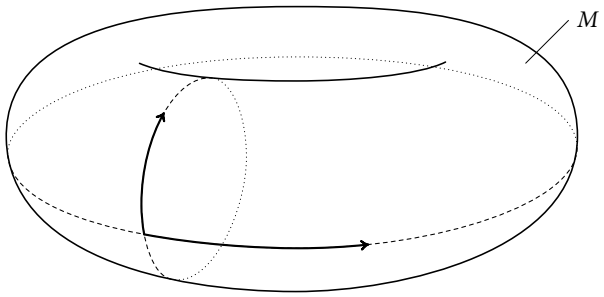


Figure 5.1 The torus T^2 is topologically the product of two circles $S^1 \times S^1$. This decomposition hints at a natural transformation group which can be used as Geometric Group $\mathcal{G}$ if it satisfies definition 5.2.3.

with $\mathfrak{X}_{\mathrm{Obs}}(M)$, that is, $\operatorname{im}\gamma = \mathfrak{X}_{\mathrm{Obs}}(M)$). Accordingly, γ as a map between $\mathfrak{L}\mathcal{G}$ and $\mathfrak{X}_{\mathrm{Obs}}(M)$ will have to be an isomorphism.

Surjectivity is guaranteed automatically if we define $\mathfrak{X}_{\mathrm{Obs}}(M) := \operatorname{im}\gamma$. To make γ injective, however, we have to require an **almost effective** group action of $\mathcal{G}$ on M. By definition 3.1.10, a group action ρ is *effective* if and only if the realisation $g \mapsto \rho_g$ is injective. Nevertheless, the explicit expression for γ, equation (5.3), uses the realisation only in a neighbourhood of the neutral element $e \in \mathcal{G}$ (since $\mathfrak{L}\mathcal{G} \cong T_e\mathcal{G}$). It is thus completely sufficient if the realisation of $\mathcal{G}$ is injective in a neighbourhood of e, and hence we need to require only an *almost effective* action.

Let us summarise all these properties in the following definition:

Definition 5.2.3 (Geometric Group) Let (M, ω) be a symplectic manifold and let $\mathcal{G}$ be a Lie group that acts on M via a smooth group action $\rho : \mathcal{G} \times M \to M$. We will call $(\mathcal{G}, \rho)$ a ***Geometric Group*** if the following conditions hold:

- $\mathcal{G}$ is a *connected* Lie group,
- ρ is a *Hamiltonian* group action in the sense of definition 5.2.1 (as a precondition, ρ must be a *symplectic* action),
- ρ is *transitive*,
- ρ is *almost effective*.

Remark 5.2.4 (Phase space as $\mathcal{G}$-space) The classical phase space M together with the structure imposed by a Geometric Group $(\mathcal{G}, \rho)$ is a $\mathcal{G}$-space in the sense of section 3.2. Since M is now in addition a *symplectic manifold*, we naturally require a Lie group $\mathcal{G}$ which acts on M in a way that is compatible with the symplectic structure.

While the definition looks quite restrictive, fact is that it doesn't determine $\mathcal{G}$ uniquely. The usual way to select a suitable Geometric Group is hence to look for a transformation group that arises 'naturally' from the structure of M. In the common case where M is a cotangent bundle $M = T^*Q$, for instance, an obvious choice is to use the *cotangent lift* from section 4.5 to induce symplectic transformations on M based on diffeomorphisms of the underlying configuration space Q and subsequently augment them by *fibre translations* to obtain a transitive group action. In other cases, it might be possible to exploit a product structure in order to obtain a natural group action, for example for the torus shown in figure 5.1.

The special case of quantization on cotangent bundles is the topic of section 5.5.

That being said, in addition to pure mathematical suitability, different Geometric groups may also correspond to different physical situations. We know that a momentum observable arises from *translations* in position space whereas angular momentum is related to *rotations*. The fact which observables better describe the classical system at hand thus enters into the choice of Geometric Group. In fact, it may in general be that an important classical observable is quantizable only in terms of a specific Geometric Group but not in terms of another, so that the choice of Geometric Group is a choice between *incompatible* quantizations. In addition, it may play a role for the problem at hand which variables *outside* the set of fundamental observables are quantizable.

Furthermore, it is important to observe that we cannot associate physical meaning to functions on phase space by considering them individually. Position and momentum, for example, are intertwined by the fact that they act on each other via translations. The physical interpretation of one observable will hence affect the possible interpretations of others. So, even if a certain observable can be quantized using several different Geometric groups, the physical interpretation of this observable may depend on what other observables are quantizable alongside.

5.3 From Geometric to Canonical Group

Let us assume that we have found a Geometric Group $\mathcal{G}$ that is compatible with a given classical phase space, at least from a mathematical point of view. How do we then proceed in order to obtain the Canonical Group?

As mentioned earlier, the quantization method does not construct the Canonical Group directly but uses a detour via the associated Lie algebras. Isham starts with the Lie algebra $\mathfrak{L}\mathcal{G}$ of the Geometric Group and finds that it must sometimes be 'extended' (via a central extension), and sometimes not. He provides a mathematical condition for when the extension is necessary, yet the physical meaning of the extension is left unclear. In fact, the central extension appears like a last resort for some 'pathological' situations – conventional quantum mechanics over $\mathbb{R}^n$ unfortunately being one of them. The specific use of a central extension to obtain $\mathfrak{L}\mathcal{C}$, however, is motivated by Isham only by the fact that a simpler type of extension doesn't work, which is not very satisfying.

5.3.1 The Relation Between $\mathfrak{L}\mathcal{C}$ and $\mathfrak{L}\mathcal{G}$

We will take a different perspective that provides a much clearer understanding. Right from the beginning, both Lie algebras have a distinct physical significance, even if they later might happen to coincide:

- $\mathfrak{L}\mathcal{C}$ abstractly describes the Lie algebra of the fundamental *observables*,
- $\mathfrak{L}\mathcal{G}$ describes the Lie algebra of the associated *Hamiltonian vector fields*.

The difference is that observables serve an additional purpose. They generate canonical transformations via their Hamiltonian vector fields, but they also serve as some kind of 'coordinates', especially with respect to measurements. More precisely, when the Poisson bracket is written as a Lie derivative:

$$\{f, g\} = \mathfrak{L}_{X_g} f \,,$$

we see that the Poisson bracket explicitly describes the *infinitesimal change of the f-values along the flow of the Hamiltonian vector field X_g* associated to the observable g. Since this

Seen like this, the coordinate change of f for *canonically conjugate observables*, with $\{f, g\} = 1$, just *matches* the flow of X_g, and vice versa. This is why conjugate observables 'fit together' exceptionally well.

second role as coordinate values is no longer explicitly contained in the algebra of the associated vector fields, it is not surprising that the algebra $\mathcal{LC}$ might in general need to be bigger than the algebra of the associated vector fields $\mathcal{LG}$.

In contrast to Isham, we will embrace the difference between $\mathcal{LC}$ and $\mathcal{LG}$, and try to determine their general relation. To this end, consider the following diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \xrightarrow{\ j\ } & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0\ . \\ & & & & \cup\mathrm{I} & & \cup\mathrm{I} & & \\ & & & & \mathrm{Obs}(M) & \xrightarrow{\ \tilde{j}\ } & \mathfrak{X}_{\mathrm{Obs}}(M) & & \\ & & & & \uparrow\wr & & \uparrow\wr\,\gamma & & \\ & & & & \mathcal{LC} & \overset{?}{\rightsquigarrow} & \mathcal{LG} & & \end{array} \tag{5.8}$$

The algebra $\mathrm{Obs}(M)$ of the fundamental observables is, of course, still unknown when we begin the quantization procedure with the Geometric Group $\mathcal{G}$. Nevertheless, the last section showed that γ must be an isomorphism between $\mathcal{LG}$ and $\mathfrak{X}_{\mathrm{Obs}}(M)$. In turn, we can thus use the image of γ to *define* the algebra $\mathfrak{X}_{\mathrm{Obs}}(M)$:

$$\mathfrak{X}_{\mathrm{Obs}}(M) := \operatorname{im}\gamma\ .$$

A recipe for the construction of $\mathrm{Obs}(M)$ and $\mathcal{LC}$ that arises from the following considerations is shown in figure 5.2.

Since our aim is to assign observables to the Hamiltonian vector fields in $\mathfrak{X}_{\mathrm{Obs}}(M)$, it makes sense to look for some kind of inverse to $\tilde{j}$. We already mentioned in section 4.3 that the 'full' mapping $j : C^\infty(M,\mathbb{R}) \to \mathfrak{X}_{\mathrm{Ham}}(M)$ isn't injective because two functions that differ only by a constant produce the same Hamiltonian vector field (remember $X_f = \mathrm{d}f^\sharp$). Nevertheless, not everything is lost. We only search a kind of inverse to the mapping $\tilde{j}$, for a subalgebra of fundamental observables $\mathrm{Obs}(M) \subseteq C^\infty(M,\mathbb{R})$.

If we assume, once again, that $\mathrm{Obs}(M)$ were already known, there are two cases that need to be distinguished. From the fact that the Lie algebra $\mathrm{Obs}(M)$ is also a vector space, we can infer that $\mathrm{Obs}(M)$ contains either *all* or *none* of the constant functions:

(i) If $\mathrm{Obs}(M)$ *doesn't contain* the constant functions, the mapping $\tilde{j}$ is *injective* and we hence have an isomorphism $\mathrm{Obs}(M) \cong \mathfrak{X}_{\mathrm{Obs}}(M)$, which gives $\mathcal{LC} \cong \mathcal{LG}$.

(ii) If $\mathrm{Obs}(M)$ *contains* the constant functions, the mapping $\tilde{j}$ is not injective. Still, since it then contains *all* constant functions the Lie algebra $\mathrm{Obs}(M)$ in this case must be a *central Lie algebra extension* of $\mathfrak{X}_{\mathrm{Obs}}(M)$ by $\mathbb{R}$ in the sense of definition 3.6.8 (this is not true in the first case!). We will see later that $\mathcal{LC}$ can under these circumstances be defined as a central Lie algebra extension of $\mathcal{LG}$.

Note that $\mathcal{LG}$ is no subalgebra of its extension $\mathcal{LC}$. Whenever the central extension is strictly necessary, $\mathcal{LG}$ cannot even be embedded into $\mathcal{LC}$!

The point is now that $\tilde{j}$ in the first case has an inverse. Of course, neither $\mathrm{Obs}(M)$ nor the restriction $\tilde{j}$ is known explicitly when we apply the quantization procedure. It is clear, however, that an invertible $\tilde{j}$ can only exist if there exists at least a *candidate* for an inverse. This is the case if we can find a (necessarily injective) Lie algebra morphism $\tilde{u} : \mathfrak{X}_{\mathrm{Obs}}(M) \to C^\infty(M,\mathbb{R})$ which satisfies $j \circ \tilde{u} = \tilde{j} \circ \tilde{u} = \mathrm{id}_{\mathfrak{X}_{\mathrm{Obs}}(M)}$:

While the explicit subalgebra $\mathrm{Obs}(M)$ is unknown at this point, we know that it must satisfy $\operatorname{im} u \subseteq \mathrm{Obs}(M)$. The precise form is given below.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \xrightarrow{\ j\ } & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0\ . \\ & & & & \cup\mathrm{I} & & \cup\mathrm{I} & & \\ & & & & \mathrm{Obs}(M) & \underset{\tilde{u}}{\overset{\tilde{j}}{\rightleftarrows}} & \mathfrak{X}_{\mathrm{Obs}}(M) = \operatorname{im}\gamma & & \\ & & & & \uparrow\wr & & \uparrow\wr\,\gamma & & \\ & & & & \mathcal{LC} & \overset{?}{\rightsquigarrow} & \mathcal{LG} & & \end{array}$$

If a morphism like $\tilde{u}$ exists we have two possibilities (Isham mentions only the first in his notes):

1. The obvious approach if $\tilde{u}$ exists is to define the algebra of fundamental quantizable observables as $\mathrm{Obs}(M) := \mathrm{im}\,\tilde{u}$. As a result, the mapping $\tilde{j}$ will be an isomorphism and we end up with the Lie algebra $\mathcal{LC} \cong \mathcal{LG}$.
2. While not required we can optionally include the constant functions by means of a central extension. Since constant functions Poisson-commute with everything they can be incorporated *trivially*, that is, we can define $\mathrm{Obs}(M) := \mathrm{im}\,\tilde{u} \oplus \mathbb{R}$ as a direct sum of Lie algebras. The algebra $\mathcal{LC} \cong \mathcal{LG} \oplus \mathbb{R}$ is thus a central extension of $\mathcal{LG}$ and the mapping $\tilde{u}$ becomes a right inverse of $\tilde{j}$, not a full inverse.

   Mathematically, the central extension in this case is, in addition, a *split extension* (definition 3.6.8), for which $\tilde{u}$ is the splitting map. The monomorphism $\tilde{u}$ embeds $\mathfrak{X}_{\mathrm{Obs}}(M)$ isomorphically into $\mathrm{Obs}(M)$. Although a split extension $\mathcal{LC}$ is a priori only a semidirect sum of Lie algebras, a split *central* extension of a Lie algebra is necessarily isomorphic to a *direct* sum of Lie algebras [Scho8, section 4.1]. Hence, the trivial extension is, up to isomorphism, the *only* possibility in this case.

The existence of this second possibility means that constant functions can *always* be quantized – in this case 'for free'.

Otherwise, a different course of action is necessary:

3. If no morphism like $\tilde{u}$ exists the algebra of fundamental observables $\mathrm{Obs}(M)$ must be a *non-splitting central extension* of $\mathfrak{X}_{\mathrm{Obs}}(M)$. The pre-images of the Hamiltonian vector fields in $\mathfrak{X}_{\mathrm{Obs}}(M)$ and the constant functions are in this case interwoven in such a non-trivial manner that no embedding of $\mathfrak{X}_{\mathrm{Obs}}(M)$ into $\mathrm{Obs}(M)$ exists.

   In contrast to the case for which a splitting map $\tilde{u}$ exists, the central extension is not optional but *strictly necessary* under these circumstances. Theorem 5.3.7, given below, provides the explicit expressions for $\mathrm{Obs}(M)$ and $\mathcal{LC}$ in this case.

That said, Isham states in his notes [Ish83, section 4.3.6] that it would be 'satisfying aesthetically if, in practice, physical systems of this type [for which a central extension is strictly necessary] never arose.' He also remarks that 'in the series of concrete examples [...] $Q = \mathbb{R}^n$ is the only "pathological" case [which requires a central extension] and in the rest (including quantum gravity!) the cocycle obligingly vanishes [no extension is necessary]'. Thus, it looks like $\mathbb{R}^n$ is 'special', but it is unclear why this is the case.

As we see it, given the relation between the two algebras $\mathrm{Obs}(M)$ and $\mathfrak{X}_{\mathrm{Obs}}(M)$ as established above, the central extension no longer appears to be 'pathological'. The connection via a central Lie algebra extension is actually the *natural* relation between observables and their associated Hamiltonian vector fields. Moreover, concerning the *necessity* to define $\mathrm{Obs}(M)$ via a central extension, the following statement holds:

Lemma 5.3.1 (Necessity of central extension)

(i) The necessity to define $\mathrm{Obs}(M)$ via a (non-splitting) central extension can only arise if constant functions are interwoven with the remaining fundamental observables in a non-trivial manner, that is, only if constant functions appear as (part of the) *result* of a Poisson bracket.

(ii) If *canonically conjugate observables* appear as fundamental quantizable observables, the algebra $\mathrm{Obs}(M)$ is necessarily a *non-splitting* central extension of $\mathfrak{X}_{\mathrm{Obs}}(M)$.

Proof of part (i). Constant functions commute with all observables due to the general definition of the Poisson bracket. Thus, constant function can only be an *essential* part of the Lie algebra $\mathrm{Obs}(M)$ if they appear as part of the *result* of a Poisson bracket of two

other observables, that is, in Poisson brackets of the form $\{f, g\} = h + c1$, with $f, g, h \in \mathrm{Obs}(M)$ and $c \in \mathbb{R} \setminus \{0\}$. Otherwise, constant functions are incorporated trivially and the central extension in this case is optional because a trivial extension splits. ■

The proof of the second statement in lemma 5.3.1 requires some basic knowledge of Lie algebra cohomology and will thus be postponed until the end of section 5.3.2.

Remark 5.3.2 (Limits of lemma 5.3.1; the Virasoro algebra) Part (i) of lemma 5.3.1 states that Poisson brackets of the form $\{f, g\} = h + c1$ with functions $f, g, h \in \mathrm{Obs}(M)$ and a constant $c \in \mathbb{R} \setminus \{0\}$ are a necessary precondition for $\mathrm{Obs}(M)$ to be a *non-splitting* central extension of $\mathfrak{X}_{\mathrm{Obs}}(M)$. An example is the Weyl–Heisenberg algebra known from section 1.6 which encodes the well-known Poisson brackets $\{q_i, p_j\} = \delta_{ij}$ associated to the configuration space $Q = \mathbb{R}^n$. The reverse statement, however, might not be true. If only the fixed combination $h + c1$ with some constant value of c appears as observable we can define $h' := h + c1$ and this new observable h' will in general not be a constant function. thus rendering a central extension unnecessary in this case.

Part (ii) of lemma 5.3.1, on the other hand, states that $\mathrm{Obs}(M)$ will *necessarily* be a non-splitting central extension when canonically conjugate observables are quantizable. Reading Isham's account, the Weyl–Heisenberg algebra also seems to be the *only* case for which a non-splitting central extension is required. Still, at least if we allow for infinite-dimensional Lie algebras we can provide a second example: the *Virasoro algebra*. The Lie brackets of the Virasoro algebra are given by:

$$[L_m, L_n] = (m - n)L_{m+n} + \tfrac{1}{12} n(n^2 - 1) Z \delta_{m+n,0} \,, \quad [L_n, Z] = 0 \,,$$

for all $m, n \in \mathbb{Z}$ (Z is an element of the centre), and therefore of the form required by part (i) of the lemma. These commutators show that none of the basis elements L_n and Z of the Virasoro algebra are canonically conjugate to each other. On the other hand, it is known that the Virasoro algebra is a non-trivial central extension of the *Witt algebra*, with commutators:

$$[L_m, L_n] = (m - n)L_{m+n} \,,$$

for all $m, n \in \mathbb{Z}$ (see [Scho8, chapter 5]). Thus, at least in the infinite-dimensional case, a non-splitting central extension does not require canonically conjugate observables.

5.3.2 Existence of a Splitting Map and Lie Algebra Cohomology

Let us now turn back to the explicit construction of the Canonical Group $\mathcal{C}$. A result that will be of great help to decide whether a Lie algebra morphism $\tilde{u}$ with the property $j \circ \tilde{u} = \mathrm{id}_{\mathfrak{X}_{Obs}(M)}$ exists is that j *always* has an (automatically injective) right inverse if we forget the Lie algebra structure and interpret j as a linear map between *vector spaces*. This is a corollary of a more general proposition. When we studied short exact sequences of groups in section 3.6, an important statement was the splitting lemma 3.6.6. A related result is that *every short exact sequence of projective modules splits* [HS71, theorem 4.7]. Therefore, since vector spaces are simple examples of projective modules, we always have the following split short exact sequence of *vector spaces*(!):

An analogous statement does not hold for Lie algebras. Otherwise, we would always end up in case (i).

$$0 \longrightarrow \mathbb{R} \hookrightarrow C^\infty(M, \mathbb{R}) \underset{u}{\overset{j}{\rightleftarrows}} \mathfrak{X}_{\mathrm{Ham}}(M) \longrightarrow 0 \,. \qquad (5.9)$$

linear map!

Given that j is not injective there is some freedom left in the specific choice of the right inverse u. Since the kernel of j is given by $\ker j \cong \mathbb{R}$ we can always add constant functions to the results of u and the new mapping will still be a right inverse of j. The question is whether this inherent freedom can be exploited to adjust the linear map u so that the restriction $u|_{\mathfrak{X}_{\mathrm{Obs}}(M)}$ becomes a Lie algebra morphism. In other words, we are trying to modify u in such a manner that the (necessarily injective) linear mapping $P := u \circ \gamma$ in the following diagram turns into a Lie algebra morphism (since γ defines an isomorphism between $\mathcal{LG}$ and $\mathfrak{X}_{\mathrm{Obs}}(M)$ the mapping P is a Lie algebra morphism if and only if this is the case for $u|_{\mathfrak{X}_{\mathrm{Obs}}(M)}$):

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(M,\mathbb{R}) & \underset{u}{\overset{j}{\rightleftarrows}} & \mathfrak{X}_{\mathrm{Ham}}(M) & \longrightarrow & 0 \\ & & & & & \overset{P}{\searrow} & \uparrow{\scriptstyle \gamma} & & \\ & & & & & & \mathcal{LG} & & \end{array} \quad , \tag{5.10}$$

that is:

$$\big\{P(A), P(B)\big\} = P\big([A,B]\big) \quad \forall A, B \in \mathcal{LG} \, . \tag{5.11}$$

Remark 5.3.3 (Sourieau's momentum map) The map P is closely related to Sourieau's *moment map* $J : M \to (\mathcal{LG})^*$, and sometimes P is called *comoment map*. Explicitly, the correspondence between P and J is given by $J(m)(A) = P(A)(m)$ for all $A \in \mathcal{LG}$ and for all $m \in M$ (see [Ish83, section 4.3.4]; more details about the momentum map can be found in [MR99, chapter 11], for example).

In order to study the situation more closely it is convenient to introduce the so-called ***obstruction cocycle*** z [Ish83, equation (4.3.17)]:

$$z(A,B) := \big\{P(A), P(B)\big\} - P\big([A,B]\big) \, , \quad A, B \in \mathcal{LG} \, . \tag{5.12}$$

Obviously, P is a Lie algebra morphism if and only if the obstruction cocycle z vanishes.

Roughly speaking, the cocycle z quantifies 'how much' P violates the equality in (5.11).

First, we will show that z really *is* a cocycle. To this end, notice that the value $z(A,B)$ of z for any choice of $A, B \in \mathcal{LG}$ is a *constant* function on M. This follows from:

$$\begin{aligned} -X_{P([A,B])} &= j \circ P([A,B]) = \gamma([A,B]) \\ &= [\gamma(A), \gamma(B)] = [X_{P(A)}, X_{P(B)}] \overset{(4.19)}{=} -X_{\{P(A),P(B)\}} \, , \end{aligned}$$

and the fact that $X_f = X_g$ implies $f = g + \text{const}$. Hence the obstruction cocycle z can be understood as a mapping:

$$z : \mathcal{LG} \times \mathcal{LG} \to \mathbb{R} \, , \quad (A,B) \mapsto z(A,B) \, .$$

Since P is required to be linear, z is actually a bilinear mapping. Moreover, z satisfies:

$$z(A,B) = -z(B,A) \, , \tag{5.13}$$

$$z\big(A,[B,C]\big) + z\big(B,[C,A]\big) + z\big(C,[A,B]\big) = 0 \, . \tag{5.14}$$

The first equation states that z is antisymmetric. Therefore, z can be read as a 2-form over the vector space $\mathcal{LG}$, that is, a 2-*cochain* $z \in \bigwedge^2(\mathcal{LG}^*)$ (the space $\mathcal{LG}^*$ is the dual of the Lie algebra $\mathcal{LG}$; it is instructive to compare this to the definition 2.7.1 of differential forms on manifolds). Furthermore, the following definition [Walo7, definition 3.3.43] helps to understand the second equation, (5.14):

Equation (5.15) looks a lot like the explicit expression for the exterior derivative, d, which is given in lemma 2.8.3.

Definition 5.3.4 (Chevalley–Eilenberg operator) Let $\mathfrak{L}G$ be a Lie algebra and consider the cochains $C^k(\mathfrak{L}G, \mathbb{R}) = \bigwedge^k(\mathfrak{L}G^*)$. For any k-cochain $\alpha \in \bigwedge^k(\mathfrak{L}G^*)$ and Lie algebra elements $A_0, \dots, A_k \in \mathfrak{L}G$ the ***Chevalley–Eilenberg operator*** δ_{CE} is given by:

$$\delta_{\text{CE}}\alpha(A_0, \dots, A_k) := \sum_{i<j}(-1)^{i+j}\alpha\big([A_i, A_j], A_0, \dots, \check{A}_i, \dots, \check{A}_j, \dots, A_k\big) \,. \tag{5.15}$$

Just like the exterior derivative in de Rham cohomology, the Chevalley–Eilenberg operator satisfies the properties of a coboundary operator, $\delta_{\text{CE}} \circ \delta_{\text{CE}} = 0$, and hence we get a cochain complex (definition 2.8.9):

$$\cdots \xrightarrow{\delta_{\text{CE}}} \bigwedge^{k-1}(\mathfrak{L}G^*) \xrightarrow{\delta_{\text{CE}}} \bigwedge^{k}(\mathfrak{L}G^*) \xrightarrow{\delta_{\text{CE}}} \bigwedge^{k+1}(\mathfrak{L}G^*) \xrightarrow{\delta_{\text{CE}}} \cdots$$

Continuing as in definition 2.8.12, we define k-cocycles $Z^k(\mathfrak{L}G, \mathbb{R})$ and k-coboundaries $B^k(\mathfrak{L}\mathcal{G}, \mathbb{R})$ as:

$$Z^k(\mathfrak{L}G, \mathbb{R}) := \ker \delta^{(k)}_{\text{CE}} \,, \quad B^k(\mathfrak{L}G, \mathbb{R}) := \operatorname{im} \delta^{(k-1)}_{\text{CE}} \,, \tag{5.16}$$

This is also known as *scalar* Lie algebra cohomology. In general, Lie algebra cohomology with coefficients in $\mathfrak{L}\mathcal{G}$-modules is defined using the Ext functor (see [HS71, section VII.2]).

as well as the cohomology classes $H^k_{\text{CE}}(\mathfrak{L}G, \mathbb{R})$ of ***Chevalley–Eilenberg cohomology***:

$$H^k_{\text{CE}}(\mathfrak{L}G, \mathbb{R}) := Z^k(\mathfrak{L}G, \mathbb{R})/B^k(\mathfrak{L}G, \mathbb{R}) \,. \tag{5.17}$$

Rewritten using the Chevalley–Eilenberg operator δ_{CE}, equation (5.14) becomes:

$$\delta_{\text{CE}} z = 0 \,, \tag{5.18}$$

and hence for any linear map P the mapping z really is a *2-cocycle* $z \in Z^2(\mathfrak{L}\mathcal{G}, \mathbb{R})$ with respect to the scalar Lie algebra cohomology of $\mathfrak{L}\mathcal{G}$, just as the name suggests.

The next step is to find out under what circumstances z can be made to vanish. We already explained that u remains a right inverse of j if we add arbitrary constants to its values. This implies that we can add an arbitrary constant $c_A \in \mathbb{R}$ to the value $P(A)$ for each $A \in \mathfrak{L}\mathcal{G}$. This freedom, however, is also the only possible modification as any other change would lead to an algebra $\text{Obs}(M)$ that no longer projects onto the given vector fields $\mathfrak{X}_{\text{Obs}}(M)$. In other words: every allowed mapping $\tilde{P}$ is related to the originally chosen mapping P via $\tilde{P}(A) = P(A) + h(A)$, where h is a function $h : \mathfrak{L}\mathcal{G} \to \mathbb{R}$ with values $h(A) := c_A$.

That said, a further restriction arises from the fact that the modified mapping $\tilde{P}$ must still be linear. As P is assumed linear, the function h must therefore really be a *linear* function on $\mathfrak{L}\mathcal{G}$. Hence, the maximal freedom in the choice of $\tilde{P}$ can be parameterised by an element of the dual Lie algebra $\alpha \in \mathfrak{L}\mathcal{G}^*$ according to:

$$\tilde{P}(A) := P(A) + \alpha(A) \,, \tag{5.19}$$

for every $A \in \mathfrak{L}\mathcal{G}$.

The obstruction cocycle $\tilde{z}$ associated to the modified mapping $\tilde{P}$ is:

$$\begin{aligned}
\tilde{z}(A, B) &= \big\{\tilde{P}(A), \tilde{P}(B)\big\} - \tilde{P}\big([A, B]\big) \\
&= \big\{P(A) + \alpha(A), P(B) + \alpha(B)\big\} - P\big([A, B]\big) - \alpha\big([A, B]\big) \\
&= \big\{P(A), P(B)\big\} - P\big([A, B]\big) - \alpha\big([A, B]\big) \\
&= z(A, B) - \alpha\big([A, B]\big) \,,
\end{aligned} \tag{5.20}$$

where we used in the next to last step that $\alpha(A)$ and $\alpha(B)$ are constant functions which Poisson-commute with everything. The result implies that the obstruction cocycles for arbitrarily chosen linear maps P and $\tilde{P}$ differ exactly by a term of the form $\alpha([A,B])$. In particular, the obstruction cocycle z can be made to vanish, i. e. $\tilde{z} = 0$, if and only if there exists some $\alpha \in \mathfrak{LG}^*$ such that z itself can be written as:

$$z(A,B) = \alpha\big([A,B]\big)\,, \tag{5.21}$$

for every $A, B \in \mathfrak{LG}$. Using the Chevalley–Eilenberg operator from above this equation becomes:

$$z(A,B) = \alpha\big([A,B]\big) = \delta_{\mathrm{CE}}\alpha(A,B)\,,$$

and hence we ultimately proved the following result:

Lemma 5.3.5 (Elimination of the obstruction cocycle) Given the notation from above, the obstruction cocycle $z \in Z^2(\mathfrak{LG},\mathbb{R})$, defined in terms of the mapping P according to equation (5.12), can be eliminated if and only if z is a *coboundary* $z \in B^2(\mathfrak{LG},\mathbb{R})$, that is, if and only if there exists a 1-form $\alpha \in \mathfrak{LG}^*$ such that:

In de Rham cohomology a coboundary is a closed form that is moreover *exact*.

$$z = \delta_{\mathrm{CE}}\alpha = \alpha\big([\,\cdot\,,\cdot\,]\big)\,. \tag{5.22}$$

If the obstruction cocycle z is of this form, an injective Lie algebra morphism from $\mathfrak{LG}$ to $C^\infty(M,\mathbb{R})$ is given by:

$$\tilde{P} : \mathfrak{LG} \to C^\infty(M,\mathbb{R})\,, \quad A \mapsto \tilde{P}(A) := P(A) + \alpha(A)\,. \tag{5.23}$$

Lemma 5.3.5 hands us a powerful method to check whether the original, at first only linear, mapping P can be adjusted to yield a Lie algebra morphism. It is even possible to make statements about whole classes of Lie algebras if we know something about their cohomology. If the cohomology class $H^2_{\mathrm{CE}}(\mathfrak{LG},\mathbb{R})$ is trivial, for example, every cochain is a coboundary and, consequently, the obstruction cocycle can always be made to vanish in this case. If the Geometric Group $\mathcal{G}$ is Abelian, on the other hand, its associated Lie algebra $\mathfrak{LG}$ is trivial, i. e. $[A,B] = 0$ for every $A, B \in \mathfrak{LG}$. Then there is no chance to adjust P and hence no Lie algebra morphism with the required properties unless P is a Lie algebra morphism in the first place.

As explained, the outcome determines how the quantization procedure continues. If the mapping P can be turned into a Lie algebra morphism we can define the Lie algebra of the fundamental observables by $\mathrm{Obs}(M) := \operatorname{im} P$. Since P is necessarily injective, the Lie algebra $\mathfrak{LG}$ is in this case *isomorphic* to $\mathrm{Obs}(M)$. Given $\mathfrak{LC} \cong \mathrm{Obs}(M)$, we therefore obtain $\mathfrak{LC} \cong \mathfrak{LG}$. It is moreover possible to choose $\mathfrak{LC} = \mathfrak{LG}$ without loss of generality, because the Canonical Lie algebra $\mathfrak{LC}$ is unique only up to an isomorphism. This yields the following diagram:

Remember that we can also choose to include the constant functions trivially in this case, just as explained in the beginning of this section.

$$\begin{array}{ccccc}
\operatorname{im} P =: \mathrm{Obs}(M) & & \xrightarrow{\;\tilde{j}\;} & & \mathfrak{X}_{\mathrm{Obs}}(M) := \operatorname{im}\gamma \\
\wr\uparrow & & \underset{\tilde{u}}{\overset{\sim}{\nwarrow}} & & \wr\uparrow\gamma \\
\mathfrak{LC} & & \xrightarrow[\mathrm{id}]{} & & \mathfrak{LG}
\end{array} \tag{5.24}$$

In short: if P can be adjusted to become a Lie morphism, we can choose the Canonical Group $\mathcal{C}$ equal to the Geometric Group $\mathcal{G}$ and our search is complete.

Although the *abstract* groups $\mathcal{G}$ and $\mathcal{C}$ can be chosen equal in this case, notice that they still bear different *meaning*!

Remark 5.3.6 (Proof of part (ii) of lemma 5.3.1) We are now in the position to prove the second part of lemma 5.3.1. To this end, remember that equation (5.20) showed that any two obstruction cocycles z and $\tilde{z}$ differ at most by a coboundary $\delta_{\mathrm{CE}}\alpha$. Therefore, the obstruction cocycles for all possible choices of the linear mapping P are part of the *same* cohomology class $[z] \in H^2_{\mathrm{CE}}(\mathcal{LG}, \mathbb{R})$. So, if a Lie algebra morphism like u exists, lemma 5.3.5 implies that *any* possible obstruction cocycle is a coboundary.

Now, consider canonically conjugate observables f and g which satisfy $\{f, g\} = 1$. Given that j is a Lie algebra epimorphism, we obtain:

$$0 = \mathrm{d}(1)^\sharp = j(1) = j\big(\{f, g\}\big) = \big[j(f), j(g)\big] = [-X_f, -X_g] = [X_f, X_g] = [A, B] \ ,$$

where the last step makes use of the definitions $A := \gamma^{-1}(X_f)$, $B := \gamma^{-1}(X_g)$ and the fact that γ is a Lie algebra isomorphism. Since j as a linear map is invertible, there exists a linear map P and constant functions c_A and c_B so that $f = P(A) + c_A$ and $g = P(B) + c_B$. The associated obstruction cocycle is:

$$z(A, B) = \big\{P(A), P(B)\big\} - P\big([A, B]\big) = \big\{f - c_A, g - c_B\big\} - P(0) = 1 \ .$$

By definition, the cocycle z is a coboundary if there exists some $\alpha \in \mathcal{LG}^*$ so that:

$$1 = z(A, B) = \delta_{\mathrm{CE}}\alpha(A, B) = \alpha\big([A, B]\big) \ ,$$

but $\alpha\big([A, B]\big) = \alpha(0) = 0$. As this is clearly a contradiction, the obstruction cocycle z is in this case *not* a coboundary. Therefore, a Lie algebra morphism like u does not exist and the central extension at hand is a *non-splitting* central extension, just as claimed.

5.3.3 The Case of a Non-Splitting Central Extension

If the linear mapping P cannot be turned into a Lie algebra morphism, we have to continue differently. We mentioned that this can only happen if the algebra $\mathrm{Obs}(M)$, which we try to determine, contains the constant functions in a non-trivial manner. Whenever P cannot be adjusted to become a Lie algebra morphism, the Lie algebra of the fundamental observables $\mathrm{Obs}(M)$ will be a non-splitting central Lie algebra extension of $\mathfrak{X}_{\mathrm{Obs}}(M)$ by $\mathbb{R}$. This means we have to look for a short exact sequence:

$$0 \longrightarrow \mathbb{R} \hookrightarrow \mathrm{Obs}(M) \xrightarrow{\ j\ } \mathfrak{X}_{\mathrm{Obs}}(M) \longrightarrow 0 \ ,$$

where $\mathbb{R}$ is embedded into the centre $\mathcal{Z}(\mathrm{Obs}(M))$.

Given the isomorphisms $\mathcal{LG} \cong \mathfrak{X}_{\mathrm{Obs}}(M)$ and $\mathcal{LC} \cong \mathrm{Obs}(M)$, we can extend this sequence into the following commutative diagram:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{R} & \hookrightarrow & \mathrm{Obs}(M) & \xrightarrow{\ j\ } & \mathfrak{X}_{\mathrm{Obs}}(M) & \longrightarrow & 0 \\
 & & & & \wr\uparrow & & \wr\uparrow \gamma & & \\
0 & \longrightarrow & \mathbb{R} & \hookrightarrow & \mathcal{LC} & \twoheadrightarrow & \mathcal{LG} & \longrightarrow & 0 \ ,
\end{array} \tag{5.25}$$

where the epimorphism between $\mathcal{LC}$ and $\mathcal{LG}$ is the epimorphism j composed with the two isomorphisms in the diagram. To determine $\mathrm{Obs}(M)$ we can thus construct a central extension $\mathcal{LC}$ of $\mathcal{LG}$ together with an Lie algebra monomorphism $\hat{P} : \mathcal{LC} \to C^\infty(M, \mathbb{R})$ for which the above diagram commutes if we define $\mathrm{Obs}(M) := \mathrm{im}\,\hat{P}$. This is accomplished by the following lemma:

Lemma 5.3.7 (Non-splitting central extension of $\mathcal{LG}$) Consider the case for which the map $P : \mathcal{LG} \to C^\infty(M, \mathbb{R})$ as given above is no Lie algebra morphism and for which the obstruction cocycle z cannot be made to vanish. A central extension of $\mathcal{LG}$ with the desired properties is then given by the direct sum of vector spaces:

$$\mathcal{LC} := \mathcal{LG} \oplus \mathbb{R}\,, \tag{5.26}$$

equipped with the Lie bracket:

$$\big[(A,\lambda),(B,\mu)\big] := \big([A,B]\,, z(A,B)\big)\,, \tag{5.27}$$

for all $A, B \in \mathcal{LG}$ and $\lambda, \mu \in \mathbb{R}$, where $[A,B]$ is the Lie bracket of $\mathcal{LG}$.

Observe how this definition of the Lie bracket leads to the fact that $\mathcal{LG}$ cannot be embedded isomorphically into $\mathcal{LC}$, unless z vanishes identically.

Furthermore, the mapping $\hat{P} : \mathcal{LC} \to C^\infty(M, \mathbb{R})$ given by $\hat{P}(A,\lambda) := P(A) + \lambda$ is an injective Lie algebra morphism with the property that diagram (5.25) commutes if we define $\mathrm{Obs}(M) := \operatorname{im}\hat{P}$.

Proof. To verify that $\mathcal{LC}$ really is a Lie algebra, observe that linearity and antisymmetry of the bracket (5.27) follow immediately because the Lie bracket of $\mathcal{LG}$ and the cocycle z both have these properties. Moreover, the Jacobi identity is satisfied, given:

$$\begin{aligned}\Big[(A,\lambda),\big[(B,\mu),(C,\nu)\big]\Big] + \text{cycl.} &= \Big[(A,\lambda),\big([B,C]\,, z(B,C)\big)\Big] + \text{cycl.}\\ &= \Big(\big[A,[B,C]\big]\,, z\big(A,[B,C]\big)\Big) + \text{cycl.} = 0\,,\end{aligned}$$

for all $A, B, C \in \mathcal{LG}$ and $\lambda, \mu, \nu \in \mathbb{R}$, where the last step employs the Jacobi identity for the Lie bracket of $\mathcal{LG}$ and the cocycle property (5.14) of z. In addition, the Lie algebra $\mathcal{LC}$ is a Lie algebra extension of $\mathcal{LG}$ because the following is a short exact sequence:

Note that the decomposition $\mathcal{LC} = \mathcal{LG} \oplus \mathbb{R}$ doesn't respect the Lie bracket. The sequence in general will not split!

$$0 \longrightarrow \mathbb{R} \xhookrightarrow{\iota_2} \mathcal{LC} = \mathcal{LG} \oplus \mathbb{R} \xrightarrow{\mathrm{pr}_1} \mathcal{LG} \longrightarrow 0\,,$$

where $\iota_2(\lambda) := (0,\lambda)$ is the natural embedding of $\mathbb{R}$ into $\mathcal{LC}$, and $\mathrm{pr}_1(A,\lambda) := A$ the natural projection of $\mathcal{LC}$ onto $\mathcal{LG}$. The extension is *central*, since $\operatorname{im}\iota_2$ lies in the centre of $\mathcal{LC}$, which follows directly from equation (5.27).

Concerning the mapping $\hat{P} : \mathcal{LC} \to C^\infty(M, \mathbb{R})$, we have:

Observe how the addition of λ in the definition of $\hat{P}$ doesn't change the value of the Poisson bracket but gives an additional term that is added to the result of the commutator $P([A,B])$.

$$\begin{aligned}&\big\{\hat{P}(A,\lambda),\hat{P}(B,\mu)\big\} - \hat{P}\Big(\big[(A,\lambda),(B,\mu)\big]\Big)\\ &\quad= \big\{P(A)+\lambda, P(B)+\mu\big\} - \Big(P\big([A,B]\big) + z(A,B)\Big)\\ &\quad= \big\{P(A),P(B)\big\} - P\big([A,B]\big) - z(A,B) = z(A,B) - z(A,B) = 0\,,\end{aligned}$$

for all $(A,\lambda),(B,\mu) \in \mathcal{LC} = \mathcal{LG} \oplus \mathbb{R}$, and hence $\hat{P}$ is a Lie algebra morphism as claimed. Furthermore, given the equality:

$$j \circ \hat{P}(A,\lambda) = j\big(P(A)+\lambda\big) = j\circ P(A) + j(\lambda) = j\circ P(A) = \gamma(A) = \gamma\circ\mathrm{pr}_1(A,\lambda)\,,$$

for all $(A,\lambda) \in \mathcal{LG} \oplus \mathbb{R} = \mathcal{LC}$, we see that the following diagram commutes:

$$\begin{array}{ccccccccc}
 & & & & \operatorname{im}\hat{P} & \xrightarrow{j|_{\operatorname{im}\hat{P}}} & \mathfrak{X}_{\mathrm{Obs}}(M) & \longrightarrow & 0\\
 & & & & \uparrow{\scriptstyle \hat{P}} & \nwarrow^{P} & \uparrow{\scriptstyle \gamma} & & \\
0 & \longrightarrow & \mathbb{R} & \xhookrightarrow{\iota_2} & \mathcal{LG}\oplus\mathbb{R} & \xrightarrow{\mathrm{pr}_1} & \mathcal{LG} & \longrightarrow & 0\,.
\end{array} \tag{α}$$

By definition, we have $\operatorname{im} P \subseteq \operatorname{im} \hat{P}$. Since the 'additional' elements, in $\operatorname{im} \hat{P} \setminus \operatorname{im} P$, are constant functions, the mapping $j|_{\operatorname{im} \hat{P}}$ remains a surjective mapping onto $\mathfrak{X}_{\mathrm{Obs}}(M)$.

To prove that the mapping $\hat{P} : \mathscr{LC} \to C^\infty(M, \mathbb{R})$ is injective consider $A, B \in \mathscr{LG}$ and $\lambda, \mu \in \mathbb{R}$ with $\hat{P}(A, \lambda) = \hat{P}(B, \mu)$. We will show $A = B$ and $\lambda = \mu$ using a simple diagram chasing argument. First, due to linearity of P we obtain:

$$\hat{P}(A, \lambda) = \hat{P}(B, \mu) \;\Leftrightarrow\; P(A) + \lambda = P(B) + \mu \;\Leftrightarrow\; P(A - B) + (\lambda - \mu) = 0\,, \qquad (\beta)$$

and application of j on both sides yields:

$$0 = j(0) = j\big(P(A - B) + (\lambda - \mu)\big) = j \circ P(A - B) = \gamma(A - B)\,.$$

Since γ is an isomorphism we get $A = B$. Inserting $P(A - B) = P(0) = 0$ back into equation (β) shows $\lambda = \mu$, and hence $(A, \lambda) = (B, \mu)$. Accordingly, the Lie algebra morphism $\hat{P} : \mathscr{LC} \to C^\infty(M, \mathbb{R})$ is injective. If we furthermore restrict the codomain of $\hat{P}$ to $\operatorname{im} \hat{P}$, the mapping $\hat{P} : \mathscr{LC} \to \operatorname{im} \hat{P}$ in diagram (α) becomes an isomorphism.

Finally, we can turn the upper row of diagram (α) into a short exact sequence by extending it to the left via the injection $\iota := \hat{P} \circ \iota_2 : \mathbb{R} \to \operatorname{im} \hat{P}$:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{R} & \xhookrightarrow{\iota} & \operatorname{im} \hat{P} & \xrightarrow{j|_{\operatorname{im} \hat{P}}} & \mathfrak{X}_{\mathrm{Obs}}(M) & \longrightarrow & 0 \\
 & & & & \wr\uparrow \hat{P} & \nearrow_{P} & \wr\uparrow \gamma & & \\
0 & \longrightarrow & \mathbb{R} & \xhookrightarrow{\iota_2} & \mathscr{LG} \oplus \mathbb{R} & \xrightarrow{\mathrm{pr}_1} & \mathscr{LG} & \longrightarrow & 0\,.
\end{array} \qquad (5.28)$$

Hence, if we write $\mathrm{Obs}(M) := \operatorname{im} \hat{P}$, this diagram turns into a diagram just like (5.25), which is exactly what we were looking for. ∎

For reference, figure 5.2 provides an overview of the whole construction of $\mathscr{LC}$.

5.4 Example: Conventional Quantum Mechanics

A FIRST APPLICATION of the quantization procedure is to reproduce conventional quantum mechanics on $Q = \mathbb{R}^n$. While Isham has already shown that the method yields the expected results (see [Ish83, section 4.4]), we nevertheless want to use this example to demonstrate how the quantization method is applied in practice. The familiar situation helps to understand what is going on.

On the other hand, despite the familiarity, conventional quantum mechanics turns out to be rather involved from the view of the quantization method. In fact, lemma 5.3.1 showed that the canonically conjugate observables q and p require a non-splitting central extension to obtain the algebra of fundamental observables $\mathrm{Obs}(M)$.

To avoid some unnecessary indices we will restrict ourselves to the 1-dimensional case. The classical configuration space is $Q = \mathbb{R}$ and the associated phase space is the trivial bundle $M = T^*\mathbb{R}$ which is globally isomorphic to $\mathbb{R}^2$ via the canonical Darboux chart from remark 4.4.2:

$$(q, p) : T^*\mathbb{R} \to \mathbb{R}^2\,.$$

The symplectic 2-form ω of $M = T^*\mathbb{R}$ is given in terms of these coordinates by:

$$\omega = \mathrm{d}q \wedge \mathrm{d}p\,.$$

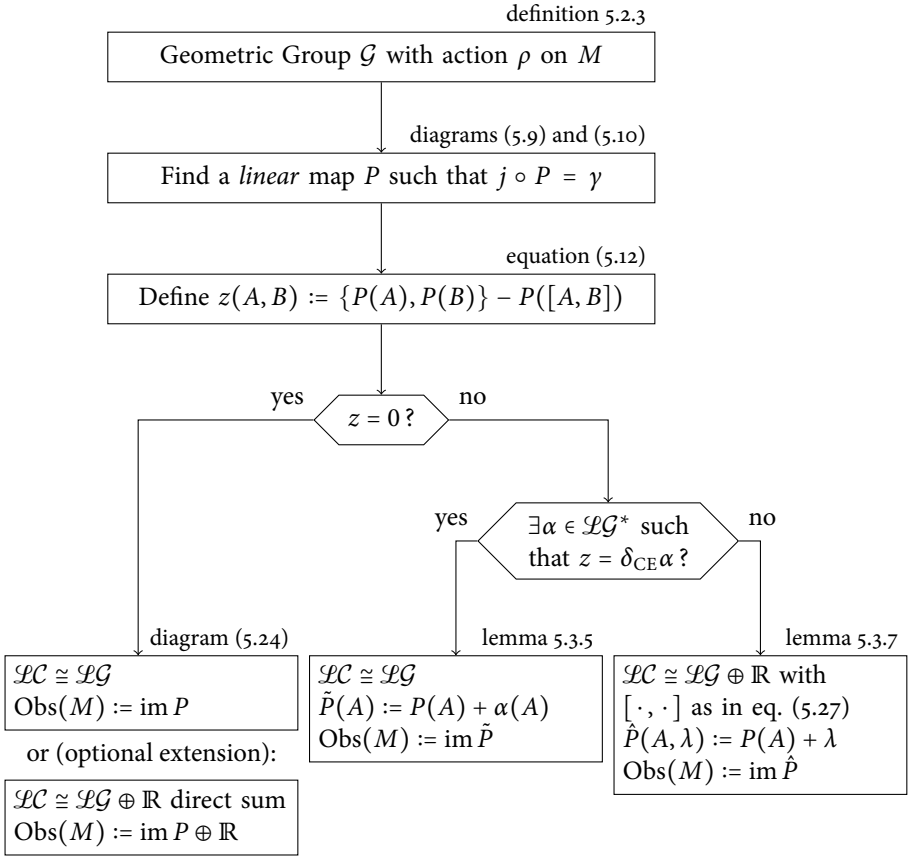


Figure 5.2 Recipe for the construction of $\mathcal{LC}$, given a Geometric Group $\mathcal{G}$.

5.4.1 Step 1: Geometric Group and Group Action

The first step is to find an appropriate Geometric Group $\mathcal{G}$ together with its group action ρ on M. One of the requirements from section 5.2 was that $\mathcal{G}$ must act transitively on M, and we also mentioned that transitivity requires $\dim \mathcal{G} \geq \dim M$. In our case this is $\dim \mathcal{G} \geq 2$. From a mathematical point of view, the simplest choice that could possibly work is hence the 2-dimensional, additive, Abelian group:

$$\mathcal{G} = (\mathbb{R}^2, +)\,, \tag{5.29}$$

which acts on $M \cong \mathbb{R}^2$ naturally by translations:

$$\rho : \mathcal{G} \times M \to M\,, \quad \rho_{(u,v)}(q,p) := (q+u, p-v)\,. \tag{5.30}$$

The signs in equation (5.30) don't affect the end result and can thus be chosen to simplify the calculations. They will later cancel out against some signs in the formula for Hamiltonian vector fields.

At the end of section 1.8 we also mentioned that a group acting on the configuration space by some kind of *translations* is necessary to define a reasonable notion of *distances*. As distances are a prerequisite to interpret points in the configuration space as *positions*, the above Geometric Group seems a good choice from a physical perspective as well.

Concerning the properties of a Geometric Group, observe that the group action ρ is *transitive* because for each two points $(q_1, p_1), (q_2, p_2) \in M$ there is always a group

element $(u, v) \in \mathcal{G}$ such that $\rho_{(u,v)}(q_1, p_1) = (q_2, p_2)$. Moreover, the action is *free* – and thus *effective* – given that $\rho_{(u,v)}(q, p) = (q, p)$ implies $(u, v) = (0, 0) = e \in \mathcal{G}$ already for any single choice of $(q, p) \in M$.

Next we need to check that the group action is *symplectic*, that is:

$$(\rho_g^* \omega)_s(X_s, Y_s) = \omega_{\rho_g s}\big(T_s\rho_g(X_s), T_s\rho_g(Y_s)\big) \stackrel{?}{=} \omega_s(X_s, Y_s) \,,$$

for all tangent vectors $X_s, Y_s \in T_sM$, and for every $s \in M$ and $g := (u, v) \in \mathcal{G}$. First, we calculate the value of the tangent map $T_s\rho_g(X_s)$ for $X_s \in T_sM$. To this end, let α be a smooth curve in M that represents the tangent vector X_s, i. e. $\alpha(0) = s$ and $\dot\alpha(0) = X_s$. For brevity, we will write $\tilde s = \rho_g(s)$, and α^i and X_s^i with $i \in \{q, p\}$ for the coordinate expressions of α and X_s, respectively. Using this notation, we get:

$$\begin{aligned} T_s\rho_g(X_s) = T_s\rho_g([\alpha]) = [\rho_{(u,v)} \circ \alpha] &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \big(\alpha^q(t) + u, \alpha^p(t) - v\big) \\ &= \big(\dot\alpha^q(0), \dot\alpha^p(0)\big) = (X_s^q, X_s^p) = X_s^q \partial_q\big|_{\tilde s} + X_s^p\big|_{\tilde s} = X_s^i \partial_i\big|_{\tilde s} \in T_{\tilde s}M \,. \end{aligned}$$

The transported tangent vector $T_s\rho_g(X_s)$ has the same 'entries', only the base point is shifted from s to $\tilde s$.

Given the coordinate expression $\omega_s = \mathrm{d}q \wedge \mathrm{d}p|_s$ of the symplectic 2-form we thereby obtain:

$$\begin{aligned} (\rho_g^* \omega)_s(X_s, Y_s) = \omega_{\tilde s}\big(T_s\rho_g(X_s), T_s\rho_g(Y_s)\big) &= \mathrm{d}q \wedge \mathrm{d}p\big|_{\tilde s}\big(X_s^i \partial_i\big|_{\tilde s}, Y_s^i \partial_i\big|_{\tilde s}\big) \\ = X_s^q X_s^p - X_s^p X_s^q = \mathrm{d}q \wedge \mathrm{d}p\big|_s\big(X_s^i \partial_i\big|_s, Y_s^i \partial_i\big|_s\big) &= \omega_s(X_s, Y_s) \,, \end{aligned}$$

and hence the group action ρ is symplectic. Conveniently, the first de Rham cohomology group $H^1_{\mathrm{dR}}(M)$ of our phase space $M = \mathbb{R}^2$ is known to be trivial, whereby remark 5.2.2 implies that ρ is in fact a *Hamiltonian* action. The Lie group $\mathcal{G}$ together with its group action ρ thus satisfies all properties required in the definition 5.2.3 of a Geometric Group.

5.4.2 Step 2: Lie Algebra and Antifundamental Vector Fields

Now that we have a Geometric Group $\mathcal{G}$, the next step is to calculate the Lie algebra $\mathfrak{L}\mathcal{G}$ and the antifundamental vector field mapping γ. The tangent space of $\mathcal{G} = \mathbb{R}^2$ at the

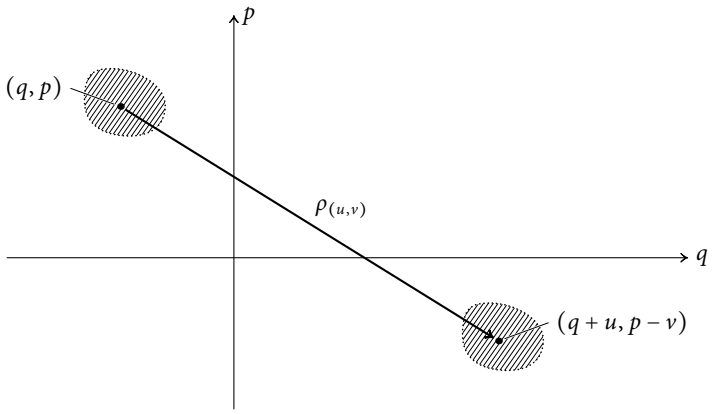


Figure 5.3 Geometric Group $\mathcal{G} = (\mathbb{R}^2, +)$ acting on $M = T^*\mathbb{R}$ by translations. The group action ρ preserves phase space volumes (drawn hatched).

point $e = (0,0) \in \mathcal{G}$ is given by $T_e\mathcal{G} \cong \mathbb{R}^2$, hence:

$$\mathcal{L}\mathcal{G} = (\mathbb{R}^2, [\,\cdot\,,\,\cdot\,])\,, \quad \text{with } [A,B] = 0 \quad \forall A, B \in \mathcal{L}\mathcal{G}\,, \tag{5.31}$$

where the vanishing Lie bracket results from the fact that $\mathcal{G}$ is Abelian (see the explicit expression for the Lie bracket in equation (3.13)).

To obtain the exponential map $\exp : \mathcal{L}\mathcal{G} \to \mathcal{G}$ we will use the left invariant vector fields $\mathfrak{X}_{\mathrm{L}}(\mathcal{G})$. If $\ell : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ denotes the left action of $\mathcal{G}$ on itself ($\ell_g h = g + h$) the left invariant vector field L^A generated by $A = [\alpha] \in T_e\mathcal{G}$ is (equation (3.6)):

In this uncomplicated situation we could just as well *guess* the exponential map and then use the uniqueness theorem 3.4.7, but we want to demonstrate the general method.

$$L^A_g = T_e\ell_g(A) = T_e\ell_g([\alpha]) = [\ell_g \circ \alpha] = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \big(g + \alpha(t)\big) = \dot\alpha(0) = A\,.$$

The integral curve γ_A through $e \in \mathcal{G}$ of this constant vector field is determined by:

$$\dot\gamma_A(t) = L^A_{\gamma_A(t)}\,, \quad \gamma_A(0) = e \quad \Rightarrow \quad \gamma_A(t) = tA\,.$$

The exponential map for our case is hence given by (definition 3.4.5):

This looks essentially like the identity map but implicitly uses the isomorphism $\mathcal{L}\mathcal{G} \cong \mathbb{R}^2 = \mathcal{G}$.

$$\exp : \mathcal{L}\mathcal{G} \to \mathcal{G}\,, \quad \exp(tA) = \gamma_{tA}(1) = \gamma_A(t) = tA\,. \tag{5.32}$$

Finally, the formula for antifundamental vector fields, equation (5.3), applied to a Lie algebra element $A = (a,b) \in \mathcal{L}\mathcal{G} \cong \mathbb{R}^2$ yields the pointwise result:

$$\gamma(A)_s = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{\exp(-tA)} s = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \big(q - ta, p + tb\big) = -a\partial_q\big|_s + b\partial_p\big|_s\,.$$

Globally, the vector field $\gamma(A)$ is:

$$\gamma(A) = -a\partial_q + b\partial_p\,. \tag{5.33}$$

5.4.3 Step 3: Construction of the Canonical Group

We are now ready to construct the Canonical Group $\mathcal{C}$ according to the recipe shown in figure 5.2. The first step is to find a *linear* mapping P that satisfies $j \circ P = \gamma$, such that the following diagram commutes:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \hookrightarrow & C^\infty(\mathbb{R}^2,\mathbb{R}) & \xrightarrow{\;j\;} & \mathfrak{X}_{\mathrm{Ham}}(\mathbb{R}^2) & \longrightarrow & 0\,. \\ & & & & & \overset{P}{\searrow} & \uparrow{\scriptstyle \gamma} & & \\ & & & & & & \mathcal{L}\mathcal{G} & & \end{array}$$

To construct such a linear mapping P, let f be a function $f \in C^\infty(\mathbb{R}^2,\mathbb{R})$ and calculate $j(f) = -X_f = -\mathrm{d}f^\sharp$. The necessary musical isomorphism $\sharp = \flat^{-1}$ can be obtained from $X^\flat = \mathrm{i}_X\omega = \mathrm{d}q \wedge \mathrm{d}p(X,\,\cdot\,)$ and is given in Darboux coordinates by:

$$\sharp : T^*_sM \to T_sM\,, \quad \sharp(a\,\mathrm{d}q + b\,\mathrm{d}p) = -a\partial_p + b\partial_q\,.$$

Therefore:

$$j(f) = -X_f = -\mathrm{d}f^\sharp = -\big((\partial_q f)\mathrm{d}q^\sharp + (\partial_p f)\mathrm{d}p^\sharp\big) = -(\partial_p f)\partial_q + (\partial_q f)\partial_p\,. \tag{5.34}$$

Continuing, we require $j(f) = \gamma(A)$ and equate the corresponding coefficients in equations (5.33) and (5.34) to obtain:

$$-a\partial_q + b\partial_p \overset{!}{=} -(\partial_p f)\partial_q + (\partial_q f)\partial_p \quad \Leftrightarrow \quad a = \partial_p f\,, \quad b = \partial_q f\,.$$

We will drop the constant for now and remember that this is an inherent freedom of the defining relation $j \circ P = \gamma$.

Accordingly, the function f related to the antifundamental vector field $\gamma(A)$ is given by $f(s) = ap(s) + bq(s) + \text{const}$. A linear map P that satisfies $j \circ P = \gamma$ is therefore:

$$P : \mathscr{LG} \to C^\infty(\mathbb{R}^2, \mathbb{R})\,, \quad A = (a,b) \mapsto P(a,b) \coloneqq ap + bq\,. \tag{5.35}$$

Given the mapping P, we can calculate the obstruction cocycle z (equation (5.12)). The Poisson bracket of $P(A_1)$ and $P(A_2)$ for $A_1 = (a_1, b_1)$ and $A_2 = (a_2, b_2)$ is:

$$\begin{aligned}\big\{P(A_1), P(A_2)\big\} &= \omega\big(X_{P(A_1)}, X_{P(A_2)}\big) = \omega\big(\gamma(A_1), \gamma(A_2)\big)\\ &= \mathrm{d}q \wedge \mathrm{d}p\big(\gamma(a_1,b_1), \gamma(a_2,b_2)\big) = b_1a_2 - b_2a_1\,.\end{aligned}$$

Since the Lie bracket of $\mathscr{LG}$ vanishes, we have $P([A_1, A_2]) = P(0) = 0$, and hence:

$$z\big((a_1,b_1),(a_2,b_2)\big) = b_1a_2 - b_2a_1\,. \tag{5.36}$$

The obstruction cocycle obviously doesn't vanish and therefore P is not a Lie algebra morphism. Worse even, since the Lie bracket of $\mathscr{LG}$ is trivial we have $\alpha([A_1, A_2]) = 0$ for any choice of $\alpha \in \mathscr{LG}^*$, which means z *cannot* be written as $z = \delta_{\mathrm{CE}}\alpha$ (the cocycle z is *not* a coboundary) and so there is no way to adjust P that would yield a Lie algebra morphism. Consequently, we have to construct $\mathscr{LC}$ as a central extension.

Fortunately, lemma 5.3.7 already provides all the details. An appropriate central Lie algebra extension of $\mathscr{LG}$ is hence given by:

$$\mathscr{LC} = \mathscr{LG} \oplus \mathbb{R} = \mathbb{R}^2 \times \mathbb{R}\,, \tag{5.37}$$

with Lie bracket $[(A_1, r_1), (A_2, r_2)] = \big([A_1, A_2]\,, z(A_1, A_2)\big)$, that is:

$$\big[(a_1,b_1;r_1),(a_2,b_2;r_2)\big] \coloneqq (0,0;b_1a_2 - b_2a_1)\,, \tag{5.38}$$

for all $(a_1, b_1; r_1), (a_2, b_2; r_2) \in \mathscr{LC} \cong \mathbb{R}^2 \times \mathbb{R}$. Moreover, there is an injective Lie algebra morphism:

$$\hat{P} : \mathscr{LC} \to C^\infty(\mathbb{R}^2, \mathbb{R})\,, \quad \hat{P}(a,b;r) \coloneqq ap + bq + r\,.$$

The algebra of fundamental classical observables $\mathrm{Obs}(M) \coloneqq \mathrm{im}\,\hat{P}$ is hence generated by the set $\{q, p, 1\}$, and the fundamental Poisson brackets of the generators are:

$$\{q,p\} = 1\,, \quad \{q,q\} = \{p,p\} = \{1,1\} = \{q,1\} = \{p,1\} = 0\,. \tag{5.39}$$

This result comes as no surprise, although the Poisson brackets involving the constant function $1 = 1_M$ are usually omitted. Nevertheless, we included them to make clear that constant functions are *not to be ignored as something gratuitous*. Constant functions are really an *essential* part of the algebra of fundamental observables $\mathrm{Obs}(M)$ in the case of conventional quantum mechanics, not optional. In fact, lemma 5.3.1 showed that this a direct consequence of q and p being canonically conjugate observables.

Furthermore, the fact that q and p can really be interpreted as the actual physical quantities of position and momentum is not as obvious as it might seem. We have, of course, chosen some suggestive notation, and we already know what to expect. Still, a

proper argument involves that $\mathcal{G}$ acts on phase space in a way so that translations in the position coordinate are generated by the momentum observable whereas translations in momentum coordinate are generated by the position observable.

In fact, the all too familiar situation makes it difficult to see an issue here. In general, however, it turns out that the mathematical appearance of *individual* observables is not enough to justify a physical interpretation. In section 6.1 we will discuss an example where the problem of interpretation proves more tangible and where the same function on phase space can be interpreted in quite different ways.

What remains for now is to construct the actual Canonical Group $\mathcal{C}$ that is associated to the Canonical Lie algebra $\mathfrak{LC}$. One possibility to obtain the group multiplication is to use some formal properties of the exponential map $\exp : \mathfrak{LC} \to \mathcal{C}$. Given a Lie algebra element $A = (a, b; r) \in \mathfrak{LC}$, let us define:

$$\exp(A) = \exp(a, b; r) =: (a, b; r)_{\mathcal{C}} = (A)_{\mathcal{C}} \,.$$

The result is formally a group element in $\mathcal{C}$. For small Lie algebra elements near 0 we can then use the Baker–Campbell–Hausdorff formula 3.4.10 to formally obtain the group operation $*$ in $\mathcal{C}$ as:

$$\begin{aligned}(A_1)_{\mathcal{C}} * (A_2)_{\mathcal{C}} &= \exp(A_1) * \exp(A_2) = \exp\big(A_1 + A_2 + \tfrac{1}{2}[A_1, A_2] + \ldots\big) \\ &= \big(A_1 + A_2 + \tfrac{1}{2}[A_1, A_2] + \ldots\big)_{\mathcal{C}} \,.\end{aligned}$$

Since the commutator (5.38) maps into the centre $\mathcal{Z}(\mathfrak{LC})$ of the Lie algebra $\mathfrak{LC} = \mathbb{R}^2 \times \mathbb{R}$, all higher commutators vanish and we obtain:

$$(a_1, b_1; r_1)_{\mathcal{C}} * (a_2, b_2; r_2)_{\mathcal{C}} = \big(a_1 + a_2, b_1 + b_2; r_1 + r_2 + \tfrac{1}{2}(b_1 a_2 - b_2 a_1)\big)_{\mathcal{C}} \,. \quad (5.40)$$

This is easily recognised as the group law of the 1-dimensional *Heisenberg group* $H(1)$ from definition 1.6.17.

Note that unlike the Geometric Group $\mathcal{G} = (\mathbb{R}^2, +)$ from which it was constructed, the Heisenberg group $H(1)$ is no longer Abelian. Nevertheless, now that we know how the 'strange' group law of the Heisenberg group $H(1)$ arises from a central Lie algebra extension of the simple translation group $(\mathbb{R}^2, +)$, the group law (5.40) looks much more natural than it did before.

5.4.4 Step 4: Representations and Canonical Commutation Relations

The final step of the quantization procedure is to study unitary, irreducible representations of the Canonical Group $\mathcal{C}$ and to define the quantum operators as the self-adjoint generators of strongly continuous 1-parameter unitary groups.

Suppose $U' : \mathcal{C} = H(1) \to U(\mathcal{H})\,,\ g \to U'(g)$ is such a strongly continuous, unitary, irreducible representation of the Heisenberg group, where $U'(g)$ is a unitary operator on the (yet unknown) Hilbert space $\mathcal{H}$. Furthermore, the uniqueness theorem for the exponential map, theorem 3.4.7, states that 1-parameter subgroups of $\mathcal{C}$ can be written as restrictions of exp to lines through the origin of $\mathfrak{LC}$. Accordingly, let us define:

$$U(a) := U'\big(\exp(a, 0; 0)\big)\,, \quad V(b) := U'\big(\exp(0, b; 0)\big)\,, \quad (5.41)$$

so that U and V are unitary representations of 1-parameter subgroups of $\mathcal{C}$. We will represent elements in the centre $\mathcal{Z}(\mathcal{C}) = \{\exp(0, 0; r) \in \mathcal{C} : r \in \mathbb{R}\}$ by:

$$U'\big(\exp(0, 0; r)\big) = e^{-i\mu r}\mathbb{1}\,, \quad \mu \in \mathbb{R}\,. \quad (5.42)$$

This way, we get the group commutation relations:

$$\begin{aligned} U(a_1)U(a_2) &= U'\big(\exp(a_1,0;0) * \exp(a_2,0;0)\big) = U'\big(\exp(a_1+a_2,0;0)\big) \\ &= U(a_1+a_2)\,, \\ V(b_1)V(b_2) &= V(b_1+b_2)\,, \end{aligned}$$

as well as:

$$\begin{aligned} U(a)V(b)U(a)^{-1}V(b)^{-1} &= U'\Big(\exp\big(\big[(a,0;0),(0,b;0)\big]+\ldots\big)\Big) \\ &= U'\big(\exp(0,0;-ba)+0\big) = e^{i\mu ab} \;\Leftrightarrow\; U(a)V(b) = V(b)U(A)e^{i\mu ab}\,. \end{aligned}$$

These relations are nothing but the Weyl form of the canonical commutation relations from lemma 1.6.16, only that U and V are now defined in terms of 1-parameter subgroups of $\mathcal{C}$, not in terms of the generating operators. Nevertheless, Stone's theorem, theorem 1.6.15, states that there *are* self-adjoint generators $\hat{q}$ and $\hat{p}$ such that:

$$e^{-ia\hat{p}} = U(a) = U'\big(\exp(a,0;0)\big)\,, \quad e^{-ib\hat{q}} = V(b) = U'\big(\exp(0,b;0)\big)\,, \tag{5.43}$$

which means we can now *define* $\hat{q}$ and $\hat{p}$ this way. If you look back at how the Canonical Group $\mathcal{C}$ was constructed, you will see that the operator $\hat{p}$ is related to translations U in position and $\hat{q}$ related to translations V in momentum, just as it should be. The operators $\hat{q}$, $\hat{p}$ and $\mathbb{1}$ also satisfy the usual commutation relations (although the parameter μ takes the place of $\hbar$).

The representations of the Heisenberg group $\mathcal{C} = H(1)$ are given by theorem 1.6.18, of Stone and von Neumann, which states that after fixing a value of $\mu \in \mathbb{R}$, all non-trivial, strongly continuous, unitary, irreducible representations are unitarily equivalent to the well-known Schrödinger representation in position space $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$, where the fundamental quantum operators are given by:

$$\hat{q} = q\,, \quad \hat{p} = -i\mu\partial_q\,, \quad \hat{1}_M = \mathbb{1}\,. \tag{5.44}$$

The quantization map $\mathcal{Q}$ is hence determined by:

$$\mathcal{Q} : ap + bq + r \mapsto a\hat{p} + b\hat{q} + r\mu\mathbb{1}\,. \tag{5.45}$$

5.4.5 Physical Units

As one might have noticed, the operators in equation (5.43) so far are dimensionless and the parameter μ is just a real number. Nevertheless, to introduce physical units we can simply choose a base unit q_0 for the length, p_0 for the momentum, and define *physical* operators as:

$$\hat{q}_\phi := q_0\hat{q}\,, \quad \hat{p}_\phi := p_0\hat{p}\,. \tag{5.46}$$

The commutator of these operators if we use the representation (5.44) is:

$$[\hat{q}_\phi, \hat{p}_\phi] = i\mu q_0 p_0 \mathbb{1}\,.$$

Finally, given that the combination $\mu q_0 p_0$ on the right-hand side has the physical unit of an action (length × momentum) it seems reasonable to identify $\mu q_0 p_0$ with the fundamental physical constant $\hbar$ that can be measured experimentally:

$$\mu q_0 p_0 =: \hbar\,. \tag{5.47}$$

Remember, once again, that the reason why $\hbar$ is measurable in experiments is that different values of μ yield unitarily *inequivalent* representations. If, on the other hand, representations belonging to different values of $\hbar$ were unitarily equivalent they would produce physically identical results and the explicit value of $\hbar$ would be irrelevant.

The physical meaning behind $\hbar$ is that it introduces a *scale dependence* into the theory, as discussed at the end of section 1.6.

Remark 5.4.1 Observe that equation (5.47) doesn't fix the physical base units q_0 and p_0 completely; it only enforces that their product $q_0 p_0$ be equal to the constant $\hbar/\mu$. Also, notice that this freedom appears already in the classical Poisson bracket $\{q, p\} = 1$. We will see this behaviour again in section 6.4.

In summary, we have shown that the method of Canonical Group Quantization is able to reproduce the theory of conventional quantum mechanics in 1 dimension, and the higher-dimensional case works analogously if we just decorate q and p with some indices. This is a reassuring result and helps to gain some trust in the method. Nevertheless, the real value of the Canonical Group Quantization method is that it can be applied to phase spaces where the conventional approach fails.

5.5 Quantization on Cotangent Bundles

Our presentation in this section again roughly follows Isham's notes [Ish83].

In the last section the classical phase space M was a cotangent bundle. This case is so common in physics that it pays off to investigate the special case of cotangent bundles more closely. Also from a mathematical point of view a cotangent bundle $M = T^*Q$ is particularly interesting because every cotangent bundle carries a *canonical* symplectic structure ω (see section 4.4):

$$\omega = -\mathrm{d}\theta \,,$$

where θ is the *Liouville* 1*-form*, given by equation (4.25):

$$\theta_m(X_m) = m(T_m\pi(X_m)) = m(\pi_* X_m) \quad \forall X_M \in T_m M \,.$$

In addition, the unique properties of cotangent bundles yield some natural constructions for symplectomorphisms that don't exist in a more general setting (see section 4.5).

5.5.1 The group $C^\infty(Q, \mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$

An example of a phase space that is *not* a cotangent bundle is the model of angle and action variables, $M = S^1 \times \mathbb{R}^+$.

From a physical standpoint, the characteristic feature of a phase space M that is a cotangent bundle, is the existence of a configuration space Q – the base manifold of the bundle $M = T^*Q$. Furthermore, there might be a natural group of transformations on the configuration space Q. In general, this must be a subgroup of the infinite-dimensional Lie group $\mathrm{Diff}(Q)$ which contains *all* global diffeomorphisms of Q. In case of conventional quantum mechanics, for example, this subgroup of diffeomorphisms of the configuration space consists of translations $\rho_u : q \mapsto q + u$ with $u \in (\mathbb{R}, +)$.

What we need for the quantization procedure, however, are not transformations on *configuration space* but transformations on *phase space*. Fortunately, the *cotangent lift* provides a solution to this problem. In section 4.5, the cotangent lift of a group action $\rho : G \to \mathrm{Diff}(Q)\,,\ g \mapsto \rho_g$ was defined as (equation (4.37)):

$$\rho^\uparrow : G \to \mathrm{Sp}(M)\,, \quad g \mapsto \rho_g^\uparrow := T_*\rho_g = T^*\rho_{g^{-1}}\,. \tag{5.48}$$

Note that we can use the group $(\mathrm{Diff}(Q), \circ)$ for G, in which case the action ρ is the identity $\rho = \mathrm{id}_{\mathrm{Diff}(Q)}$, with $\rho_\varphi := \varphi$.

In particular, lemma 4.5.4 states that the cotangent lift $\rho^\uparrow$ of a left (right) action ρ on Q is a symplectic left (right) action on phase space M. The above action $\rho_u : q \mapsto q + u$, for

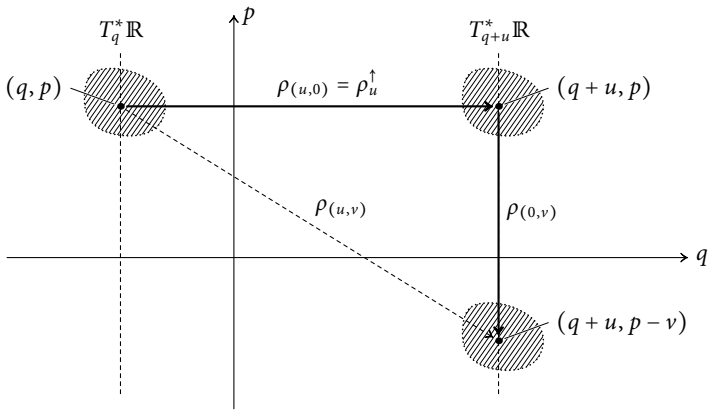


Figure 5.4 The group action on $\mathbb{R}^2$ decomposes into the *cotangent lift* $\rho_{(u,0)}$ of a translation ρ_u in configuration space, augmented by a *fibre translation* $\rho_{(0,v)}$. The q-axis in this drawing is actually the embedding of the base manifold $Q = \mathbb{R}$ via the zero section $Q \to T^*Q\,,\ q \mapsto (q, 0_q)$.

example, acting on the configuration space $Q = \mathbb{R}$, is thereby lifted to the group action $\rho^{\uparrow}_u : (q,p) \mapsto (q+u,p)$ which acts symplectically on phase space $M = T^*\mathbb{R} \cong \mathbb{R}^2$.

That being said, we also know that the cotangent lift of an action is by itself never transitive but can be augmented by *fibre translations* to obtain a transitive group action. The fibre translation by a 1-form $\beta \in \Omega^1(Q)$ is given by equation (4.38):

The decomposition of the group action $\rho_{(u,v)}$ from before is shown in figure 5.4.

$$\tau_\beta : T^*Q \to T^*Q\,, \quad m \mapsto m + \beta_{\pi_Q(m)}\,,$$

where m is a point $m \in M = T^*Q$ and where $\pi_Q : T^*Q \to Q$ is the bundle projection. Remark 4.5.9 then showed that the mapping τ:

$$\tau : G \to \mathrm{Diff}(M)\,, \quad \beta \mapsto \tau_\beta\,, \tag{5.49}$$

is a group action if we see the vector space of 1-forms as a Lie group $G = (\Omega^1(Q), +)$. In addition, lemma 4.5.8 tells us that this action is *symplectic* if and only if β is closed, and *Hamiltonian* if and only if β is *exact*.

Because G is a vector space, the Lie algebra $\mathcal{L}G$ is isomorphic to G itself, and the exponential is given by $\exp(t\beta) = t\beta$.

Since a Hamiltonian group action is what we need for the quantization method, we are interested only in the case where β is an exact 1-form, that is, we look at the case where there exists a function $f \in C^\infty(Q,\mathbb{R}) = \Omega^0(Q)$ such that $\beta = \mathrm{d}f$. Accordingly, we could just let the additive group of functions $C^\infty(Q,\mathbb{R})$ act on M via $f \mapsto \tau_{\mathrm{d}f}$, which gives a Hamiltonian action. However, since two functions f and g that differ only by a constant produce the same 1-form $\beta = \mathrm{d}f = \mathrm{d}g$, this group action is not effective. Instead, we will thus use the quotient group $C^\infty(Q,\mathbb{R})/\mathbb{R}$, where constant functions have been divided out. The resulting group action of $(C^\infty(Q,\mathbb{R})/\mathbb{R}, +)$ on M is:

$$\tau : C^\infty(Q,\mathbb{R})/\mathbb{R} \to \mathrm{Sp}(M)\,, \quad \alpha \mapsto \tau_{\mathrm{d}\alpha}\,, \tag{5.50}$$

The exterior derivative of an equivalence class $\alpha = [f]$ is defined as $\mathrm{d}\alpha := [\mathrm{d}f]$.

with:

$$\tau_{\mathrm{d}\alpha} : M \to M\,, \quad \tau_{\mathrm{d}\alpha}(m) = m + (\mathrm{d}\alpha)_{\pi_Q(m)}\,. \tag{5.51}$$

Due to the construction, this group action τ is both *effective* and *Hamiltonian*.

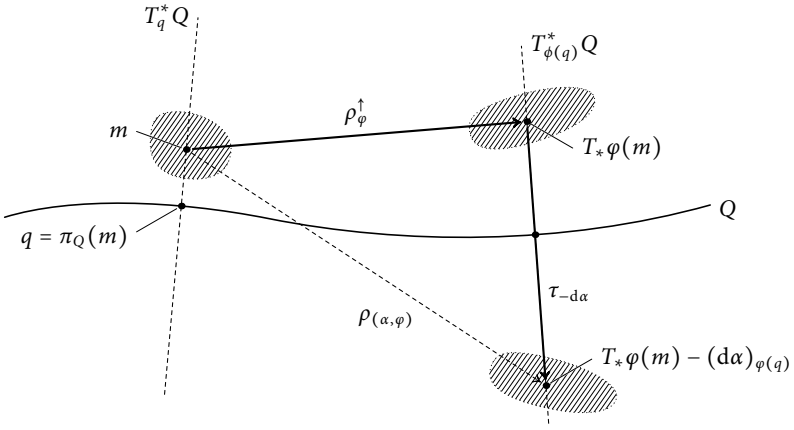


Figure 5.5 Group action of the semidirect product $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ on the cotangent bundle $M = T^*Q$. Observe the similarity to figure 5.4.

Finally, the group actions $\rho^{\uparrow}$, obtained via the cotangent lift, and τ, induced by fibre translations, can be combined into a transitive, symplectic action ρ on M:

$$\rho : C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q) \to \mathrm{Sp}(M)\,, \quad (\alpha,\varphi) \mapsto \rho_{(\alpha,\varphi)}\,, \tag{5.52}$$

with:

$$\rho_{(\alpha,\varphi)} : M \to M\,, \quad \rho_{(\alpha,\varphi)}(m) := T_*\varphi(m) - (\mathrm{d}\alpha)_{\varphi\circ\pi_Q(m)}\,. \tag{5.53}$$

The group law of $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ turns out as:

$$(\alpha_2,\varphi_2) * (\alpha_1,\varphi_1) = \left(\alpha_2 + \alpha_1 \circ \varphi_2^{-1}, \varphi_2 \circ \varphi_1\right), \tag{5.54}$$

which is the usual form of group law for an (external) semidirect product (see section 3.6, specifically definition 3.6.1). The neutral and inverse elements of this group are:

$$e = (0, \mathrm{id})\,, \quad (\alpha,\varphi)^{-1} = (-\alpha\circ\varphi, \varphi^{-1})\,. \tag{5.55}$$

In general, a suitable Geometric Group $\mathcal{G}$ for a cotangent bundle T^*Q will be a finite-dimensional subgroup of the infinite-dimensional Lie group $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$.

It is actually possible to use the group $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ directly, instead of a subgroup (see [Ish83, section 4.7.3]). The infinite-dimensional group however yields infinitely many fundamental observables, which is usually not what we want.

5.5.2 Lie algebra of $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$

Since the full group $C^\infty(Q,\mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$ is rather complicated we will build up its Lie algebra in several consecutive steps.

Let us begin with the Lie algebra of the diffeomorphism group $\mathrm{Diff}(Q)$. The group law of $\mathrm{Diff}(Q)$ is the composition of diffeomorphisms, the neutral element the identity $e = \mathrm{id}_Q$, and the inverse of a diffeomorphism φ is given by φ^{-1}. To obtain the Lie algebra, note that each 1-parameter subgroup of global diffeomorphisms in $\mathrm{Diff}(Q)$ can be read as a *complete flow* $\Phi : \mathbb{R}\times Q \to Q$. The mapping $\alpha : \mathbb{R} \to \mathrm{Diff}(Q)\,,\ t \mapsto \alpha(t) = \Phi_t$ is a smooth curve in $\mathrm{Diff}(Q)$ which satisfies $\alpha(0) = \mathrm{id}_Q = e \in \mathrm{Diff}(Q)$. The Lie algebra then consists of tangent vectors to curves like α.

To understand how the tangent vector for a curve of diffeomorphisms looks like, observe that for each $q \in Q$ the mapping $\Phi_q : \mathbb{R} \to Q\,,\ t \mapsto \Phi_q(t) = \Phi_t(q)$ is a complete

curve in Q which passes through the point $\Phi_q(0) = q$. The tangent vector for the curve Φ_q is a tangent vector $v_q \in T_qQ$ in the usual sense. This is the *pointwise* relation that must hold between α and its tangent vector V, i. e. $V_q = v_q$ for each $q \in Q$. Accordingly, the tangent vector V for the curve α of diffeomorphisms can be understood as a *vector field* on Q – more precisely, V is the vector field that generates the flow Φ. Furthermore, since Φ is a *complete* flow, we see that the Lie algebra of $\mathrm{Diff}(Q)$ is isomorphic to the algebra of *complete vector fields* on Q, that is:

$$\mathscr{L}\,\mathrm{Diff}(Q) = T_{\mathrm{id}}\,\mathrm{Diff}(Q) \cong \mathfrak{X}_{\mathrm{cpl}}(Q)\,. \tag{5.56}$$

What remains is to calculate the Lie bracket of $\mathscr{L}\,\mathrm{Diff}(Q)$. One way to do this is via the adjoint action and theorem 3.3.10. Accordingly, we begin with the conjugation automorphism I_φ:

$$I_\varphi(\psi) = \varphi \circ \psi \circ \varphi^{-1}\,, \quad \varphi, \psi \in \mathrm{Diff}(Q)\,. \tag{5.57}$$

If we then write $\alpha : t \mapsto \alpha_t$ and $\beta : t \mapsto \beta_t$ for the diffeomorphism-valued curves with $\alpha_0 = \beta_0 = \mathrm{id}$ that represent the tangent vectors $A, B \in T_{\mathrm{id}}\,\mathrm{Diff}(Q)$, respectively, we get the adjoint action of the Lie algebra $\mathscr{L}\,\mathrm{Diff}(Q)$ on itself from equation (3.13) as:

The calculation employs the product rule and the fact that $\alpha_{-t} = \alpha_t^{-1}$ is the flow of $-A$.

$$\begin{aligned}
\mathrm{ad}(A)(B)(f)(q) &= T_{\mathrm{id}}\,\mathrm{Ad}([\alpha])([\beta])(f)(q) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{Ad}_{\alpha_t}([\beta])(f)(q) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f \circ \alpha_t \circ \beta_s \circ \alpha_t^{-1}(q) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} A(f)\big(\beta_s \circ \alpha_0^{-1}(q)\big) + \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} B(f \circ \alpha_0)\big(\alpha_t^{-1}(q)\big) \\
&= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} A(f)\big(\beta_s(q)\big) + \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} B(f)\big(\alpha_t^{-1}(q)\big) \\
&= B \circ A(f)(q) - A \circ B(f)(q)\,,
\end{aligned}$$

for all functions $f \in C^\infty(Q, \mathbb{R})$ and for each point $q \in Q$. A little surprising, the Lie bracket of $\mathscr{L}\,\mathrm{Diff}(Q)$ is hence the *negative* of the Jacobi–Lie bracket of vector fields:

$$[A, B]_{\mathscr{L}} = B \circ A - A \circ B = -[A, B]\,, \tag{5.58}$$

for $A, B \in \mathscr{L}\,\mathrm{Diff}(Q) = \mathfrak{X}_{\mathrm{cpl}}(Q)$.

Now that we have the Lie algebra of $\mathrm{Diff}(Q)$, let us proceed with the Lie algebra of the full group $C^\infty(Q, \mathbb{R})/\mathbb{R} \rtimes \mathrm{Diff}(Q)$. The conjugation automorphism $I_{(\alpha,\varphi)}$ in this case is given by:

$$I_{(\alpha,\varphi)}(\beta, \psi) = \big(\alpha + \beta \circ \varphi^{-1} - \alpha \circ \varphi \circ \psi^{-1} \circ \varphi^{-1}, \varphi \circ \psi \circ \psi^{-1}\big)\,, \tag{5.59}$$

for group elements (α, φ) and (β, ψ). There are the following special cases:

(i) For $\alpha = \beta = 0$ we get:

$$I_{(0,\varphi)}(0, \psi) = (0, \varphi \circ \psi \circ \varphi^{-1})\,.$$

As the second entry is the same as the conjugation automorphism of $\mathrm{Diff}(Q)$ from equation (5.57), the corresponding Lie bracket is:

$$\big[(0, A), (0, B)\big]_{\mathscr{L}} = \big(0, -[A, B]\big)\,. \tag{5.60}$$

(ii) The case $A = B = 0$ is even easier. Given that $C^\infty(Q,\mathbb{R})/\mathbb{R}$ is vector space, the Lie algebra can be identified with the group itself. The exponential is $\exp(\alpha) = \alpha$ and the Lie bracket is trivial because the group is Abelian:

$$\big[(\alpha,0),(\beta,0)\big]_{\mathfrak{L}} = (0,0)\,. \tag{5.61}$$

(iii) For a mixed combination of group elements, $(0,\varphi)$ and (β,id), the conjugation is:

$$I_{(0,\varphi)}(\beta,\mathrm{id}) = \big(0+\beta\circ\varphi^{-1}-0,\varphi\circ\mathrm{id}\circ\varphi^{-1}\big) = (\beta\circ\varphi^{-1},\mathrm{id})\,.$$

Accordingly, for a Lie algebra element $(\beta,0)$ we get:

$$\mathrm{Ad}_{(0,\varphi)}(\beta,0) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} I_{(0,\varphi)}\big(t\beta,\exp(0)\big) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big((t\beta)\circ\varphi^{-1},\mathrm{id}\big) = (\beta\circ\varphi^{-1},0)\,.$$

Furthermore, if we represent A by a curve $t\mapsto\alpha_t$ as before, we obtain:

$$\mathrm{ad}(0,A)(\beta,0) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\mathrm{Ad}_{(0,\alpha_t)}(\beta,0) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\beta\circ\alpha_t^{-1},0) = (-A(\beta),0)\,.$$

Given $A(\beta) = \mathrm{d}\beta(A)$, we see that the choice of representative for β doesn't really matter.

From this, we get the mixed Lie bracket:

$$\big[(0,A),(\beta,0)\big]_{\mathfrak{L}} = \big(-A(\beta),0\big)\,. \tag{5.62}$$

Finally, the general Lie bracket of $C^\infty(Q,\mathbb{R})/\mathbb{R}\rtimes\mathrm{Diff}(Q)$ is now easy to calculate if we exploit linearity and antisymmetry of the bracket. The result is:

$$\big[(\alpha,A),(\beta,B)\big]_{\mathfrak{L}} = \big(B(\alpha)-A(\beta),-[A,B]\big)\,. \tag{5.63}$$

5.5.3 Antifundamental Vector Fields

Right now, we have the group $C^\infty(Q,\mathbb{R})/\mathbb{R}\rtimes\mathrm{Diff}(Q)$ that acts on phase space via the group action (5.52). We know that the group action is effective, transitive and symplectic. The next step is to calculate the antifundamental vector fields and check whether they are Hamiltonian.

Again, let us begin with the 'Diff(Q) part' of the group. To this end, consider a Lie algebra element (a complete vector field) $V\in\mathfrak{X}_{\mathrm{cpl}}(Q) = \mathfrak{L}\mathrm{Diff}(Q)$ and write $\hat{V}$ for the associated antifundamental vector field $\gamma(V)\in\mathfrak{X}(M)$. Since the action of a diffeomorphism $\varphi\in\mathrm{Diff}(Q)$ on M is given by $\rho_{(0,\varphi)}(m) = T_*\varphi(m)$, according to equation (5.53), the vector field $\hat{V}$ is:

$$\hat{V}_m = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\rho_{(0,\exp(-tA))}(m) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}T_*\big(\exp(-tA)\big)(m) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}T^*\big(\exp(tA)\big)(m)\,,$$

pointwise, for each $m\in M = T^*Q$.

Concerning the symplectic 2-form ω of the cotangent bundle T^*Q, we know from section 4.4 that it fulfils:

$$\omega(Y,Z) = -\mathrm{d}\theta(Y,Z) = Z\big(\theta(Y)\big) - Y\big(\theta(Z)\big) + \theta\big([Y,Z]\big)\,,$$

for all $Y,Z\in\mathfrak{X}(T^*Q)$, where θ is the Liouville 1-form. The Lie derivative of θ along Y, on the other hand, is:

$$\mathfrak{L}_Y\theta(Z) = \mathfrak{L}_Y\big(\theta(Z)\big) - \theta(\mathfrak{L}_YZ) = Y\big(\theta(Z)\big) - \theta\big([Y,Z]\big)\,.$$

Combining these two equations, we obtain:

$$\hat{V}^\flat(Z) = \omega(\hat{V}, Z) = Z\big(\theta(\hat{V})\big) - \mathcal{L}_{\hat{V}}\theta(Z) = Z\big(\theta(\hat{V})\big) = \mathrm{d}\big(\theta(\hat{V})\big)(Z)\,, \tag{5.64}$$

where the symplectic vector field $\hat{V}$ preserves the symplectic 1-form, and hence $\mathcal{L}_{\hat{V}}\theta = 0$. Accordingly, the antifundamental vector field $\hat{V}$ is a Hamiltonian vector field:

$$\gamma(V) = \hat{V} = X_{\theta(\hat{V})}\,. \tag{5.65}$$

While this proves that the 'Diff(Q) part' of the group action is Hamiltonian, it would be nice if we had an explicit expression for the function $\theta(\hat{V})$ in terms of the original Lie algebra element V. To this end, remember the definition of the Liouville form from equation (4.24):

$$\theta(\hat{V})_m = m(\pi_*\hat{V})_{\pi(m)}\,,$$

where $\pi_*\hat{V}$ is now given by:

$$(\pi_*\hat{V})(f) = \hat{V}(f\circ\pi) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\pi\circ T^*\exp(tV)\,.$$

Lemma 4.5.2 states that the cotangent lift $T^*\varphi$ of a diffeomorphism φ is a vector bundle isomorphism covering φ^{-1}. The following is hence a commutative diagram:

$$\begin{array}{ccc} T^*Q & \xrightarrow{T^*\varphi} & T^*Q \\ \downarrow{\scriptstyle\pi} & & \downarrow{\scriptstyle\pi} \\ Q & \xleftarrow{\varphi} & Q \end{array}\,,$$

or, in other words, $\pi\circ T^*\varphi(m) = \varphi^{-1}\circ\pi(m)$ for all $m\in T^*Q$. Applied to the above expression for $\pi_*\hat{V}$, this yields:

$$(\pi_*\hat{V})(f) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\pi\circ T^*\exp(tV) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} f\circ\exp(-tV)\circ\pi\,.$$

Since $\exp(-tV)$ is the flow of the vector field $-V$, we get $\pi_*\hat{V}_m = -V_{\pi(m)}$. An explicit expression for the function $\theta(\hat{V})\in C^\infty(M,\mathbb{R})$ that yields the Hamiltonian vector field $\hat{V}$ in equation (5.65) is hence:

$$\theta(\hat{V})(m) = -m(V_{\pi(m)})\,. \tag{5.66}$$

For the 'function part' of the group we explained that the Lie algebra is isomorphic to, and thus can be identified with the group $C^\infty(Q,\mathbb{R})/\mathbb{R}$ itself. For a Lie algebra element α the exponential is hence $\exp(\alpha) = \alpha$. According to equation (5.53), the action of α on M is given by $\rho_{(\alpha,\mathrm{id})}(m) = m - (\mathrm{d}\alpha)_{\pi(m)}$ for all $m\in M$. The antifundamental vector field $\hat{A} := \gamma(\alpha)\in\mathfrak{X}(M)$ is therefore:

$$\hat{A}_m = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\rho_{(\exp(-t\alpha),\mathrm{id})}(m) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\big(m + t(\mathrm{d}\alpha)_{\pi(m)}\big)\,. \tag{5.67}$$

Actually, we already know this vector field to be Hamiltonian. We mentioned the relation between fundamental vector fields that arise from fibre translations and vertical

lifts in section 4.5 (specifically remark 4.5.9). Since our group action has an additional minus sign, however, it is now the *anti*fundamental vector field $\hat{A}$ that is equal to the vertical lift $V^{\mathrm{d}\alpha}$. Furthermore, part (iii) of lemma 4.5.8 states that the vertical lift $V^{\mathrm{d}\alpha}$ is Hamiltonian, and equation (4.41) yields the explicit expression:

The Hamiltonian vector field does not depend on the representative chosen for α because the definition $X_{\alpha\circ\pi} = \mathrm{d}(\alpha\circ\pi)^\sharp$ contains an exterior derivative.

$$\gamma(\alpha) = \hat{A} = V^{\mathrm{d}\alpha} = -X_{\pi^*\alpha} = -X_{\alpha\circ\pi} \,. \tag{5.68}$$

Since all the antifundamental vector fields so turned out to be Hamiltonian, the full group action (5.52) of $C^\infty(Q,\mathbb{R})/\mathbb{R}\rtimes\mathrm{Diff}(Q)$ is a Hamiltonian action. Summing up, the group $C^\infty(Q,\mathbb{R})/\mathbb{R}\rtimes\mathrm{Diff}(Q)$ with group action ρ hence possesses all the properties of a Geometric Group for the cotangent bundle T^*Q. The only difficulty is that it is infinite dimensional and hence would generate a set of infinitely many fundamental observables. Instead, a finite-dimensional subgroup whose action is still transitive will thus serve as the Geometric Group $\mathcal{G}$ in the quantization procedure. Such a subgroup will in general be a semidirect product of the form $\mathcal{G} = W\rtimes G \subseteq C^\infty(Q,\mathbb{R})/\mathbb{R}\rtimes\mathrm{Diff}(Q)$ for some group $G\subseteq\mathrm{Diff}(Q)$ and some vector space $W\subseteq C^\infty(Q,\mathbb{R})/\mathbb{R}$.

5.5.4 The Mapping *P* and the Obstruction Cocycle

It is possible to make some general statements about the obstruction cocycle in this setting. Again, we will look at the two parts of the group independently.

For an element V in the Lie algebra $\mathfrak{L}\,\mathrm{Diff}(Q)$ of the 'Diff(Q) part' of the group, the restriction of the mapping P that satisfies $\gamma = j\circ P$ is determined by equations (5.65) and (5.66) as:

$$P_D(V)(m) := P|_{\mathfrak{L}\,\mathrm{Diff}(Q)}(V)(m) := -\theta(\hat{V})(m) = m(V_{\pi(m)}) \,. \tag{5.69}$$

We will see that the corresponding obstruction cocycle vanishes. In fact, we will show that P_D is a Lie algebra morphism. For $V, W\in\mathfrak{L}\,\mathrm{Diff}(Q)$ we have:

$$\{P_D(V), P_D(W)\} = \{\theta(\hat{V}),\theta(\hat{W})\} = \omega\big(X_{\theta(\hat{V})}, X_{\theta(\hat{W})}\big) = \omega(\hat{V},\hat{W}) \,,$$

where we used equation (5.65) in the last step. Due to equation (5.64) this evaluates to:

$$\omega(\hat{V},\hat{W}) = \hat{W}\big(\theta(\hat{V})\big) = \mathcal{L}_{\hat{W}}\big(\theta(\hat{V})\big) = \mathcal{L}_{\hat{W}}\theta(\hat{V}) + \theta(\mathcal{L}_{\hat{W}}\hat{V}) = \theta\big([\hat{W},\hat{V}]\big) \,.$$

Continuing, the commutator of $\hat{W}$ and $\hat{V}$ expressed in terms of W and V is:

$$\begin{aligned}
[\hat{W},\hat{V}](f) = L_{\hat{W}}\hat{V}(f) &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(\Phi_t^{\hat{W}\#}\hat{V})(f) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\Phi_t^{\hat{W}*}\circ\hat{V}\circ\Phi_{t*}^{\hat{W}}f \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\hat{V}(f\circ\Phi_{-t}^{\hat{W}})\circ\Phi_t^{\hat{W}} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\circ\Phi_{-t}^{\hat{W}}\circ\Phi_s^{\hat{V}}\circ\Phi_t^{\hat{W}} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\circ\Phi_{-t}^{W*}\circ\Phi_s^{V*}\circ\Phi_t^{W*} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} f\circ(\Phi_t^W\circ\Phi_s^V\circ\Phi_{-t}^W)^* \\
&= \mathrm{d}f\circ\left(\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\Phi_t^W\circ\Phi_s^V\circ\Phi_{-t}^W\right)\widehat{\ } = \mathrm{d}f\circ\big([V,W]\widehat{\ }\big) \\
&= [V,W]\widehat{\ }(f) = -[V,W]^{\widehat{}}_{\mathfrak{L}}(f) = [W,V]^{\widehat{}}_{\mathfrak{L}}(f) \,,
\end{aligned}$$

where we used the notation $(A)\hat{} = \hat{A}$ for the lifted vector field corresponding to A, and where Φ_t^X is the flow of the vector field X. The final result for the Poisson bracket hence becomes:

$$\{P_D(V), P_D(W)\} = \theta\big([\hat{W}, \hat{V}]\big) = \theta\big([W, V]\hat{}_{\mathcal{L}}\big) = P\big([V, W]_{\mathcal{L}}\big) \,. \tag{5.70}$$

Accordingly, the mapping P_D is a Lie algebra morphism and the obstruction cocycle for the 'Diff(Q) part' of the group vanishes.

The 'function part' of the group taken individually is unproblematic as well. Nevertheless, there is a catch when we try to define the mapping P. In fact, for a Lie algebra element $\alpha \in \mathcal{L}(C^\infty(Q,\mathbb{R})/\mathbb{R}) \cong C^\infty(Q,\mathbb{R})/\mathbb{R}$ equation (5.68) suggests that we should define the corresponding restriction of P by:

$$P_C(\alpha) := P|_{\mathcal{L}(C^\infty(Q,\mathbb{R})/\mathbb{R})}(\alpha) \stackrel{?}{:=} \alpha \circ \pi \,.$$

This mapping, however, is ill-defined. The Lie algebra element α is an *equivalence class* and the value $P_C(\alpha)$ clearly depends on the actual representative chosen for α. The only solution to fix this problem is hence to somehow select a *specific* representative $\bar{\alpha} \in \alpha$ for each Lie algebra element α and define P_C by:

$$P_C(\alpha) := \bar{\alpha} \circ \pi \,. \tag{5.71}$$

Using this definition we can calculate the Poisson bracket. For equivalence classes $\alpha, \beta \in \mathcal{L}(C^\infty(Q,\mathbb{R})/\mathbb{R})$ pick some representatives $\bar{\alpha} \in \alpha$ and $\bar{\beta} \in \beta$, then:

$$\begin{aligned}\{P_C(\alpha), P_C(\beta)\} &= \omega\big(X_{P_C(\alpha)}, X_{P_C(\beta)}\big) \\ &= \omega\big(X_{\bar{\alpha}\circ\pi}, X_{\bar{\beta}\circ\pi}\big) = \omega(\hat{A}, \hat{B}) = -\mathcal{L}_{\hat{A}}\theta(\hat{B}) = 0 \,,\end{aligned}$$

where we used the notation $\hat{A} := \gamma(\alpha)$ and $\hat{B} := \gamma(\beta)$ for the antifundamental vector fields as before. Since the commutator of $\mathcal{L}(C^\infty(Q,\mathbb{R})/\mathbb{R})$ is trivial, i. e. $[\alpha, \beta] = 0$ for all α, β, the mapping P_C is a Lie algebra morphism so that the obstruction cocycle vanishes for the 'function part' of the group as well.

A problem, however, arises from mixing both parts. For $m \in M$ we get the Poisson bracket:

$$\begin{aligned}\{P_D(V), P_C(\alpha)\}(m) &= \omega(\hat{V}, \hat{A})(m) = \mathrm{d}(\bar{\alpha} \circ \pi)(\hat{V})(m) \\ &= \hat{V}_m(\bar{\alpha} \circ \pi) = (\pi_* \hat{V}_m)(\bar{\alpha}) = -V_{\pi(m)}(\bar{\alpha}) = -V(\bar{\alpha}) \circ \pi(m) \,.\end{aligned}$$

On the other hand, equation (5.62) yields the commutator of V and α as:

$$[V, \alpha] = \big[(0, V), (\alpha, 0)\big] = \big(-V(\alpha), 0\big) \,.$$

Since the result is in the 'function part' of the Lie algebra, we obtain:

$$P\big([V, \alpha]\big) = P_C\big(-V(\alpha)\big) = -\overline{V(\alpha)} \circ \pi \,.$$

Consequently, the obstruction cocycle in this case is:

$$\begin{aligned}z\big((0, V), (\alpha, 0)\big) &= \{P(0, V), P(\alpha, 0)\} - P\big([(0, V), (\alpha, 0)]\big) \\ &= -\big(V(\bar{\alpha}) - \overline{V(\alpha)}\big) \circ \pi \,.\end{aligned} \tag{5.72}$$

The obstruction cocycle hence vanishes if and only if the representative $\overline{V(\alpha)}$ for the equivalence class $V(\alpha)$ is equal to the function $V(\bar{\alpha})$. Nevertheless, the representative chosen for $V(\alpha)$ must match the choice made for some $\beta \in C^\infty(Q, \mathbb{R})/\mathbb{R}$. The choice for $V(\beta)$ must match the specific representative chosen for some γ, and so on. Although this might be possible for some specific subgroups of $\mathrm{Diff}(Q)$ and $C^\infty(Q, \mathbb{R})/\mathbb{R}$, we are almost guaranteed to run into problems in the general case.

6 Quantization on the Half-Line

> Whatever I have up till now accepted as most true I have acquired either from the senses or through the senses. But from time to time I have found that the senses deceive, and it is prudent never to trust completely those who have deceived us even once.
>
> *René Descartes*, Meditations on First Philosophy

In this chapter we will study in depth the question how to quantize a particle that is restricted to a half-line. Obviously, the configuration space of such a particle is given by $\mathbb{R}^+$ and the corresponding classical phase space must clearly be $T^*\mathbb{R}^+$. Quantization on $T^*\mathbb{R}^+$, however, has already been discussed by Isham in [Ish83], and it would seem pointless to dedicate an entire chapter to a solved problem.

Anyway, the toy model of a particle moving on a half-line provides us with the rare opportunity to verify the result and we will see that the quantum theory constructed in [Ish83] offers only an incomplete description of this problem. As guiding principle, we expect a sensible quantum theory for the half-line model to be in some sense equivalent to the quantum theory obtained if the particle gets restricted to $\mathbb{R}^+ \subset \mathbb{R}$ by a 'hard wall' potential, but this relation is left completely in the dark. In fact, the representation of the quantum operators in [Ish83] requires a scale-invariant measure on $\mathbb{R}^+$ which suggests a rather different interpretation. We will present Isham's quantization attempt based on $T^*\mathbb{R}^+$ in section 6.1 and discuss issues that arise in section 6.2. In particular, we will establish the relation to the reference model obtained by restriction to $\mathbb{R}^+$.

Once the connection to the reference model is made clear, section 6.3 will consider the same problem from its quantum-mechanical perspective and we will see that the quantum theory obtained by quantization starting with $T^*\mathbb{R}^+$ is insufficient to explain the reflection at the end of the half-line. Isham's main intention in [Ish83] is quantum gravity for which $\mathbb{R}^+$ can be seen as special case of $\mathrm{GL}^+(n,\mathbb{R})$, so this shortcoming did probably not occur to him. Still, in [Ish83, section 2.2.5] he introduced the half-line by means of a potential and the contact interaction with the 'hard wall' is quite interesting in its own right [FCT02; GK04; BW10]. Surprisingly, there exists a self-adjoint quantum operator $\hat{p}^2$ over $\mathbb{R}^+$, yet the classical observable p^2 on $T^*\mathbb{R}^+$ cannot be quantized.

We will show that this is not a shortcoming of the quantization technique. Instead, we will construct a different phase space that implements the classical contact interaction *topologically*: the orbifold $\mathbb{R}^2/\mathbb{Z}_2$. Section 6.4 discusses quantization on $\mathbb{R}^2/\mathbb{Z}_2$ using covering groups of $\mathrm{SO}^\uparrow(1,2)$. In particular, we will identify a specific Canonical Group that reproduces the desired self-adjoint operator $\hat{p}^2$ on $\mathbb{R}^+$.

6.1 Isham's Quantization on the Phase Space $T^*\mathbb{R}^+$

We begin with an account of Isham's quantization on the phase space $T^*\mathbb{R}^+$ – the cotangent bundle over the configuration space $Q = \mathbb{R}^+$ of strictly positive real numbers. Isham's original treatment can be found in [Ish83, section 4.5] and the calculations turn out to be pretty straightforward from a mathematical point of view. In particular, there is no need to define the Canonical Group via a central extension. The present section primarily provides a walk-through of these calculations and shows how the quantization procedure can be applied to formally construct a quantum theory on $\mathbb{R}^+$. We will also intersperse some remarks where problems arise but leave the discussion for later.

6.1.1 Step 1: Geometric Group and Group Action

As for conventional quantum mechanics, the first step is to find an appropriate Geometric Group $\mathcal{G}$. Now following the results of section 5.5, though, we will not try to determine the whole group with its action directly. Instead, we begin with a Lie group G that acts on the *configuration space* $Q = \mathbb{R}^+$ in a 'natural' way and then use the cotangent lift to obtain a group action on $M = T^*Q$. Afterwards, this action will be augmented by fibre translations to obtain a suitable Geometric Group $\mathcal{G}$.

To start with, let us have a closer look at the classical phase space. Since the fibres of the cotangent bundle $M = \mathbb{T}^*\mathbb{R}^+$ in this case have to be 1-dimensional vector spaces, M is isomorphic to $M = T^*\mathbb{R}^+ \cong \mathbb{R}^+ \times \mathbb{R}$. The isomorphism is a global Darboux chart:

We write q for points in Q, but we will see that q may require a physical interpretation different from a spatial position!

$$(q, p) : M \to \mathbb{R}^2 , \quad m \mapsto \big(q(m), p(m)\big) , \tag{6.1}$$

in which the symplectic form ω takes on the usual form:

$$\omega = \mathrm{d}q \wedge \mathrm{d}p .$$

Now for the group action. First, remember that a smooth group action ρ of a Lie group G on the configuration space Q sends group elements $g \in G$ to diffeomorphisms $\rho_g \in \mathrm{Diff}(Q)$. For this reason we cannot simply use translations $q \mapsto q + u$, like we did for conventional quantum mechanics. For large enough negative values of $u \in \mathbb{R}$ the translations are no longer in $\mathrm{Diff}(\mathbb{R}^+)$. On the other hand, we cannot restrict ourselves to only positive values of u because $\mathbb{R}^+$ is not a group under addition.

Instead, Isham points out that the multiplicative group $G = (\mathbb{R}^+, \cdot)$ acts naturally on the configuration space $Q = \mathbb{R}^+$, via:

Mind the qualitative difference: In opposition to the neutral element $e = 1$ of $G = \mathbb{R}^+$ when seen as a group, there is no preferred point in the configuration space $Q = \mathbb{R}^+$.

$$\rho : G \times Q \to Q , \quad (\lambda, q) \mapsto \rho_\lambda(q) := \lambda q . \tag{6.2}$$

Remark 6.1.1 Although this group action is a natural choice from a mathematical point of view, from a physical perspective it looks strange to multiply position coordinates with numbers. We will discuss this point in the next section.

Given the group action ρ, we apply the machinery from the last section and see what happens. The cotangent lift $\rho^\uparrow$ of the group action ρ is given by:

$$\rho^\uparrow_\lambda = T_*\rho_\lambda = T^*\rho_{\lambda^{-1}} .$$

In terms of the Darboux chart from above this is:

This explicit expression shows nicely how the scaling by λ in the q-direction combined with the inverse scaling by λ^{-1} in the p-direction preserves phase space volumes.

$$\rho^\uparrow_\lambda(q, p) = (\lambda q, \lambda^{-1} p) . \tag{6.3}$$

A fibre translation τ looks identical to the one we used in the example of conventional quantum mechanics:

$$\tau_v(q, p) = (q, p - v) , \quad v \in \mathbb{R} .$$

The group W behind this action is the Abelian group $(\mathbb{R}, +)$.

Both groups can be combined via the semidirect product to obtain the Geometric Group $\mathcal{G} = W \rtimes G = \mathbb{R} \rtimes \mathbb{R}^+$, which acts on phase space $M = T^*\mathbb{R}^+$ by the (according to its construction) *Hamiltonian*, *transitive* and *effective* action:

The group action on phase space is shown in figure 6.1 on the next page.

$$\rho : \mathcal{G} \times M \to M , \quad \rho_{(v,\lambda)}(q, p) := (\lambda q, \lambda^{-1} p - v) . \tag{6.4}$$

The multiplication law of $\mathcal{G}$, which agrees with this group action, is:

$$(v_2, \lambda_2) * (v_1, \lambda_1) = \big(v_2 + \lambda_2^{-1} v_1, \lambda_2 \lambda_1\big) , \tag{6.5}$$

for all $(v_1, \lambda_1), (v_2, \lambda_2) \in \mathcal{G}$. The neutral element of the Geometric Group $\mathcal{G}$ is given by $e = (0, 1)$, and the inverse of a group element (v, λ) is $(v, \lambda)^{-1} = (-\lambda v, \lambda^{-1})$.

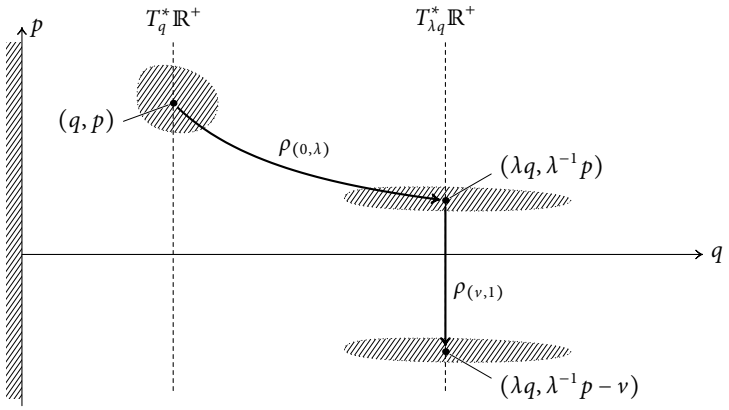


Figure 6.1 Group action of the Geometric Group $\mathcal{G} = \mathbb{R} \rtimes \mathbb{R}^+$ on classical phase space $M = T^*\mathbb{R}^+$. Note how symplectic volume elements are preserved.

6.1.2 Step 2: Lie Algebra and Antifundamental Vector Fields

THE NEXT STEP is to calculate the Lie algebra $\mathcal{LG}$ associated to $\mathcal{G}$, and to determine the antifundamental vector field mapping γ. The vector space underlying $\mathcal{LG}$ is:

$$\mathcal{LG} = T_e\mathcal{G} = T_{(0,1)}(\mathbb{R} \times \mathbb{R}^+) \cong T_0\mathbb{R} \oplus T_1\mathbb{R}^+ \cong \mathbb{R} \times \mathbb{R} = \mathbb{R}^2 \,.$$

For the Lie bracket, we start with the conjugation I_g. Explicitly, we have:

$$I_{(v,\lambda)}(w,\mu) = (v,\lambda) * (w,\mu) * (-\lambda v, \lambda^{-1}) = \left(v + \lambda^{-1}w - \mu^{-1}v, \mu\right),$$

for all $(v,\lambda), (w,\mu) \in \mathcal{G}$. Next, for Lie algebra elements $A = (A^1, A^2)$ and $B = (B^1, B^2)$ in $\mathcal{LG}$, represented by the curves $\alpha = (\alpha_v, \alpha_\lambda)$ and $\beta = (\beta_v, \beta_\lambda)$, respectively, we get:

$$\begin{aligned}\mathrm{Ad}_{(v,\lambda)}(B) &= T_e I_{(v,\lambda)}(B) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} I_{(v,\lambda)}\beta(t) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(v + \lambda^{-1}\beta_v(t) - \beta_\lambda(t)^{-1}v, \beta_\lambda(t)\right) = \left(\lambda^{-1}B^1 + B^2 v, B^2\right).\end{aligned}$$

Accordingly:

$$\begin{aligned}\mathrm{ad}(A)(B) &= T_e\,\mathrm{Ad}(A)(B) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{Ad}_{\alpha(t)}(B) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(\alpha_\lambda(t)^{-1}B^1 + B^2\alpha_v(t), B^2\right) = \left(-A^2B^1 + B^2A^1, 0\right).\end{aligned}$$

If we slightly change the notation, the Lie bracket $[\,\cdot\,,\,\cdot\,]_{\mathcal{L}} = \mathrm{ad}(\,\cdot\,)(\,\cdot\,)$ hence becomes:

$$\left[(b_1, r_1), (b_2, r_2)\right]_{\mathcal{L}} = (r_2 b_1 - r_1 b_2, 0)\,. \tag{6.6}$$

This Lie bracket is, in fact, a special case of the general formula (5.63). An important point about this result is that the Lie bracket of $\mathcal{LG}$ doesn't vanish this time – in contrast to how it was for conventional quantum mechanics.

In addition to the Lie algebra we need to find the corresponding antifundamental vector field mapping γ. In order to achieve this, our next step is to determine the exponential map $\exp : \mathfrak{LG} \to \mathcal{G}$. Given a Lie algebra element $A = [\alpha] = (b, r) \in T_e\mathcal{G}$, the left invariant vector field $L^A \in \mathfrak{X}_{\mathrm{L}}(\mathcal{G})$ generated by A, evaluated at the point $(v, \lambda) \in \mathcal{G}$ is:

$$\begin{aligned} L^A_{(v,\lambda)} &= T_e\ell_{(v,\lambda)}(A) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \ell_{(v,\lambda)}\alpha(t) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \big(v + \lambda^{-1}\alpha_v(t), \lambda\alpha_\lambda(t)\big) = (\lambda^{-1}b, \lambda r)\,. \end{aligned}$$

A short calculation yields the integral curve γ_A of the left invariant vector field L^A as:

$$\gamma_A(t) = \begin{cases} \big((b/r)(1 - e^{-rt}), e^{rt}\big) & \text{for } r \neq 0\,, \\ (bt, 1) & \text{for } r = 0\,. \end{cases}$$

Accordingly, the exponential map is:

$$\exp(b, r) = \begin{cases} \big((b/r)(1 - e^{-r}), e^{r}\big) & \text{for } r \neq 0\,, \\ (b, 1) & \text{for } r = 0\,. \end{cases} \tag{6.7}$$

It is interesting to look at those special cases in the light of the individual exponential maps of the two factors of $\mathcal{G} = \mathbb{R} \rtimes \mathbb{R}^+$.

Two important special cases are $\exp(b, 0) = (b, 1)$ and $\exp(0, r) = (0, e^r)$.

Finally, the antifundamental vector fields are easy to calculate. We get:

$$\gamma(b,0)_{(q,p)} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{\exp(-t(b,0))}(q,p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{(-tb,1)}(q,p) = (0,b) = b\partial_p\big|_{(q,p)}\,,$$

and:

$$\begin{aligned} \gamma(0,r)_{(q,p)} &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{\exp(-t(0,r))}(q,p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \rho_{(0,e^{-r})}(q,p) = (-rq, rp) \\ &= -rq\partial_q\big|_{(q,p)} + rp\partial_p\big|_{(q,p)}\,. \end{aligned}$$

Combining these two intermediate results we end up with the antifundamental vector field mapping γ according to:

$$\gamma(b, r) = -rq\partial_q + (b + rp)\partial_p\,, \tag{6.8}$$

for all $(b, r) \in \mathfrak{LG}$.

6.1.3 Step 3: Construction of the Canonical Group

Obtaining the Canonical Group $\mathcal{C}$ is now just a matter of following the recipe given in figure 5.2. As before, the first step is to find a linear mapping P that satisfies $j \circ P = \gamma$. Since d and $\sharp$ are both 'local' operations, the mapping j looks identical to the one in equation (5.34):

$$j(f) = -X_f = -\mathrm{d}f^\sharp = -(\partial_p f)\partial_q + (\partial_q f)\partial_p\,, \tag{6.9}$$

only the function f now lives in $C^\infty(T^*\mathbb{R}^+, \mathbb{R})$ instead of in $C^\infty(T^*\mathbb{R}, \mathbb{R})$. Equating the corresponding coefficients between equations (6.8) and (6.9) leaves us with:

$$-rq\partial_q + (b + rp)\partial_p \stackrel{!}{=} -(\partial_p f)\partial_q + (\partial_q f)\partial_p \quad \Leftrightarrow \quad rq = \partial_p f\,, \quad (b + rp) = \partial_q f\,.$$

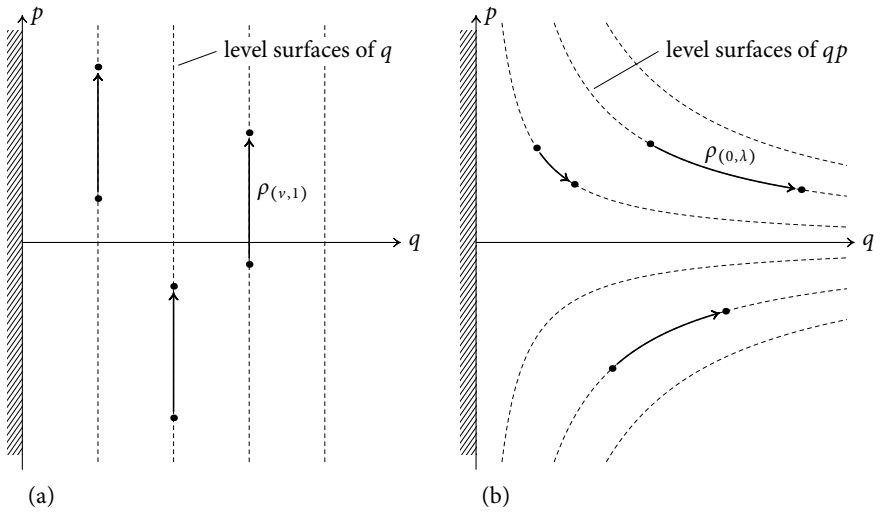


Figure 6.2 Remark 4.2.3 can be used to gain an intuitive understanding how the group action of $\mathbb{R} \rtimes \mathbb{R}^+$ on the phase space $T^*\mathbb{R}^+$ is related to the fundamental observables. In fact, the one-parameter subgroups of $\mathbb{R} \rtimes \mathbb{R}^+$ which generate the observables q and qp are realised as transformations along the *level surfaces* of q and qp, respectively.

Up to some constant, this set of partial differential equations is solved by $f = bq + rqp$. Hence we will define the searched-for mapping P as:

$$P : \mathcal{LG} \to C^\infty(M, \mathbb{R}) , \quad (b, r) \mapsto P\big((b, r)\big) := bq + rqp . \tag{6.10}$$

To check whether P is a Lie algebra morphism or not, we calculate the obstruction cocycle z. For Lie algebra elements $A_1 = (b_1, r_1)$ and $A_2 = (b_2, r_2)$ in $\mathcal{LG}$, the Poisson bracket of their images under P yields:

$$\begin{aligned}\big\{P(A_1), P(A_2)\big\} &= \omega\big(\gamma(A_1), \gamma(A_2)\big) = \mathrm{d}q \wedge \mathrm{d}p\big(\gamma(b_1, r_1), \gamma(b_2, r_2)\big) \\ &= -r_1 q(b_2 + r_2 p) + r_2 q(b_1 + r_1 p) \\ &= (r_2 b_1 - r_1 b_2) q .\end{aligned}$$

On the other hand, the image of the commutator is:

$$\begin{aligned}P\big([A_1, A_2]\big) &= P\big(\big[(b_1, r_1), (b_2, r_2)\big]\big) = P\big((r_2 b_1 - r_1 b_2, 0)\big) \\ &= (r_2 b_1 - r_1 b_2) q .\end{aligned}$$

The obstruction cocycle $z(A_1, A_2)$ hence vanishes for all $A_1, A_2 \in \mathcal{LG}$ so that the (a priori only known to be linear) mapping P turns out to be a Lie algebra morphism. According to the discussion in section 5.3 we can under these circumstances choose $\mathcal{LC}$ equal to $\mathcal{LG}$, which gives $\mathcal{C} = \mathcal{G}$. The algebra of fundamental observables can be defined as $\mathrm{Obs}(M) := \mathrm{im}\, P$ and is generated by the functions q and qp (see also figure 6.2).

6.1.4 Step 4: Representations and Commutation Relations

In order to study the unitary representations of the Canonical Group $\mathcal{C} = \mathbb{R} \rtimes \mathbb{R}^+$, assume that $U' : \mathcal{C} \to U(\mathcal{H})$ is a strongly continuous, unitary, irreducible representation

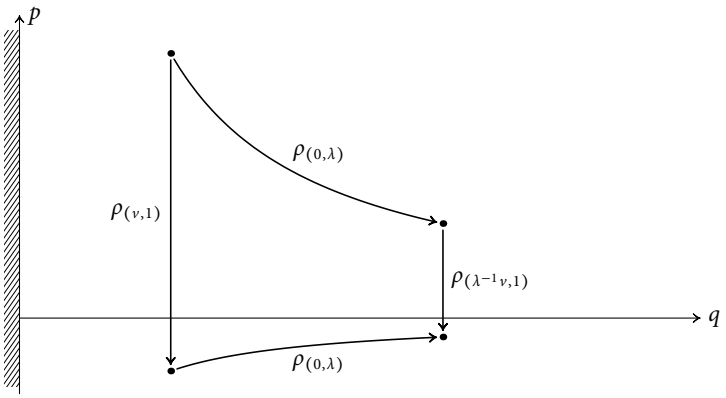


Figure 6.3 Concerning the commutation relation (6.12) of $U(\lambda)$ and $V(b)$. The transformation $\rho_{(0,\lambda)}$ scales p-coordinates by λ^{-1}, so that the transformations carried out in the opposite order require a *different* fibre translation, namely by $\lambda^{-1}v$. The transformation $\rho_{(0,\lambda)}$, however, stays the *same* for the opposite order. A consequence of this is the unitary equivalence $V(b) \sim V(\lambda^{-1}b)$ for any choice of $\lambda \in \mathbb{R}^+$ (see equation (6.20)).

The calculation is possible and yields the same end result if we use $\tilde{U}(r) := U'(\exp(0,r))$ instead of $U(\lambda)$. The choice of $U(\lambda)$ is merely a matter of convenience.

of $\mathcal{C}$ on some complex separable Hilbert space $\mathcal{H}$. Isham in [Ish83, section 4.5.2] then goes on and defines 1-parameter subgroups of unitary operators by:

$$U(\lambda) := U'\big((0,\lambda)\big)\,, \quad V(b) := U'\big(\exp(b,0)\big)\,,$$

for $\lambda \in \mathbb{R}^+$ and $b \in \mathbb{R}$. The commutation relations for these operators are:

$$\begin{aligned} U(\lambda_1)U(\lambda_2) &= U'\big((0,\lambda_1) * (0,\lambda_2)\big) = U(\lambda_1\lambda_2)\,, \\ V(b_1)V(b_2) &= U'\big((b_1,1) * (b_2,1)\big) = U'\big((b_1+b_2,1)\big) = V(b_1+b_2)\,, \end{aligned}$$

as well as:

$$\begin{aligned} U(\lambda)V(b) &= U'\big((0,\lambda) * (b,1)\big) = U'\big((\lambda^{-1}b,\lambda)\big) \\ &= U'\big((\lambda^{-1}b,1) * (0,\lambda)\big) = V(\lambda^{-1}b)U(\lambda)\,. \end{aligned}$$

In summary (see also figure 6.3):

$$U(\lambda_1)U(\lambda_2) = U(\lambda_1\lambda_2)\,, \quad V(b_1)V(b_2) = V(b_1+b_2)\,, \tag{6.11}$$

$$U(\lambda)V(b) = V(\lambda^{-1}b)U(\lambda)\,. \tag{6.12}$$

Again, Stone's theorem guarantees that there exist self-adjoint generators $\hat{q}$ and $\hat{\pi}$ of the one-parameter unitary groups $U(\lambda)$ and $V(b)$, such that:

$$e^{-ir\hat{\pi}} = U(e^r) = U'\big(\exp(0,r)\big)\,, \quad e^{-ib\hat{q}} = V(b) = U'\big(\exp(b,0)\big)\,, \tag{6.13}$$

If we also have an explicit representation, the quantization map will be determined by:

$$\mathcal{Q} : bq + rqp \mapsto b\hat{q} + r\hat{\pi}\,. \tag{6.14}$$

The quantum operator $\hat{\pi}$ hence corresponds to the classical observable qp, the operator $\hat{q}$ to the observable q. Given the quantization map, the commutation relations of $\hat{q}$ and

$\hat{\pi}$ (valid on a dense subset of wave functions) follow from the classical Poisson brackets of q and qp as:

$$[\hat{q}, \hat{q}] = 0 \,, \quad [\hat{\pi}, \hat{\pi}] = 0 \,, \quad [\hat{q}, \hat{\pi}] = i\hat{q} \,. \tag{6.15}$$

Remark 6.1.2 (Constant functions) In contrast to conventional quantum mechanics the commutation relations (6.15) do not include constant functions as an essential part. As mentioned earlier, this is possible because the observables q and qp are not canonically conjugate to each other.

What remains is to find a *concrete* representation for the quantum operators. Isham remarks that a representation in this case can be constructed using the analogy to the Schrödinger representation of conventional quantum mechanics.

Recall that in section 5.4, the example of conventional quantum mechanics in one dimension, wave functions were elements of the Hilbert space $L^2(\mathbb{R}, \mathrm{d}q)$. The groups of unitary transformations $U(a)$ and $V(b)$ from conventional quantum mechanics act on these wave functions according to:

$$\big(U(a)\psi\big)(q) = \psi(q - \mu a) \,, \quad \big(V(b)\psi\big)(q) = e^{-ibq}\psi(q) \,.$$

The main difference of the current example is that the group $\mathbb{R}^+$ is not additive but *multiplicative*. Isham hence deems it reasonable to define the following realisation:

$$\big(U(\lambda)\psi\big)(q) := \psi(\lambda^{-1}q) \,, \quad \big(V(b)\psi\big)(q) := e^{-ibq}\psi(q) \,, \tag{6.16}$$

where the wave functions ψ are square-integrable, complex-valued functions over the configuration space $Q = \mathbb{R}^+$.

However, Isham remarks in [Ish83, section 4.5.2] that the operator $U(\lambda)$ given in equation (6.16) isn't unitary if the inner product of $\mathcal{H}$ is defined in the usual way via the Lebesgue measure $\mathrm{d}q$ on $\mathbb{R}^+$. He explains that in order to solve this problem we have to introduce a different, scale-invariant, measure $\mathrm{d}q/q$ on $\mathbb{R}^+$ and define the inner product of wave functions φ, ψ by:

Mathematically, the measure $\mathrm{d}q/q$ is a left and right invariant Radon measure on $(\mathbb{R}^+, \cdot\,)$ (see for example [Els11, chapter VIII, example 3.10]).

$$\langle \psi | \varphi \rangle := \int_{\mathbb{R}^+} \psi^*(q)\varphi(q)\,\frac{\mathrm{d}q}{q} \,.$$

The operators $U(\lambda)$ and $V(b)$ are both unitary if the Hilbert space $\mathcal{H}$ uses this new inner product. According to Isham, the appropriate Hilbert space for the representations of the Canonical Group $\mathcal{C} = \mathbb{R} \rtimes \mathbb{R}^+$ is hence:

$$\mathcal{H} = L^2(\mathbb{R}^+, \mathrm{d}q/q) \,. \tag{6.17}$$

The generators $\hat{q}$ and $\hat{\pi}$ are given explicitly on this Hilbert space by:

$$(\hat{q}\psi)(q) = q\psi(q) \,, \quad (\hat{\pi}\psi)(q) = -iq(\partial_q\psi)(q) \,. \tag{6.18}$$

Remark 6.1.3 In contrast to $\hat{p} = -i\partial_q$, the operator $\hat{\pi}$ in equation (6.18), defined on a suitable domain, is a self-adjoint operator on $L^2(\mathbb{R}^+, \mathrm{d}q/q)$. Intuitively, the additional factor of q in $\hat{\pi} = -iq\partial_q$ works in such a way that the transformations generated by $\hat{\pi}$ 'slow down' the nearer one gets to the origin $q = 0$. It is therefore no longer possible to 'shift wave functions past the origin', which is what prevents the momentum operator $\hat{p}$ from being self-adjoint on $\mathbb{R}^+$ (see remark 1.2.4).

Note that $\hat{\pi}$ as in (6.18) is not even symmetric if $\mathbb{R}^+$ uses the Lebesgue measure $\mathrm{d}q$.

Remark 6.1.4 (Concerning the unusual measure) A problem with this result lies in the necessity of the unusual measure $\mathrm{d}q/q$ on $\mathbb{R}^+$. The operator $\hat{q}$ given in equation (6.18) is diagonal and associated to the classical position observable which suggests that we are in the position space representation. On the other hand, it if we want to interpret the result as particle on a half-line it makes sense to calculate distances between different spatial positions and calculating distances requires a metric. Furthermore, when $\mathbb{R}^+$ is really the position space it should be possible to view $\mathbb{R}^+$ as submanifold of the full line $\mathbb{R}$ and the metric on $\mathbb{R}^+$ should hence be the usual Euclidean metric. Nevertheless, the measure $\mathrm{d}q/q$ is incompatible with the Euclidean metric. This certainly raises some doubts regarding the interpretation of $\mathbb{R}^+$ as position space.

Postponing the issues with the measure $\mathrm{d}q/q$ until the next section, we need to find out whether there are additional, unitarily inequivalent representations of the Canonical Group $\mathcal{C}$. For this, observe that the commutation relations (6.15) of $\hat{q}$ and $\hat{\pi}$ still hold if we replace $\hat{q}$ by $\alpha\hat{q}$, for some arbitrary value of $\alpha \in \mathbb{R}$. Hence there is actually a whole family of unitary representations:

$$\big(U(\lambda)\psi\big)(q) := \psi(\lambda^{-1}q)\,, \quad \big(V(b)\psi\big)(q) := e^{-i\alpha bq}\psi(q)\,, \tag{6.19}$$

for $\psi \in \mathcal{H} = L^2(\mathbb{R}^+, \mathrm{d}q/q)$, parameterised by $\alpha \in \mathbb{R}$.

Nevertheless, not all the representations are inequivalent. Due to the commutation relation (6.12) we have (see also figure 6.3):

$$U^{-1}(\lambda)V(b)U(\lambda) = V(\lambda b) \quad \forall\lambda \in \mathbb{R}^+\,. \tag{6.20}$$

Thus we see that the unitary operator $V(b)$ in equation (6.19) is unitarily equivalent to $V'(b) = e^{-i\alpha\lambda bq}$ for an arbitrary value of $\lambda \in \mathbb{R}^+$. Accordingly, the representations fall into three classes:

- representations with $\alpha > 0$,
- representations with $\alpha < 0$,
- the representation with $\alpha = 0$.

Representations within the same class are unitarily equivalent. If we look at the spectrum of $\hat{q} = \alpha q$, on the other hand, which is equal to $\mathbb{R}^+$, $\mathbb{R}^-$ or $\{0\}$, respectively, it becomes clear that unitary transformations *between* different classes cannot exist.

Remark 6.1.5 (Physical relevance of α) In the case of conventional quantum mechanics there appeared a parameter $\mu \in \mathbb{R}$ during the study of the representations, at the end of section 5.4. There we argued that the choice of μ influences the physical predictions and hence must be a measurable physical quantity. Consequently, we explained that μ introduces a *preferred scale* into the theory, which we later identified with $\hbar$.

The parameter $\alpha \in \mathbb{R}$ in this example is of a different quality. Due to the unitary equivalence (6.20) it matters only whether α is positive, negative or zero, but the actual magnitude of α is *physically irrelevant*! A value of $\alpha = 1$ gives exactly the same physical predictions as, say, $\alpha = 42$. In contrast to conventional quantum mechanics, there is therefore *no preferred scale* in the quantum theory that we obtained via the Geometric Group $\mathcal{G} = \mathbb{R} \rtimes \mathbb{R}^+$ acting on $T^*\mathbb{R}^+$.

This approach isn't in general applicable to construct all possible representations of a given group. Nevertheless, Isham assures that no additional irreducible representations exist [Ish83, section 4.5.2]:

> '*A priori*, there might be many other irreducible representations, but when Mackey's techniques are applied to the group $[\mathbb{R} \rtimes \mathbb{R}^+]$ they show that the three we have just found—[…]—are the entire set.'

Remark 6.1.6 The supposed nonexistence of other irreducible representations raises another question. For a particle on a half-line we expect it possible to obtain the quantum theory in an alternative manner as some kind of 'restriction' of the well-known quantum theory for a particle on the full line. In particular, while we expect position space wave functions to be functions over $\mathbb{R}^+$, it should be possible to inherit the *usual* inner product from conventional quantum mechanics over $\mathbb{R}$.

Isham's statement, however, suggests that there is no way to get rid of the unusual measure $\mathrm{d}q/q$ on $\mathbb{R}^+$, which would mean that it is impossible to obtain the quantum theory for a particle on a half-line as a restriction. We will see in the following section that this isn't true. We will actually construct a representation for the operators $\hat{q}$ and $\hat{\pi}$ via a restriction so that the inner product is the usual one, which yields a representation of $\mathbb{R} \rtimes \mathbb{R}^+$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$.

6.1.5 Physical Units

It is again possible to replace the dimensionless operators $\hat{q}$ and $\hat{\pi}$ by some new operators $\hat{q}_\phi$ and $\hat{\pi}_\phi$ carrying physical dimension. Like at the end of section 5.4, we can choose base units q_0 and π_0 and define:

$$\hat{q}_\phi := q_0 \hat{q} \,, \quad \hat{\pi}_\phi := \pi_0 \hat{\pi} \,. \tag{6.21}$$

The canonical commutation relations of these operators are:

$$[\hat{q}_\phi, \hat{q}_\phi] = 0 \,, \quad [\hat{\pi}_\phi, \hat{\pi}_\phi] = 0 \,, \quad [\hat{q}_\phi, \hat{\pi}_\phi] = i\pi_0 \hat{q}_\phi \,. \tag{6.22}$$

What is strange about these commutation relations is that the base unit π_0 appears *explicitly* on the right-hand side of the commutator. In contrast to representations of the Heisenberg group, there is this time no parameter like μ that can be used to 'absorb' the base unit π_0. The unit q_0, on the other hand, doesn't appear at all in equation (6.22) and can be chosen at will. We can replace the operator $\hat{q}_\phi$ by $\lambda \hat{q}_\phi$ for any $\lambda \in \mathbb{R}^+$ without measurable effect since this amounts to choosing a unitarily equivalent representation. Consequently, *the magnitude of $\hat{q}_\phi$ can not be measured using the operators we have*!

It is interesting to consider the current situation in the light of remark 5.4.1.

6.1.6 A Canonical Transformation

Isham presents an argument in [Ish83, section 4.5.4] with intention to show that π_0 in equation (6.22) must be chosen equal to $\hbar$. His argument is based on the existence of a canonical transformation between the phase spaces $T^*\mathbb{R}$ and $T^*\mathbb{R}^+$ – essentially the cotangent lift of the exponential map $\exp : \mathbb{R} \to \mathbb{R}^+$, only that we have to take care of the dimensions if we want to map *physical* operators. Accordingly, if we temporarily write $\hat{\tilde{q}}$ for the operator $\hat{q}_\phi$ on $\mathbb{R}^+$ and $\hat{q}$ for the operator $\hat{q}_\phi$ on $\mathbb{R}$ from section 5.4 we have the following diffeomorphism between the configuration spaces:

$$\Phi : \mathbb{R} \to \mathbb{R}^+ \,, \quad \Phi(q) := \tilde{q}_0 e^{q/q_0} \,, \tag{6.23}$$

The existence of a *global* canonical transformation in this case is a coincidence. In general, a phase space M is only *locally* symplectomorphic to a subset of $T^*\mathbb{R}^{2n}$ (Darboux's theorem).

where $\tilde{q}_0$ is the base unit of $\hat{\tilde{q}}$ in $\mathbb{R}^+$ and q_0 the base unit of $\hat{q}$ in $\mathbb{R}$. The inverse is:

$$\Phi^{-1} : \mathbb{R}^+ \to \mathbb{R} \,, \quad \Phi^{-1}(\tilde{q}) = q_0 \ln(\tilde{q}/\tilde{q}_0) \,.$$

The cotangent lift yields symplectomorphisms between the cotangent bundles:

$$\begin{aligned} T^*\Phi : T^*\mathbb{R}^+ \to T^*\mathbb{R}\,, &\quad (\tilde{q},\tilde{p}) \mapsto \big(q_0 \ln(\tilde{q}/\tilde{q}_0), \tilde{q}\tilde{p}/q_0\big)\,, \\ T^*\Phi^{-1} : \ T^*\mathbb{R} \to T^*\mathbb{R}^+, &\quad (q,p) \mapsto \big(\tilde{q}_0 e^{q/q_0}, e^{-q/q_0} q_0 p/\tilde{q}_0\big)\,. \end{aligned} \tag{6.24}$$

$$\begin{array}{ccc} T^*\mathbb{R} & \xleftarrow{T^*\Phi} & T^*\mathbb{R}^+ \\ {\scriptstyle \rho_{(u,v)}}\downarrow & \Rightarrow & \downarrow{\scriptstyle \tilde{\rho}_{(u,v)}} \\ T^*\mathbb{R} & \xrightarrow[T^*\Phi^{-1}]{} & T^*\mathbb{R}^+ \end{array}$$

This canonical transformation allows us to use the group action $\rho : \mathcal{G} \times T^*\mathbb{R} \to T^*\mathbb{R}$ of the Geometric Group $\mathcal{G} = (\mathbb{R}^2, +)$ from conventional quantum mechanics to induce a group action $\tilde{\rho} : \mathcal{G} \times T^*\mathbb{R}^+ \to \mathbb{R}^+$, acting on the phase space $T^*\mathbb{R}^+$, via:

$$\tilde{\rho}_{(u,v)}(\tilde{q},\tilde{p}) := T^*\Phi^{-1} \circ \rho_{(u,v)} \circ T^*\Phi(\tilde{q},\tilde{p}) = \big(\tilde{q}e^{u/q_0}, e^{-u/q_0}(\tilde{p} - q_0 v/\tilde{q})\big)\,, \tag{6.25}$$

for all $(\tilde{q},\tilde{p}) \in T^*\mathbb{R}^+$ and all $(u,v) \in \mathcal{G} = (\mathbb{R}^2,+)$. This group action certainly isn't what one would call 'natural' from a mathematical point of view. Still, it is a transitive, effective and Hamiltonian action on $T^*\mathbb{R}^+$, which is what we need for the quantization procedure.

When applying the recipe, we can reuse most of the results from section 5.4. The obstruction cocycle z doesn't vanish and is no coboundary so that the Canonical Group $\mathcal{C}$, which then has to be a central extension of $\mathcal{G}$, again turns out to be the Heisenberg group. We obtain a Lie algebra morphism:

$$\hat{P} : \mathfrak{LC} \to C^\infty(M,\mathbb{R})\,, \quad \hat{P}(a,b,r) = a(\tilde{q}\tilde{p}/q_0) + bq_0 \ln(\tilde{q}/\tilde{q}_0) + r\,,$$

and the quantization map can be defined as:

$$\mathcal{Q} : a(\tilde{q}\tilde{p}/q_0) + bq_0 \ln(\tilde{q}/\tilde{q}_0) + r \mapsto a\hat{p} + b\hat{q} + i\hbar\mathbb{1}\,, \tag{6.26}$$

where $\hat{q}$, $\hat{p}$ and $\mathbb{1}$ satisfy the same commutation relations as in conventional quantum mechanics, $[\hat{q},\hat{p}] = i\hbar\mathbb{1}$, but the operators are now defined on $\mathcal{H}(\mathbb{R}^+, \mathrm{d}q/q)$.

Note that Isham's argument is based on the mathematical form of a single observable alone. We consider this to be problematic at best because different quantizations using different Geometric groups might yield physically *incompatible* quantum theories so that it doesn't make sense to compare the operators.

Isham argues that the product $\tilde{q}\tilde{p}$ is a quantizable observable for *both* Geometric Groups, $\mathcal{G} = \mathbb{R}^2$ as well as $\mathcal{G} = \mathbb{R} \rtimes \mathbb{R}^+$, and thus he demands that the corresponding quantum operators should match. Comparing the quantization maps (6.26) and (6.14), this means $\hat{\pi} = q_0\hat{p}$ and $\hat{\tilde{q}} = \tilde{q}_0 e^{\hat{q}/q_0}$. Accordingly, he obtains:

$$[\hat{\tilde{q}},\hat{\pi}] = [\tilde{q}_0 e^{\hat{q}/q_0}, q_0\hat{p}] = \tilde{q}_0 e^{\hat{q}/q_0} i\hbar = i\hbar\hat{\tilde{q}}\,.$$

If we go back to the notation from before, this result is equivalent to:

$$[\hat{q}_\phi, \hat{\pi}_\phi] = i\hbar\hat{q}_\phi\,, \tag{6.27}$$

and hence Isham concludes that π_0 must be equal to $\hbar$.

Remark 6.1.7 This result seems rather surprising from a physical perspective. Contrary to the way how physical units were introduced, Isham's argument says that we *cannot choose* a base unit for $\hat{\pi}_\phi$ but that the base unit π_0 *must* be chosen equal to $\hbar$.

On the other hand, the argument doesn't fix the base unit for the operator $\hat{q}_\phi$. The value q_0 can still be chosen at will because all the corresponding representations of the Canonical Group $\mathbb{R} \rtimes \mathbb{R}^+$ are unitarily equivalent. From a physical point of view this is not really what one would expect for a position coordinate of a particle.

6.2 Physical Interpretation as Particle on a Half-Line?

After going through the calculations it is now time to think about the physical interpretation of the result. Even though Isham's prime motive is quantum gravity, he repeatedly calls $\mathbb{R}^+$ the configuration space of a particle moving on the positive real line [Ish83, sections 2.2.5 and 4.5]. However, the scale-invariant measure $\mathrm{d}q/q$ which was required for unitarity is not what we expect in this case. Specifically:

- The measure $\mathrm{d}q/q$ conflicts with a Euclidean metric (remark 6.1.4),
- The quantum theory cannot be obtained from the position space representation of conventional quantum mechanics on a full line by restriction (remark 6.1.6).

As a result, it becomes questionable whether the resulting quantum theory can still be interpreted as describing a particle on a half-line. We want to investigate whether this is a problem of the quantization procedure or a fault of our expectations.

6.2.1 Physical Interpretation of Isham's Result

The reason which leads Isham to call $Q = \mathbb{R}^+$ the configuration space of a particle moving on the positive real line – a *half-line* or *semi-axis*, conceptually – is obvious. If we consider a particle moving on a full line, with configuration space $Q = \mathbb{R}$, we can impose the constraint $q > 0$ which results in the submanifold $\mathbb{R}^+ \subset \mathbb{R}$. Consequently, the function $q \in C^\infty(T^*\mathbb{R}^+, \mathbb{R})$ is a restriction of $q \in C^\infty(T^*\mathbb{R}, \mathbb{R})$ and it is reasonable to attribute to the restricted function the same physical meaning as to the unrestricted one: the position coordinate of a particle.

Nevertheless, the canonical transformation (6.24) between $T^*\mathbb{R}$ and $T^*\mathbb{R}^+$ from above paints a different picture. With regard to the canonical transformation, the base space $\mathbb{R}^+$ is *not* a submanifold of $\mathbb{R}$; instead, the phase spaces $T^*\mathbb{R}$ and $T^*\mathbb{R}^+$ are put on an equal footing – they are globally isomorphic symplectic manifolds. If we adopt this view, the function $q \in C^\infty(T^*\mathbb{R}^+, \mathbb{R})$, however, requires a different physical interpretation. Due to equation (6.23) it is no longer q but (up to some constants) the *logarithm* of $q \in C^\infty(T^*\mathbb{R}^+, \mathbb{R})$ which represents the position coordinate of a particle. A particle on a *full* line, not a half-line, though.

The exponential yields *intrinsically* positive numbers. The positivity isn't the result of a constraint (see also section 1.4).

Now, the manifolds $\mathbb{R}^+$ and $\mathbb{R}$ are diffeomorphic and hence *indistinguishable* from a mathematical point of view. From a physical point of view, however, a half-line is clearly *not* the same as a full line. We expect a half-line to 'end' at some point (whatever that may mean mathematically), whereas the full line 'doesn't end'; the full line should extend to infinity in both directions.

One possibility to implement a distinction also mathematically, as a property of the configuration space, arises when we view points in configuration space as positions and assume that *distances* between spatial positions have physical relevance. At least in the present example this seems a reasonable and not overly restrictive assumption.

Mathematically, to calculate distances we need a metric. If we introduce a Euclidean metric on the whole line $\mathbb{R}$, the two interpretations yield different metrics on $\mathbb{R}^+$:

A Euclidean metric on $\mathbb{R}$ seems reasonable if $\mathbb{R}$ represents the position of a particle on a line, but the argument doesn't really depend on this specific choice. A different metric yields the same conclusions.

(i) When $\mathbb{R}^+$ is seen as submanifold $\mathbb{R}^+ \subset \mathbb{R}$, the space $\mathbb{R}^+$ inherits the Euclidean metric $d_{\mathbb{R}}(x, y) = \|x - y\|$ from $\mathbb{R}$. The measure induced by the Euclidean metric is the Lebesgue measure $\mathrm{d}q$.

(ii) In contrast, the symplectomorphism $T^*\mathbb{R}^+ \cong T^*\mathbb{R}$ arises from the diffeomorphism $\exp : \mathbb{R} \to \mathbb{R}^+$. Since it can't matter whether we calculate distances in terms of the

original or in terms of transformed coordinates, this diffeomorphism must be an *isometry* between the configuration spaces. We hence get an induced metric $d_{\mathbb{R}^+}$ on $\mathbb{R}^+$ which satisfies:

$$d_{\mathbb{R}}(x, y) \overset{!}{=} d_{\mathbb{R}^+}\big(\exp(x), \exp(y)\big) ,$$

for all $x, y \in \mathbb{R}$. In terms of the Euclidean norm $\|\cdot\|$ on $\mathbb{R}$, this is:

$$d_{\mathbb{R}^+}(\tilde{x}, \tilde{y}) = d_{\mathbb{R}}\big(\ln(\tilde{x}), \ln(\tilde{y})\big) = \|\ln(\tilde{x}) - \ln(\tilde{y})\| = \|\ln(\tilde{x}/\tilde{y})\| , \tag{6.28}$$

for all $\tilde{x}, \tilde{y} \in \mathbb{R}^+$. In agreement with this, the measure on $\mathbb{R}^+$ becomes:

$$\mathrm{d}q \mapsto \mathrm{d}(\ln \tilde{q}) = \frac{\mathrm{d}\tilde{q}}{\tilde{q}} . \tag{6.29}$$

Remark 6.2.1 (Compatibility with the symplectic form) It is an interesting fact that the same symplectic form on phase space is compatible with *both* measures on configuration space. The given symplectomorphism between $T^*\mathbb{R}$ and $T^*\mathbb{R}^+$ not only maps $\mathrm{d}q$ to $\mathrm{d}\tilde{q}/\tilde{q}$ but also transforms $\mathrm{d}p$ to $\tilde{q}\,\mathrm{d}\tilde{p}$. Due to the multilinearity of differential forms with respect to functions, we obtain:

$$\mathrm{d}q \wedge \mathrm{d}p \mapsto (\mathrm{d}\tilde{q}/\tilde{q}) \wedge (\tilde{q}\,\mathrm{d}\tilde{p}) = \mathrm{d}\tilde{q} \wedge \mathrm{d}\tilde{p} ,$$

which is the same symplectic form that we get when $\mathbb{R}^+$ is obtained via the restriction.

The phase space provides enough freedom to absorb the different measures on base space in the 'fibre part' of the symplectic form, but we don't necessarily have this freedom in the associated quantum theory. Wave functions are defined on only 'one half' of the classical phase space. There is hence no guarantee that two classically equivalent interpretations are still both compatible with the quantized theory.

When we now take another look at the action (6.4) of $\mathcal{G} = \mathbb{R} \rtimes \mathbb{R}^+$ on $T^*\mathbb{R}^+$ we can consider the transformations at the level of the configuration space. Using the bundle projection π_Q of $T^*\mathbb{R}^+$ and an embedding of the configuration space $Q = \mathbb{R}^+$ via the zero section $q \mapsto (q, 0)$, we obtain:

$$\rho^{\downarrow} : \mathcal{G} \times Q \to Q , \quad \rho^{\downarrow}_{(v,\lambda)}(q) := \pi_Q(\lambda q, \lambda^{-1} 0 - v) = \lambda q .$$

Depending on the metric of $\mathbb{R}^+$, the 'meaning' of these transformations is different:

(i) When the configuration space carries the Euclidean metric $d_{\mathbb{R}}$ restricted to $\mathbb{R}^+$, the distance between points $x, y \in \mathbb{R}^+$ is *scaled* by such a transformation:

$$d_{\mathbb{R}}(x, y) \mapsto d_{\mathbb{R}}(\lambda x, \lambda y) = \|\lambda x - \lambda y\| = \lambda \, \|x - y\| = \lambda \, d_{\mathbb{R}}(x, y) .$$

This is what we expect if q is a spatial position.

(ii) On the other hand, using the metric $d_{\mathbb{R}^+}$ induced on $\mathbb{R}^+$ via the exponential, we get a different result:

$$d_{\mathbb{R}^+}(x, y) \mapsto d_{\mathbb{R}^+}(\lambda x, \lambda y) = \|\ln(\lambda x/\lambda y)\| = d_{\mathbb{R}^+}(x, y) .$$

Distances defined via $d_{\mathbb{R}^+}$ are *preserved* under this transformation. Accordingly, the transformation should in this case really be understood as a *translation* in unusual coordinates, not as some 'scaling'.

Given these two possibilities, the representation (6.16) of $\mathbb{R} \rtimes \mathbb{R}^+$ seems to belong only to the second interpretation, not the first (this wouldn't be a problem if we were not in the position space representation, but the realisation of $\hat{q}$ in equation (6.18) as a diagonal operator suggests that we are). In other words:

> We have strong indications that the representation of the Canonical Group $\mathbb{R} \rtimes \mathbb{R}^+$ in equation (6.16) is *not* the correct position space representation for a particle on a half-line. Instead, the result seems to describe a particle on the *whole line* – it is essentially conventional quantum mechanics over $\mathbb{R}$ with an unusual choice of coordinates.

We will see in a moment that this is really the case.

These doubts are not meant to be dismissed carelessly. Apart from conventional quantum mechanics over $\mathbb{R}^n$ it is difficult to find simple classical systems which have a quantum analogue that can be realised in practice and where the relation between classical and quantized system is easy to understand (recall the discussion of non-trivial phase spaces and boundary conditions in section 1.4). We can *hope* that the quantization procedure is justified, but, given the non-intuitive nature of quantum mechanics, results not matching our expectations could just as well be the consequence of some misguided assumptions made along the way.

It certainly seems worthwhile to investigate *why* the quantization method yields a result that doesn't really match our expectations. *We need the classical theory to provide a framework for the physical interpretation of the associated quantum theory*, and it is also important to know the *limits* of this analogy. Cases exist where classically equivalent systems yield inequivalent quantum theories. An interesting example is the harmonic oscillator, where it is possible to obtain genuinely different predictions for the ground state energy depending on whether the harmonic oscillator is quantized using Cartesian or Action-and-angle variables (see [Kas03; Kas07; Hun10]). The 'unusual' result isn't a fault of the quantization procedure, however, but a *different kind* of quantum harmonic oscillator. According to Kastrup [Kas03; Kas06; Kas07], there seem to exist applications of such models in the field of quantum optics. Also, an explicit model for a quantum oscillator with an unusual ground state eigenvalue will appear in section 6.4.

6.2.2 The Position Space Representation for a Particle on a Half-Line

Let us take a step back and consider the situation from a different perspective. From a physical point of view we expect the particle on a half-line $\mathbb{R}^+$ to be essentially the same as a particle on a whole line when confined to $\mathbb{R}^+ \subset \mathbb{R}$ by a suitable potential V:

$$V(q) = \begin{cases} \infty & q \leq 0\,, \\ 0 & q > 0\,. \end{cases} \tag{6.30}$$

This infinitely high potential cannot be seen as limiting case of a finite one in this context. For any finite step potential the momentum has a unique self-adjoint extension. For an infinite potential there exists *no* self-adjoint momentum operator [BFV01, section 7.4].

In other words, we assume that the operations of quantization and restriction *commute* in this situation (in general, however, there also exist counterexamples, e. g. [Lol90]; this assumption should not be taken too lightly!):

$$\begin{array}{ccc} \text{classical theory over } \mathbb{R} & \xrightarrow{\text{quantization}} & \text{quantum theory over } \mathbb{R} \\ \Big\downarrow{\scriptstyle \text{restriction}} & & \Big\downarrow{\scriptstyle \text{restriction}} \\ \text{classical theory over } \mathbb{R}^+ & \xrightarrow{\text{quantization}} & \text{quantum theory over } \mathbb{R}^+ \end{array}$$

To obtain a quantum theory for a particle on a half-line we will hence start with the usual position and momentum operators on $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$, acting on wave functions according to (for $\hbar = 1$):

$$(\hat{q}\psi)(q) = q\psi(q) \,, \quad (\hat{p}\psi)(q) = -i\partial_q\psi(q) \,.$$

Continuing with this train of thought, a wave function $\psi \in \mathcal{H}$ is compatible with the above potential if and only if it satisfies $\psi(q) = 0$ for all $q \leq 0$. Since functions that do not satisfy this condition cannot contribute to the solutions, we can ignore the forbidden wave functions and restrict ourselves to the subspace of allowed wave functions:

$$\tilde{\mathcal{H}} := \{\psi \in \mathcal{H} : \psi(q) = 0 \; \forall q \leq 0\} \subset \mathcal{H} \,.$$

Furthermore, it turns out that the resulting subspace is isomorphic to the Hilbert space $\mathcal{H}^+ = L^2(\mathbb{R}^+, \mathrm{d}q)$. The isomorphism is:

$$\mathcal{H}^+ \to \tilde{\mathcal{H}} \,, \quad \psi \mapsto \tilde{\psi} \,, \quad \tilde{\psi}(q) := \begin{cases} 0 & q \leq 0 \,, \\ \psi(q) & q > 0 \,. \end{cases} \tag{6.31}$$

This remark seems trivial but we will come back to this point in remark 6.3.1.

Remark 6.2.2 It might look like continuing a function in a non-continuous manner as in equation (6.31) could be problematic. Nevertheless, while functions in $L^2(\ldots)$ must be measurable, they are *not* required to be continuous. As a result, the mapping (6.31) really is an isomorphism between Hilbert spaces.

This also means that there is *no momentum-space representation* for a particle on a half-line!

Remark 6.2.3 (Concerning the momentum) We already know that the momentum is no longer self-adjoint on $\mathcal{H}^+$. Formally, the conclusion is that 'the momentum is not a measurable quantity in that situation' [BFV01, section 5.2]. While this truth at first seems almost too hard to be truly acceptable, viewing $\mathbb{R}^+$ like proposed above, as $\mathbb{R}$ together with a suitable potential, makes it more plausible. As a matter of fact, the momentum eigenfunctions – plane waves – are incompatible with the given potential and hence can't contribute to any solutions.

On the other hand, it may be possible to define something like an 'intrinsically unsharp' momentum observable in such situations (my guess is that wavelets might be a step in this direction; Morlet wavelets, for example, are basically Gaussian wave packets, and they have already been applied in conventional quantum mechanics [Ash10], albeit for different reasons). Nevertheless, we would then have to make a distinction between operators that generate dynamics and observables that appear in measurements. This would require some changes to the framework of quantum mechanics that we are not prepared to discuss in this thesis.

To construct a quantum theory with observables $\hat{q}$ and $\hat{\pi}$ for the particle on a half-line $\mathbb{R}^+ \subset \mathbb{R}$ and to find out whether a relation exists to the result of the last section it seems reasonable to proceed in the following manner:

1. Quantize the classical observables q and qp over $\mathbb{R}$ using the well-known position space representations of $\hat{q}$ and $\hat{p}$ given above,
2. Check whether the obtained operators are still self-adjoint when we restrict them to act on the subspace $\tilde{\mathcal{H}} \subset \mathcal{H}$, isomorphic to $\mathcal{H}^+$,
3. Integrate the self-adjoint generators to 1-parameter unitary groups on $\mathcal{H}^+$ and study the relation to the representation (6.16) of $\mathbb{R} \rtimes \mathbb{R}^+$.

For the first step, we already have the operator $\hat{q}$ associated to the observable q of spatial position. Concerning the observable $\pi := qp$, we know that $(AB)^\dagger = B^\dagger A^\dagger$, so neither $\hat{q}\hat{p}$ nor $\hat{p}\hat{q}$ alone can be symmetric because $\hat{q}$ and $\hat{p}$ don't commute. Instead we must take the symmetric linear combination to obtain a self-adjoint operator on $\mathcal{H}$:

$$\pi \mapsto \hat{\pi} := \tfrac{1}{2}\big(\hat{q}\hat{p} + \hat{p}\hat{q}\big) = \hat{q}\hat{p} - \tfrac{1}{2}i\,, \tag{6.32}$$

which acts on (a dense subset of) position space wave functions $\psi \in \mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$ according to:

$$(\hat{\pi}\psi)(q) = -i\big(q\partial_q + \tfrac{1}{2}\big)\psi(q)\,. \tag{6.33}$$

This result differs from the operator $\hat{\pi}$ in equation (6.18). The additional term of $-i/2$ doesn't appear in the original result.

Given the operators on $\mathcal{H}$, the next step is to find out what happens when we restrict $\hat{q}$ and $\hat{\pi}$ to operators acting on wave functions in $\mathcal{H}^+$. It is easy to see that $\hat{q}$ is self-adjoint on $\mathcal{H}^+$, but we know that the momentum operator $\hat{p}$ restricted to $\mathcal{H}^+$ is not. Nevertheless, since self-adjointness is rather complicated, this doesn't say much about $\hat{\pi}$ itself.

We first have to check that $\hat{\pi}$ is symmetric on $\mathcal{H}^+$. Integration by parts yields:

$$\begin{aligned}
\langle\psi|\hat{\pi}\varphi\rangle &= \int_0^\infty \psi^*(q)(\hat{\pi}\varphi)(q)\,\mathrm{d}q = \int_0^\infty \psi^*(q)(-i)\big(q\partial_q + \tfrac{1}{2}\big)\varphi(q)\,\mathrm{d}q \\
&= -i\int_0^\infty \psi^*(q)q\partial_q\varphi(q)\,\mathrm{d}q - \frac{i}{2}\int_0^\infty \psi^*(q)\varphi(q)\,\mathrm{d}q \\
&= -i\Big[\psi^*(q)q\varphi(q)\Big]_{q=0}^\infty + i\int_0^\infty \partial_q(q\psi)^*(q)\varphi(q)\,dq - \frac{i}{2}\langle\psi|\varphi\rangle\,.
\end{aligned}$$

Using $\partial_q(q\psi)^*(q) = \psi^*(q) + q\partial_q\psi^*(q)$, the integral in the second term evaluates to:

$$\int_0^\infty \partial_q(q\psi)^*(q)\varphi(q)\,\mathrm{d}q = \int_0^\infty (q\partial_q\psi)^*(q)\varphi(q)\,\mathrm{d}q + \int_0^\infty \psi^*(q)\varphi(q)\,\mathrm{d}q\,.$$

Reinserting this result into the above equation we get:

$$\langle\psi|\hat{\pi}\varphi\rangle = -i\Big[q\,\psi^*(q)\varphi(q)\Big]_{q=0}^\infty + \langle\hat{\pi}\psi|\varphi\rangle\,. \tag{6.34}$$

The surface term cannot be argued away for arbitrary wave functions in $L^2(\mathbb{R}^+, \mathrm{d}q)$. Nevertheless, ψ and φ are not arbitrary. They must both be elements of an appropriately chosen domain of $\hat{\pi}$. For example, $\hat{\pi}$ is symmetric on the domain $D(\hat{\pi}) = C_0^\infty(0;\infty)$ of smooth functions with compact support contained in $(0;\infty)$, and this domain is dense in $\mathcal{H}^+$ [Gro88, section 3.2].

Note that the typical argument which assumes that functions in $L^2(\mathbb{R}, \mathrm{d}q)$ vanish for $q \to \pm\infty$ is plainly wrong (see [Gieoo] for a counterexample). Even the supposedly simple quantum mechanics over $\mathbb{R}$ requires a more sophisticated argument involving operator domains.

If it were practical to specify explicitly a domain on which $\hat{\pi}$ is symmetric and *closed*, we could check for self-adjointness by calculating the deficiency indices as explained in section 1.6. Nevertheless, the condition that $\hat{\pi}$ be closed is crucial, and we would also have to determine the explicit domain of the adjoint $\hat{\pi}^\dagger$ to execute the calculation. Fortunately, there is an easier way in the current situation. Using the identity $q\partial_q = \partial_{\ln(q)}$ for $q \in \mathbb{R}^+$ we can calculate $\tilde{U}(a) := e^{-ia\hat{\pi}}$:

$$\begin{aligned}
\big(\tilde{U}(a)\psi\big)(q) &= \big(e^{-ia\hat{\pi}}\psi\big)(q) = \big(e^{-ia(\partial_{\ln(q)}+1/2)}\psi\big)(q) = e^{-a/2}\big(e^{-a\partial_{\ln(q)}}\psi\big)(q) \\
&= e^{-a/2}\sum_{n=0}^\infty \frac{(-a)^n}{n!}(\partial_{\ln(q)}^n\psi)(q) = e^{-a/2}\sum_{n=0}^\infty \frac{(-a)^n}{n!}(\partial_z^n\psi)(e^z) \\
&= e^{-a/2}\psi(e^{z-a}) = e^{-a/2}\psi(e^{-a}q)\,.
\end{aligned}$$

Although the calculation probably looks acceptable, it turns out that we have to be more careful. It is *only a formal calculation*: the usual power series which we used to calculate the exponential is not well-defined for unbounded self-adjoint operators because the series isn't guaranteed to converge on a big enough set of wave functions (see [RS.I, sections VIII.4 and VIII.5]). Nevertheless, we can in the reverse direction calculate the generator of the formally obtained $\tilde{U}(a)$ via the strong derivative to verify the result:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}(-ia)}\bigg|_{a=0}\big(\tilde{U}(a)\psi\big)(q) &= i\frac{\mathrm{d}}{\mathrm{d}a}\bigg|_{a=0}\big(e^{-a/2}\psi(e^{-a}q)\big)\\ &= i\Big[-\tfrac{1}{2}e^{-a/2}\psi(e^{-a}q) + e^{-a/2}(-q)\partial_q\psi(e^{-a}q)\Big]_{a=0}\\ &= -i\big(q\partial_q + \tfrac{1}{2}\big)\psi(q)\,.\end{aligned}$$

Since we regained the original $\hat{\pi}$ we see that the expression for $\tilde{U}(a)$ produced by the formal calculation is, in this case, actually the correct one:

$$\big(\tilde{U}(a)\psi\big)(q) = e^{-a/2}\psi(e^{-a}q)\,. \tag{6.35}$$

Moreover, $\tilde{U}(a)$ is a unitary operator on $\mathcal{H}^+$ for each $a \in \mathbb{R}$, given:

$$\begin{aligned}\langle\tilde{U}(a)\psi|\tilde{U}(a)\varphi\rangle &= \int_0^\infty \big(\tilde{U}(a)\psi\big)^*(q)\big(\tilde{U}(a)\varphi\big)(q)\,\mathrm{d}q\\ &= \int_0^\infty e^{-a/2}\psi^*(e^{-a}q)e^{-a/2}\varphi(e^{-a}q)\,\mathrm{d}q\\ &= \int_0^\infty e^{-a}\psi^*(z)\varphi(z)e^{a}\,\mathrm{d}z \qquad (z = e^{-a}q)\\ &= \int_0^\infty \psi^*(q)\varphi(q)\,\mathrm{d}q = \langle\psi|\varphi\rangle\,.\end{aligned}$$

Due to Stone's theorem we can hence conclude that there exists a dense domain on which the generator $\hat{\pi}$ is self-adjoint (this is shown in the proof of Stone's theorem given in [RS.I, theorem VIII.8]; for the domain see also [RS.I, theorem VIII.7]).

Looking at the explicit representation, we can see directly that the operators $\hat{q}$ and $\hat{\pi}$ on $\mathcal{H}^+$ obtained via this method satisfy the same commutation relations as the equally named operators of Isham's representation on $L^2(\mathbb{R}^+, \mathrm{d}q/q)$ given in equation (6.18). They define a self-adjoint representation of the Lie algebra $\mathfrak{L}(\mathbb{R} \rtimes \mathbb{R}^+)$, only this time on the Hilbert space $L^2(\mathbb{R}^+, \mathrm{d}q)$ with the usual Lebesgue measure $\mathrm{d}q$.

Finally, if we define $U(\lambda = e^a) := U(a)$, equivalent to $U(\lambda) = \tilde{U}(\ln\lambda)$ for $\lambda \in \mathbb{R}^+$, we can rewrite equation (6.35) in the form:

$$\big(U(\lambda)\psi\big)(q) = \sqrt{\lambda^{-1}}\,\psi(\lambda^{-1}q)\,. \tag{6.36}$$

In addition, we have:

$$\big(V(b)\psi\big)(q) = e^{-ibq}\psi(q)\,. \tag{6.37}$$

A simple calculation shows that both families of operators together give rise to a unitary irreducible representation of $\mathbb{R} \rtimes \mathbb{R}^+$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$, and due to its construction this is the **correct position space representation** of the Canonical Group $\mathcal{C} = \mathbb{R} \rtimes \mathbb{R}^+$ when the operators $\hat{q}$ and $\hat{\pi}$ are used to describe a particle moving on a half-line.

6.2.3 The Relation to Isham's Result

FINALLY, WE WANT to investigate how the correct position space representation of the Canonical Group $\mathbb{R} \rtimes \mathbb{R}^+$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ for a particle on a half-line is related to Isham's representation (6.16) on $L^2(\mathbb{R}^+, \mathrm{d}q/q)$. There exists, actually, a unitary transformation between the two different Hilbert spaces:

$$W : L^2(\mathbb{R}^+, \mathrm{d}q/q) \to L^2(\mathbb{R}^+, \mathrm{d}q) \,, \quad (W\psi)(q) := \frac{1}{\sqrt{q}}\psi(q) \,. \tag{6.38}$$

The realisation of the one-parameter group of unitary transformations $V(b) = e^{-ibq}$ generated by the operator $\hat{q}$ is not affected by this transformation; since W doesn't alter the argument of the wave functions we have:

$$\big(WV(b)W^{-1}\psi\big)(q) = \frac{1}{\sqrt{q}} e^{-ibq}\sqrt{q}\,\psi(q) = e^{-ibq}\psi(q) \,, \tag{6.39}$$

for all $\psi \in L^2(\mathbb{R}^+, \mathrm{d}q)$. The realisation of the self-adjoint generator $\hat{q}$ is hence diagonal in *both* representations.

Remark 6.2.4 Note that the *domain* of $V(b)$ after the transformation is not the same as before. Wave functions that are square-integrable with respect to the measure $\mathrm{d}q/q$ will in general not be square-integrable with respect to the measure $\mathrm{d}q$, and vice versa. As a consequence, although the generator $\hat{q}$ acts on wave functions in both Hilbert spaces by the exact same prescription $(\hat{q}\psi)(q) = q\psi(q)$, the eigenfunctions of $\hat{q}$ to the *same* eigenvalues will be *different* for $L^2(\mathbb{R}^+, \mathrm{d}q/q)$ and $L^2(\mathbb{R}^+, \mathrm{d}q)$. The reason is that a unitary transformation, in particular, affects the way how *boundary conditions* are realised.

The realisation (6.16) of the one-parameter unitary group $U(\lambda)$ on $L^2(\mathbb{R}^+, \mathrm{d}q/q)$, on the other hand, transforms according to:

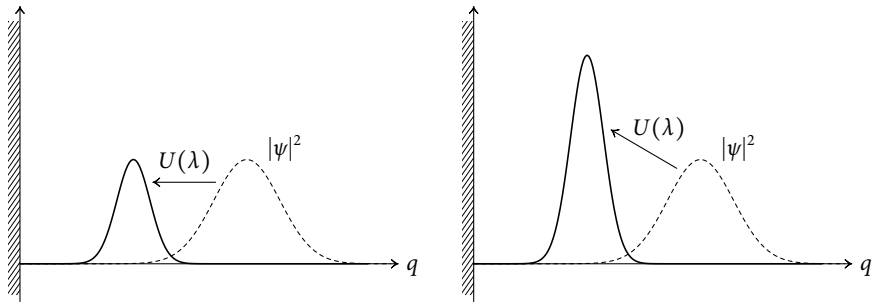
$$\big(WU(\lambda)W^{-1}\psi\big)(q) = \frac{1}{\sqrt{q}}\sqrt{\lambda^{-1}q}\,\psi(\lambda^{-1}q) = \sqrt{\lambda^{-1}}\,\psi(\lambda^{-1}q) \,, \tag{6.40}$$

for all $\psi \in L^2(\mathbb{R}^+, \mathrm{d}q)$. The right-hand side of this formula is just the realisation (6.36) of $U(\lambda)$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ that we obtained above. Moreover, the realisation of the generator $\hat{\pi} = -iq\partial_q$ on $L^2(\mathbb{R}^+, \mathrm{d}q/q)$ is transformed to $\hat{\pi} = -i(q\partial_q + 1/2)$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$, which matches equation (6.33).

> Given these results, the two different representations of $\mathbb{R} \rtimes \mathbb{R}^+$ – the one by Isham and the representation constructed above – turn out to be **unitarily equivalent**. Even though Isham's representation is not the correct position space representation for a particle on a half-line, it hence produces the same physical predictions.

The fact that different notions of distances in the q-coordinate are equivalent is probably a consequence of the canonically conjugate momentum p not being quantizable. When there is no observable that generates translations in the q-coordinate, the measure isn't required to respect these translations.

The mathematical reason why such a unitary equivalence exists is that a measure is not required to be strictly invariant under the group action for a unitary representation to be possible; in general, a *quasi-invariant* measure μ is completely sufficient (Isham actually discusses the point of quasi-invariance in section 5.2 of [Ish83], but he doesn't seem to recognise its relevance to the present case; as a matter of fact, the interpretation of the quantum theory as describing a particle moving on a half-line would break down if the Lebesgue measure $\mathrm{d}q$ on $\mathbb{R}^+$ were not quasi-invariant!).



(a) $L^2(\mathbb{R}^+, \mathrm{d}q/q)$ (b) $L^2(\mathbb{R}^+, \mathrm{d}q)$

Figure 6.4 Unitary representations of the transformation $U(\lambda)$ for the different measures on $\mathbb{R}^+$. The operator $U(\lambda)$ actually preserves the area under the wave packet in both figures. The reason is the unusual measure $\mathrm{d}q/q$ used in (a).

Measures with the same sets of measure zero are usually called *equivalent* measures.

In general, a measure μ is called G-***quasi-invariant*** if the transformed measure μ_g for all $g \in G$ has the same sets of measure zero as the original measure μ. The measure μ_g is thereby defined as $\mu_g(B) := (\rho_{g*}\mu)(B) := \mu(\rho_g^{-1}(B))$ for all Borel sets $B \subseteq Q$, where ρ_g is the group action of G on Q [Ish83, equation (5.2.5)].

Given a G-quasi-invariant measure μ on Q, the space $L^2(Q, \mathrm{d}\mu)$ carries a unitary representation of the group G:

$$\big(U(g)\psi\big)(q) := \left(\frac{\mathrm{d}\mu_g}{\mathrm{d}\mu}(q)\right)^{1/2} \psi(\rho_{g^{-1}}q)\,, \tag{6.41}$$

where the factor $(\mathrm{d}\mu_g/\mathrm{d}\mu)(q)$ is the so-called ***Radon–Nikodým derivative*** (a detailed treatment can be found, for example, in [Els11, section VII.2.3]; be aware that $\mathrm{d}\mu_g/\mathrm{d}\mu$ is not an actual derivative, yet it formally fulfils most of the properties).

Concerning the particle on a half-line, we can see that $U(\lambda)$ in equation (6.36) is of the specific form (6.41). The subgroup $G = (\mathbb{R}^+, \cdot)$ in this case acts on $Q = \mathbb{R}^+$ by simple multiplication, and the Lebesgue measure $\mathrm{d}q$ turns out to be G-quasi-invariant under this action on $\mathbb{R}^+$. In particular, the Radon–Nikodým derivative with respect to the group action is given by $(\mathrm{d}\mu_g/\mathrm{d}\mu)(q) = \lambda^{-1}$, which results in the additional factor of $\lambda^{-1/2}$ compared to Isham's representation.

In the present case, figure 6.4 shows what is going on. We know that any unitary operator $U(\lambda)$ must preserve probabilities, that is, the area under the probability density function $|\psi|^2$ must be preserved. Isham's representation, shown in figure 6.4(a), is only one way to accomplish this. If only the argument of the wave function is transformed, preservation of the area requires a measure on $\mathbb{R}^+$ which takes the transformation into account. For Isham's representation the measure $\mathrm{d}q/q$ is hence strictly necessary.

Nonetheless, we have shown in this section that it *is* possible to keep the Lebesgue measure $\mathrm{d}q$ on $\mathbb{R}^+$ if we rescale the wave functions afterwards, as shown in figure 6.4(b). It turned out that the correct factor for the rescaling to make $U(\lambda)$ unitary is given by the Radon–Nikodým derivative. In particular, if we require the usual notion of distances in the q-coordinate based on the Euclidean metric, the figure shows that the additional factor corresponding to the Radon–Nikodým derivative arises naturally from a simple geo*metric*al consideration.

Remark 6.2.5 After writing section 6.2, I found that the possibility to 'get rid of the non-trivial measure' by means of the unitary transformation (6.38) has been mentioned by

Bojowald and Strobl in [BS00, section 2.2.1]. While no complete proof is provided, they also remark that the operator $\hat{\pi}$ in equation (6.32) is self-adjoint on $L^2(\mathbb{R}^+, \mathrm{d}q)$. The way we obtain the representation of $\mathbb{R} \rtimes \mathbb{R}^+$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ in section 6.2.2 in fact can be seen as an example of the 'projection method' of quantization they describe in [BS00, section 2.5]. Nevertheless, the 'projection method' does not provide any short cuts to our calculations as it does not solve the question of self-adjointness. For every operator restricted to a subspace, self-adjointness has to be checked 'the hard way', as we have done in section 6.2.2. Unfortunately, their method also offers no systematic way how to obtain 'restrictable' operators.

That said, it seems the authors of [BS00] view the nontrivial measure $\mathrm{d}q/q$ just as a mathematical curiosity without any physical significance. Contrary to this, we view the measure on $\mathbb{R}^+$ as crucial to the physical interpretation. As has been discussed in section 6.2.1, a metric that must be compatible with the given measure decides whether the same transformation $q \mapsto \lambda q$ has to be interpreted as a 'scaling' or as a 'translation'. The existence of the two possibilities is particularly relevant to the present situation for which the canonical transformation between $T^*\mathbb{R}$ and $T^*\mathbb{R}^+$ can also be interpreted as a simple change of coordinates, a point that hasn't been mentioned before. Seen this way, the quantum theory over $\mathbb{R}^+$ endowed with the measure $\mathrm{d}q/q$ describes a particle on the *full line* (with unusual coordinates) but q is *not* a position coordinate. Only for a representation on $L^2(\mathbb{R}^+, \mathrm{d}q)$ with the measure inherited from $L^2(\mathbb{R}, \mathrm{d}q)$ it is possible to interpret the result as position space representation of the half-line and q as position of a particle as only in this case the Hilbert space is a subspace $L^2(\mathbb{R}^+, \mathrm{d}q) \subset L^2(\mathbb{R}, \ dq)$, that is, actually 'one half' of a full line.

6.3 Dynamics and the Topological Realisation of Boundary Conditions

So far, the previous section gives the impression that quantization on the phase space $T^*\mathbb{R}^+$ via the Canonical Group $\mathbb{R} \rtimes \mathbb{R}^+$ yields a quantum theory that might, after all, be used to describe a particle moving on a half-line. Nevertheless, we will see in a second that exactly the word 'moving' in the last sentence poses a serious problem. An important dynamical aspect of the quantum particle on a half-line cannot be explained using the obtained quantum theory: the reflection at the end. The quantum theory constructed in section 6.1 is in this sense incomplete. To solve this shortcoming, we will construct a different classical phase space for a particle on a half-line and we will show that this phase space emerges naturally on closer examination of the problem at hand. Our proposal is to implement the 'hard wall' (Dirichlet) boundary conditions as a *topological* feature of the classical phase space.

6.3.1 Dynamics and a 'Free' Particle on the Half-Line

When we think about dynamics, a typical Hamiltonian of a classical particle is given by the sum of kinetic energy and some potential:

$$H(q,p) = \frac{p^2}{2m} + V(q)\ . \tag{6.42}$$

We run straight into a problem: For a particle restricted to a half-line the momentum cannot be quantized. The Hamiltonian vector field X_p on $T^*\mathbb{R}^+$ is incomplete and hence we cannot associate a self-adjoint operator to p. Self-adjoint operators always generate *groups* of unitary transformations but flows of incomplete vector fields give rise only to

*pseudo*groups. Still, since there is no way to make $\hat{p} = -i\partial_q$ a self-adjoint operator on $L^2(\mathbb{R}^+, \mathrm{d}q)$ this is no fault of the quantization method.

At first, the same seems to be the case for the squared momentum p^2. Just like the vector field X_p, the Hamiltonian vector field X_{p^2} is incomplete when restricted to the submanifold $T^*\mathbb{R}^+ \subset T^*\mathbb{R}$ (see figure 6.5). Therefore, the quantization method cannot possibly map the observable p^2 defined on $T^*\mathbb{R}^+$ to a self-adjoint operator.

A rather unexpected thing, however, happens when we consider the same problem from the quantum-mechanical perspective and try to restrict the operator $\hat{p}^2 = -(\partial_q)^2$ known from conventional quantum mechanics on $\mathbb{R}$ to the half-line $\mathbb{R}^+ \subset \mathbb{R}$. One finds that the operator $\hat{p}^2$ indeed *admits* self-adjoint extensions after it has been restricted to the Hilbert space $L^2(\mathbb{R}^+, \mathrm{d}q)$ [RS.II; BFV01; FCT02; GK04; BW10]. This is a key point of this section:

> The Hamiltonian (6.42) without any potential is a valid Hamiltonian for a quantum particle moving on a half-line!

We could, of course, try to ignore this mathematically rather technical result as being unphysical. Nevertheless, we will see that exactly the Hamiltonian without any potential describes reflection of an otherwise free quantum particle at the origin and this surely isn't entirely unphysical.

6.3.2 Self-Adjoint Extensions of $\hat{p}^2$ on a Half-Line

A DETAILED treatment of the squared momentum operator $\hat{p}^2$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ can be found in Reed and Simon's book [RS.II, section X.1, example 2]. We will give an account of the points that are important to our ongoing discussion.

The starting point is the operator $\hat{p}^2 = -(\partial_q)^2$ with domain $D_0 = C_0^\infty(0;\infty)$ of smooth functions with compact support contained in $\mathbb{R}^+ = (0;\infty)$. This operator is symmetric and closed on D_0. Also, the domain D_0 is dense in $L^2(\mathbb{R}^+, \mathrm{d}q)$. To avoid some clumsy notations, we write $A = \hat{p}^2$ in the following.

The next step is to calculate the deficiency indices $n_\pm$ of A. We try to find solutions to the eigenvalue equation $A^\dagger \varphi_\pm = \mp i\varphi_\pm$, where $A^\dagger$ is the adjoint operator. A priori,

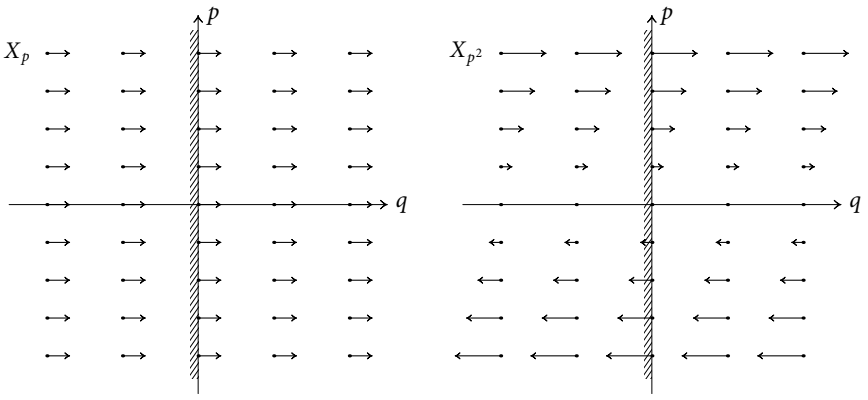


Figure 6.5 The Hamiltonian vector fields X_p and X_{p^2} are both *incomplete* when restricted to $T^*\mathbb{R}^+$. The flow curves leave $T^*\mathbb{R}^+$ at some point and cross into the forbidden region $q \leq 0$.

given the domain of A, the eigenfunctions $\varphi_\pm$ could be generalised solutions (evaluated on test functions in $C_0^\infty(0;\infty)$). Nevertheless, there exists a regularity theorem in this particular case [RS.II, theorem IX.25 (Weyl's lemma)] which states that only ordinary solutions with $\varphi_\pm \in C^\infty(0,\infty)$ are possible.

A basis for the ordinary solutions of $-\partial_q^2\varphi_+(q) = -i\varphi_+(q)$ is given by:

$$\varphi_{+,1}(q) = \exp\big((1+i)q/\sqrt{2}\big)\,, \quad \varphi_{+,2}(q) = \exp\big(-(1+i)q/\sqrt{2}\big)\,,$$

for $-\partial_q^2\varphi_-(q) = +i\varphi_-(q)$ we have:

$$\varphi_{-,1}(q) = \exp\big((1-i)q/\sqrt{2}\big)\,, \quad \varphi_{-,2}(q) = \exp\big(-(1-i)q/\sqrt{2}\big)\,.$$

Of these solutions, only $\varphi_{+,2}$ and $\varphi_{-,2}$ are in $L^2(\mathbb{R}^+,\mathrm{d}q)$ whereas $\varphi_{-,1}$ and $\varphi_{+,1}$ diverge for $q\to\infty$. The deficiency subspaces $\mathfrak{r}_\pm$ are hence both 1-dimensional, so that we obtain the deficiency indices $(n_+,n_-) = (1,1)$. As a consequence, the operator A admits self-adjoint extensions.

The specific extension we are interested in is the one that corresponds to a reflection by an infinitely high hard wall potential at the origin, because this is the case which we expect to have a well-defined classical counterpart (see also [FCT02]). The domain of this extension is:

$$D_\infty = \big\{\psi\in L^2(\mathbb{R}^+,\mathrm{d}q) : \psi\in AC^2[0,\infty]\,,\ \psi(0)=0\big\}\,. \tag{6.43}$$

It is important to notice that the boundary condition $\psi(0) = 0$ is **not optional** in the definition of D_∞. Unlike in conventional quantum mechanics over $\mathbb{R}$, $\hat{p}^2$ will not be self-adjoint if the condition is omitted.

The space $AC^2[0;\infty]$ is the set of square-integrable functions in $L^2([0;\infty],\mathrm{d}q)$ whose weak derivatives are in $AC[0;\infty]$. The space $AC[0;\infty]$ denotes absolutely continuous functions over $[0;\infty]$.

Absolutely continuous functions over X are differentiable μ-almost everywhere and their derivatives are in $L^2(X,\mu)$ (see [Gro88, section 3.2] or [RS.I, section VIII.1] for a definition).

Remark 6.3.1 (Boundary conditions) The domain D_∞ poses conditions for functions over the *closed* interval $[0;\infty]$ although the boundary points 0 and ∞ are *not* contained in $\mathbb{R}^+$. This doesn't seem to make sense at first, because $\psi\in L^2(\mathbb{R}^+,\mathrm{d}q)$ isn't even defined at $q=0$, hence it should be impossible to require $\psi(0)=0$. However, the condition in equation (6.43) must actually be read as 'ψ is an equivalence class in $L^2(\mathbb{R}^+,\mathrm{d}q)$ that admits a representative which can be continued to a function in $AC^2[0;\infty]$ that satisfies $\psi(0)=0$.' As a result, it is possible for wave functions ψ in the domain D_∞ to kind of 'feel' a potential like the one in equation (6.30) which is infinitely high for $q\le 0$, and this *although the potential vanishes identically on* $\mathbb{R}^+$!

The necessity of the boundary condition is also interesting in the light of our earlier remark 6.2.2. There, we pointed out that functions in the Hilbert space $L^2(\mathbb{R}^+,\mathrm{d}q)$ don't have to be continuous, just measurable. The boundary condition $\psi(0)=0$ alone isn't a restriction for functions in $L^2(\mathbb{R}^+,\mathrm{d}q)$ and equation (6.31) defines an isomorphism between Hilbert spaces. Nevertheless, we see now that the restriction to the subspace of wave functions $L^2(\mathbb{R}^+,\mathrm{d}q)\subset L^2(\mathbb{R},\mathrm{d}q)$ compatible with the hard wall potential V doesn't 'fully implement' the effects of the potential. In fact, it is in general not enough to consider the effects on wave functions in Hilbert space, but it is necessary to have in mind the effect of restrictions on the *domains of operators*.

Regarding the physical interpretation of the operator extension to D_∞, it is clear that a plane wave $\psi(q) = e^{ipq}$ doesn't satisfy the boundary condition $\psi(0)=0$. There is, however, a linear combination of plane waves that works, namely:

$$\psi(q) = e^{ipq} - e^{-ipq}\,. \tag{6.44}$$

Although p appears in this formula as a variable, keep in mind that the momentum on the half-line isn't observable (see the discussion below).

Remark 6.3.2 (Eigenfunctions and spectrum of $\hat{p}^2$) The functions ψ in equation (6.44) are the (generalised) eigenfunctions of $\hat{p}^2$. While they are not square-integrable, they satisfy:

$$\hat{p}^2\psi(q) = -\partial_q^2\big(e^{ipq} - e^{-ipq}\big) = p^2\big(e^{ipq} - e^{-ipq}\big) = p^2\psi(q)\,. \tag{6.45}$$

Consequently, the (purely continuous) spectrum of $\hat{p}^2$ on D_∞ is given by:

$$\sigma(\hat{p}^2) = (0;\infty)\,. \tag{6.46}$$

From a physical perspective, the eigenfunction ψ is effectively the superposition of an incoming plane wave with negative momentum $-p$ and a reflected, outgoing plane wave with positive momentum p. Hence we see that the self-adjoint operator $\hat{p}^2$ with domain D_∞ generates the dynamics of a quantum particle that gets reflected by an infinitely high hard wall potential at the origin, just as claimed.

That being said, with respect to the quantization scheme it is important to note that neither the incoming, nor the outgoing wave alone is compatible with the boundary condition for $\hat{p}^2$ on D_∞. Accordingly, although ψ *looks* like a superposition of plane waves, it *cannot actually be decomposed* in this way. We know that the momentum operator restricted to the half-line isn't a valid observable, and, indeed, the eigenfunction ψ in equation (6.44) doesn't depend on the sign of p. Although the wave function ψ contains information about the *magnitude* $|p|$ of the momentum, it doesn't select a direction. This agrees perfectly with the fact that only the kinetic energy, proportional to p^2, is preserved by an elastic reflection at the origin, but not the momentum p, which changes its sign.

6.3.3 A Limit of the Quantization Procedure

We have seen that the squared momentum operator $\hat{p}^2$ on an appropriately chosen domain is a self-adjoint operator on $L^2(\mathbb{R}^+, \mathrm{d}q)$. Moreover, when used as Hamiltonian (we will drop the constant $2m$ in the following) the operator $\hat{p}^2$ on D_∞ generates some perfectly reasonable dynamics for a quantum particle moving on a half-line. In fact, an elastic reflection at the origin is not just a reasonable but a *characteristic* trait if we want the theory for a particle on the half-line to be equivalent to a particle restricted to $\mathbb{R}^+$ as if by the infinitely high step potential V given in equation (6.30).

The quantization method, on the other hand, when applied to the phase space $T^*\mathbb{R}^+$ cannot justify the self-adjoint operator $\hat{p}^2$ at all. We know that the Hamiltonian vector field X_{p^2} on $T^*\mathbb{R}^+$, just like X_p, is incomplete (figure 6.5), and since p^2 doesn't generate a complete group of transformations on phase space the quantization procedure cannot assign a self-adjoint operator to p^2. In other words:

> There is no Canonical Group $\mathcal{C}$ for which the squared momentum p^2 on the classical phase space $T^*\mathbb{R}^+$ is a quantizable observable!

The fact that the quantization procedure is unable to reproduce such an important and characteristic aspect of the dynamics of a quantum particle moving on a half-line is a severe shortcoming. It must be either a fault of the quantization method, or a limit of the chosen classical phase space $T^*\mathbb{R}^+$. While we will argue that a different classical phase space can be used to obtain the self-adjoint operator $\hat{p}^2$ over $\mathbb{R}^+$ (and we will explain why this phase space might be a better choice overall with respect to the characteristic properties of a half-line), *there are strict limits as to what can be accomplished*:

Lemma 6.3.3 There exists no classical phase space M and no Canonical Group $\mathcal{C}$ for which the quantization procedure maps classical fundamental observables q and p^2 to the self-adjoint quantum operators $\hat{q}$ and $\hat{p}^2$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$, respectively.

Proof. The commutator of $\hat{q}$ and $\hat{p}^2$ is given by $[\hat{q}, \hat{p}^2] = 2i\hbar\hat{p}$, so that the Poisson bracket of the corresponding classical observables q and p^2 must be $\{q, p^2\} = 2p$. Thus, p as well becomes an element of the algebra of fundamental observables whenever both q and p^2 are in $\mathrm{Obs}(M)$.

On the other hand, it is a requirement of the quantization procedure that classical observables in $\mathrm{Obs}(M)$ are mapped to (essentially) self-adjoint operators. Nevertheless, such a map cannot exist in the present situation because there is no way to make $\hat{p}$ a self-adjoint operator on $L^2(\mathbb{R}^+, \mathrm{d}q)$. ∎

Accordingly, as long as we require the full set of all quantizable observables to close under the Poisson bracket, not just the fundamental ones (observe that this is important to guarantee a consistent quantization map), lemma 6.3.3 implies that *there is no way to obtain $\hat{q}$ and $\hat{p}^2$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ at the same time*. So, if we want the quantization method to reproduce the squared momentum operator $\hat{p}^2$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ it becomes clear that we will have to give up the position operator $\hat{q}$.

Remark 6.3.4 It is important to consider the implications of lemma 6.3.3 carefully. In general, only the *full* Hamiltonian $\hat{H}$ needs to be self-adjoint, not individual terms. It is, for example, sometimes possible to counter-balance non-matching deficiency indices of one operator by adding or multiplying with another (this happens for the operator $\hat{p}^2$ for which $\hat{p}$ on $C_0^\infty(0;\infty)$ has deficiency indices $(n_+, n_-) = (0, 1)$, yet $\hat{p}^2$ admits self-adjoint extensions; this is also the reason why the eigenfunction (6.44) cannot be decomposed over the half-line). In a more general situation, it could therefore be that only a Hamiltonian of the form (6.42) with some *non-vanishing* potential $V(q)$ generates physically acceptable dynamics. In the specific case at hand, however, we showed that $\hat{p}^2$ is already self-adjoint *by itself* on the domain D_∞, so no additional terms are needed. Moreover, we know that $\hat{p}^2$ generates physically relevant dynamics.

6.3.4 A Different Phase Space for a Particle on a Half-Line

Giving up the position operator $\hat{q}$ in favour of $\hat{p}^2$ isn't easy to accept. Nevertheless, we have seen numerous times that it can be futile to insist on prejudices, especially with regard to quantum mechanics. In fact, we will show that a different classical phase space together with a suitable Canonical Group captures the characteristics of a half-line in a much better way. Moreover, when searching for fundamental observables compatible with p^2 on this phase space, we find q^2 as a natural candidate. Although q^2 in general fixes the position coordinate q only up to a sign, this is not an issue for $q \in \mathbb{R}^+$, for which the observable q and q^2 contain the same information.

Let us begin with the classical phase space. The problem with the original phase space $T^*\mathbb{R}^+$ is that the Hamiltonian vector field X_{p^2} on $T^*\mathbb{R}^+$ is incomplete. On the other hand, this is nothing but expected. We know that the Hamiltonian $H = p^2/(2m)$ in classical mechanics generates the motion of a free particle. So, if we just 'look' at the half-line $\mathbb{R}^+ \subset \mathbb{R}$ but *don't impose any boundary condition* it is no wonder that a free particle will be able to leave this subset and cross into the 'forbidden' area.

When we now consider the vector field X_{p^2} in figure 6.5 again, with this in mind, an interesting possibility opens up. We can try to implement the boundary condition that describes elastic reflection of an otherwise free particle at the origin by gluing together the phase space as in figure 6.6. If we identify $(0, -p)$ with $(0, p)$ for all momenta p, the flow curve of the Hamiltonian vector field X_{p^2} on which an incoming free particle with

The flow curves of X_H are the solutions of the canonical equations with Hamiltonian H.

negative momentum $-p$ arrives (and would otherwise leave the submanifold $T^*\mathbb{R}^+$) gets connected to the flow curve of an outgoing particle with momentum $+p$.

While this is basically the correct idea, there are some technical details that require our attention:

- We have to take care that the 'gluing together' happens smoothly. We must be able to calculate the Hamiltonian vector field X_{p^2} on the resulting phase space and, in particular, the symplectic form ω needs to survive the process.
- It turns out that it is impossible to obtain globally a smooth manifold. We will see that the space after the identification looks like a half-cone (think 'forward light cone'), which has a singularity at the tip, corresponding to the point $(q, p) = (0, 0)$. Our options are either to exclude this point in order to get a smooth manifold, or to think about a suitable generalisation of the quantization method.

If we ignore the issue with the origin for the moment, the solution to the first problem is straightforward. As shown in figure 6.7, we just have to make sure that for each point $(0, p)$ a neighbourhood U contained in a small strip around $q = 0$ gets identified *in the correct direction* with a neighbourhood U' of $(0, -p)$. We can see that this requires the neighbourhoods U and U' to be related by a point reflection at the origin.

That being said, this strip around $q = 0$ doesn't actually have to be 'small'. Without loss of generality we can, in fact, extend this strip to the whole space $T^*\mathbb{R}$, as shown in figure 6.8, if we identify all points in $T^*\mathbb{R} \cong \mathbb{R}^2$ related by a point reflection (a group action of $\mathbb{Z}_2$ on $T^*\mathbb{R}$):

Notice that the reflection (6.47) restricts to the identification of $(0, p)$ with $(0, -p)$ for $q = 0$.

$$\mathbb{R}^2 \to \mathbb{R}^2 , \quad (q, p) \mapsto (-q, -p) . \tag{6.47}$$

Each equivalence class of points then has a representative (q, p) with $q \geq 0$ and the resulting phase space $\mathbb{R}^2/\mathbb{Z}_2$ has the mathematical structure of an *orbifold*. Furthermore, we can see that it is possible to keep the symplectic 2-form $\omega = \mathrm{d}q \wedge \mathrm{d}p$ of $T^*\mathbb{R} \cong \mathbb{R}^2$ on the orbifold $\mathbb{R}^2/\mathbb{Z}_2$ since $\mathrm{d}q \wedge \mathrm{d}p$ is invariant under the reflection (6.47):

An ***orbifold*** M/Γ is basically the quotient of a manifold M by a finite group Γ. The actual definition enforces the quotient structure only locally, of course. Moerdijk [MM03, section 2.4] and Hepworth [Hep09, section 3] give proper definitions, but we do not need the details in the following.

$$\mathrm{d}q \wedge \mathrm{d}p \mapsto \mathrm{d}(-q) \wedge \mathrm{d}(-p) = \mathrm{d}q \wedge \mathrm{d}p .$$

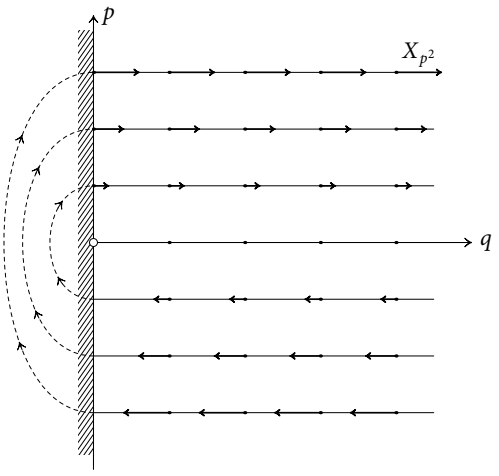


Figure 6.6 It is possible to make the Hamiltonian vector field X_{p^2} complete by gluing together the phase space so that the flow curves of X_{p^2} that would leave the submanifold $T^*\mathbb{R}^+$ at the point $(0, -p)$ enter again at $(0, p)$.

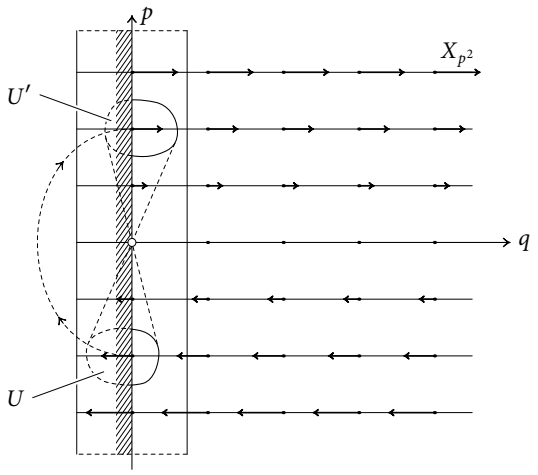


Figure 6.7 Mathematically, to ensure that the identification happens smoothly, we have to include a small strip around $q = 0$ where a neighbourhood U is related to U' by the point reflection (6.47).

Remark 6.3.5 The phase space $\mathbb{R}^2/\mathbb{Z}_2$ has an interesting physical interpretation. It is possible to think of the original particle at the position q with momentum p as being accompanied by a 'mirror particle' at the point $-q$ with the opposite momentum $-p$. When the original particle with (negative) momentum p leaves the half-line, the mirror particle enters with the opposite momentum $-p$, and in terms of representatives with $q \geq 0$ this looks exactly like a reflection at the point $q = 0$. Doncheski and Robinett carried out a numerical study of wave packets reflecting at the origin using an informal version of this 'mirror particle' model in [DR99].

Concerning observables on this space, notice that the Hamiltonian vector field X_q associated to the position observable q on $T^*\mathbb{R}$ doesn't survive the identification process because tangent vectors with opposite direction would have to be identified. In fact, we see that the observable *q itself* is no longer well-defined on $\mathbb{R}^2/\mathbb{Z}_2$.

Due to lemma 6.3.3, we can not have q as observable anyway, if we want p^2 to be quantizable.

In general, only smooth functions on $T^*\mathbb{R}$ invariant under the reflection (6.47) are still well-defined on the resulting phase space and may in principle be quantizable. Valid observables are hence functions $f \in C^\infty(T^*\mathbb{R}, \mathbb{R})$ with the symmetry property:

$$f(q,p) = f(-q,-p)\,. \tag{6.48}$$

This is the case, for example, for polynomials of *even* degree in q and p (including mixed polynomials like qp), but rules out all odd polynomials.

So, what remains is to tackle the problem with the singularity. To this end, observe that all the Hamiltonian vector fields of valid observables on $\mathbb{R}^2/\mathbb{Z}_2$ vanish at this point:

Lemma 6.3.6 (Hamiltonian vector fields on $\mathbb{R}^2/\mathbb{Z}_2$) Let f be a smooth function in $C^\infty(\mathbb{R}^2, \mathbb{R})$ which satisfies the symmetry property (6.48). Then, the Hamiltonian vector field X_f vanishes at the point $(q,p) = (0,0)$.

Proof. First, observe that the derivative of a differentiable function g which satisfies the

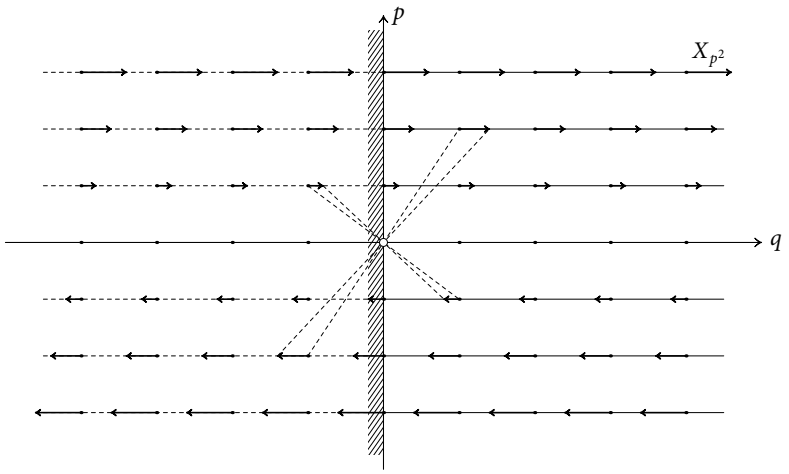


Figure 6.8 An alternative construction for the phase space in figure 6.6 is to start with $T^*\mathbb{R} \cong \mathbb{R}^2$ and identify (q, p) with $(-q, -p)$. Again, it is easy to see that the Hamiltonian vector field X_{p^2} is complete on this space.

symmetry property $g(x) = g(-x)$ vanishes at the point $x = 0$, due to:

$$g'(0) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon}\Big(g(0+\varepsilon) - g(0-\varepsilon)\Big) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon}\Big(g(\varepsilon) - g(\varepsilon)\Big) = 0\,.$$

When we apply this result to the mappings $q \mapsto f(q, 0)$ and $p \mapsto f(0, p)$, respectively, it follows that all functions which satisfy the property (6.48) will have vanishing partial derivatives $\partial_q f$ and $\partial_p f$ at the point $(q, p) = (0, 0)$. Therefore, given the definition of X_f as $X_f = \mathrm{d}f^\sharp = (\partial_q f)\,\mathrm{d}q^\sharp + (\partial_p f)\,\mathrm{d}p^\sharp$, we see that the Hamiltonian vector field X_f vanishes at the point $(0, 0)$. ■

Since all Hamiltonian vector fields vanish at the singularity, the stabiliser subgroup $G_{(0,0)}$ of this point for any group action of a Lie group G that generates Hamiltonian vector fields is the whole group $G_{(0,0)} = G$. Hence, according to the orbit–stabiliser theorem 3.2.2, the G-orbit of $(0, 0)$ is given by $G(0, 0) \cong G/G_{(0,0)} \cong \{0\}$. Consequently, the singularity *will always be an orbit by itself*.

Given this result, it doesn't really matter for the quantization procedure whether we keep or exclude the singularity:

- If we choose to *exclude* the singularity, we get the manifold $(\mathbb{R}^2 \setminus \{0\})/\mathbb{Z}_2$ and the quantization method can be applied exactly as before. Moreover, since the singularity is a separate orbit and a group action on a single point can only be trivial, excluding this point doesn't affect the choice of Geometric Group.
- On the other hand, if we consider the orbifold $\mathbb{R}^2/\mathbb{Z}_2$ *including* the singularity, we cannot require the action of the Geometric Group $\mathcal{G}$ to be transitive on the whole orbifold. The best we can do is to require the $\mathcal{G}$-action to be transitive on the subset $(\mathbb{R}^2/\mathbb{Z}_2) \setminus \{0\}$. The quantization method simply *avoids the singularity* and doesn't really 'feel' the orbifold structure.

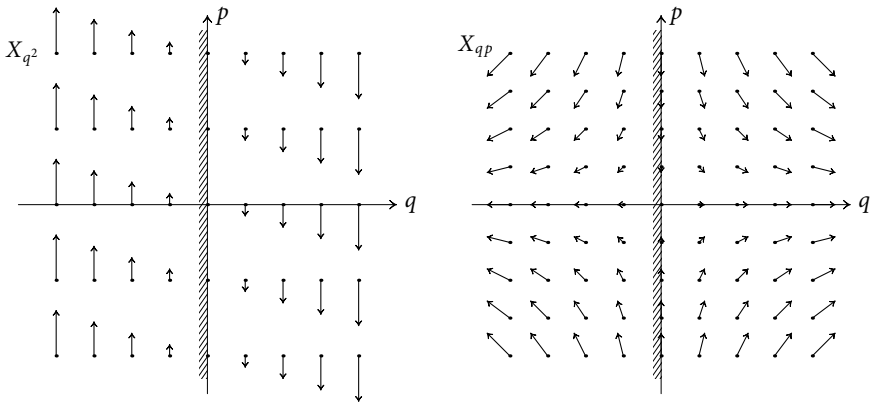


Figure 6.9 The Hamiltonian vector fields X_{q^2} and X_{qp} are compatible with the reflection (6.47) and hence well-defined on the orbifold $\mathbb{R}^2/\mathbb{Z}_2$. In combination with X_{p^2} they generate a transitive group action on $\mathbb{R}^2 \setminus \{0\}$.

6.4 Quantization on $\mathbb{R}^2/\mathbb{Z}_2$ and 'Half a Harmonic Oscillator'

After constructing a phase space that describes the elastic reflection at the end of a half-line classically, we try to find a Canonical Group that reproduces the squared momentum operator $\hat{p}^2$ (with domain D_∞) in the quantum theory on $L^2(\mathbb{R}^+, \mathrm{d}q)$. We will, however, apply the quantization method a little differently than usual. Instead of searching for a Geometric Group $\mathcal{G}$, we will directly construct a suitable Lie algebra of fundamental observables $\mathrm{Obs}(M)$ that contains p^2. We will then use the quantization method to 'reverse-engineer' the Canonical Group on $\mathbb{R}^2/\mathbb{Z}_2$ that maps p^2 to $\hat{p}^2$.

6.4.1 An Algebra of Fundamental Quantizable Observables

As mentioned before, the symmetry condition (6.48) for observables on $\mathbb{R}^2/\mathbb{Z}_2$ rules out all odd polynomials in q and p. The most natural supplement to p^2 hence seems to be q^2. Nevertheless, when we calculate the Poisson bracket of q^2 and p^2 we find that the mixed polynomial qp must be included as well. After adding qp, tough, the Lie algebra $\mathrm{Obs}(M)$ generated by the fundamental observables $\{q^2, p^2, qp\}$ closes, and the Poisson brackets are:

$$\{q^2, p^2\} = 4qp\,, \quad \{qp, q^2\} = -2q^2\,, \quad \{qp, p^2\} = +2p^2\,. \tag{6.49}$$

The Hamiltonian vector fields X_{q^2}, X_{p^2} and X_{qp} are well-defined on $\mathbb{R}^2/\mathbb{Z}_2$, complete, and they span a 2-dimensional tangent space at every point except $(q, p) = (0, 0)$ (see figure 6.9). Therefore, given Palais' theorem 5.1.1, these vector fields generate a transitive group action of some Lie group on $(\mathbb{R}^2/\mathbb{Z}_2) \setminus \{0\}$.

A look at the literature (e. g. Fulton and Harris [FH91, section 10.3]) reveals that the Poisson brackets in equation (6.49) are brackets in the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$:

$$[N_+, N_-] = A\,, \quad [A, N_+] = +2N_+\,, \quad [A, N_-] = -2N_-\,, \tag{6.50}$$

The algebra $\mathfrak{sl}(2, \mathbb{R})$ consists of traceless matrices. This follows from $\det \circ \exp = \exp \circ \mathrm{tr}$ and $\det(g) = 1$ for all $g \in \mathrm{SL}_2\mathbb{R}$.

for which the basis elements A, N_+ and N_- have matrix representations:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad N_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad N_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{6.51}$$

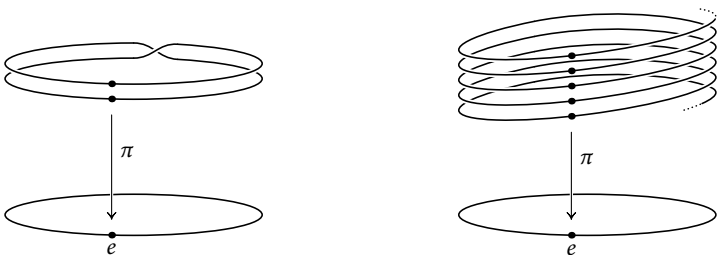


Figure 6.10 Two-fold and universal covering group of $\mathrm{SO}(2) \cong S^1$.

Explicitly, if we identify the Poisson bracket with the Lie bracket in (6.50), $[\,\cdot\,,\,\cdot\,] \hat{=} \{\,\cdot\,,\,\cdot\,\}$, equation (6.50) results from the following identifications:

$$A \hat{=} -qp\,, \quad N_+ \hat{=} \tfrac{1}{2}\alpha q^2\,, \quad N_- \hat{=} -\tfrac{1}{2}\alpha^{-1}p^2\,, \tag{6.52}$$

with $\alpha \in \mathbb{R}^\times$. We will later use $\alpha = 1$, but notice that this is the same freedom as in the choice of physical base units q_0 and p_0 for the Heisenberg group (remark 5.4.1).

Like Fulton and Harris [FH91], we denote the *real symplectic group* of rank n by $\mathrm{Sp}_{2n}\mathbb{R}$. Be aware that some authors prefer to write $\mathrm{Sp}_n\mathbb{R}$ in place of our $\mathrm{Sp}_{2n}\mathbb{R}$ (e. g. Taylor [Tay86]).

A Lie group associated to the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ is the *special linear group* $\mathrm{SL}_2\mathbb{R}$, with isomorphisms $\mathrm{SL}_2\mathbb{R} = \mathrm{Sp}_2\mathbb{R} \cong \mathrm{SU}(1,1) \cong \mathrm{Spin}^+(1,2)$. The centre of $\mathrm{SL}_2\mathbb{R}$ is $\mathcal{Z}(\mathrm{SL}_2\mathbb{R}) = \{\pm\mathbb{1}\} \cong \mathbb{Z}_2$, and the quotient $\mathrm{PSL}_2\mathbb{R} := \mathrm{SL}_2\mathbb{R})/\mathbb{Z}_2$, called the *projective special linear group*, produces the same Lie algebra. $\mathrm{PSL}_2\mathbb{R}$ is isomorphic to the *proper orthochronous Lorentz group* $\mathrm{SO}^\uparrow(1,2) = \mathrm{SO}_e(1,2)$ *in 2+1 (space–time) dimensions*. Of course, all covering groups of $\mathrm{SO}^\uparrow(1,2)$ have $\mathfrak{sl}(2,\mathbb{R})$ as their Lie algebra as well.

The basic theorems concerning covering groups can be found in [Sim96, section VII.6] and in [FH91, section 7.3]. Duistermaat and Kolk [DK00, chapter 1], on the other hand, give a more detailed account.

Remark 6.4.1 (Covering groups) It is well known that the double cover $\mathrm{Spin}(n)$ of the rotation group $\mathrm{SO}(n) = \mathrm{Spin}(n)/\mathbb{Z}_2$ is the *universal covering group* (the unique, simply connected covering group) for $n \geq 3$, because the fundamental group of $\mathrm{SO}(n)$ is equal to $\pi_1(\mathrm{SO}(n)) \cong \mathbb{Z}_2$ for all $n \geq 3$ [FH91, proposition 23.1]. In particular, there are for $n \geq 3$ no additional covering groups of $\mathrm{SO}(n)$ other than the spin group $\mathrm{Spin}(n)$.

For $\mathrm{SO}^\uparrow(1,2)$, however, the double cover $\mathrm{SL}_2\mathbb{R}$ is *not* the universal covering group and the situation gets a lot more interesting. The fundamental group of $\mathrm{SL}_2\mathbb{R}$ is given by $\pi_1(\mathrm{SL}_2\mathbb{R}) \cong \mathbb{Z}$ (the group $\mathrm{SL}_2\mathbb{R}$ contains $\mathrm{SO}(2) \cong S^1$ as a maximal compact subgroup and the fundamental group of the circle is $\pi_1(S^1) \cong \mathbb{Z}$). The universal covering group $\widetilde{G}$ of $G = \mathrm{SL}_2\mathbb{R}$ is thus infinitely sheeted. As a result, there exist n-fold covering groups of $\mathrm{SO}^\uparrow(1,2)$ for every number $n \in \mathbb{N}$. The n-fold cover G_n of a connected Lie group G is a *central group extension* by $\mathbb{Z}_n$ [Sim96, theorem VII.6.2]:

$$0 \longrightarrow \mathbb{Z}_n \hookrightarrow G_n \twoheadrightarrow G \longrightarrow 0\ .$$

Given that the centre of $\mathrm{SO}^\uparrow(1,2)$ is trivial, it follows that the n-fold covering group G_n of $G = \mathrm{SO}^\uparrow(1,2)$ has centre $\mathcal{Z}(G_n) = \mathbb{Z}_n \subseteq \mathbb{Z}$.

The metaplectic group $\mathrm{Mp}_2\mathbb{R}$, in contrast to $\mathrm{SO}^\uparrow(1,2)$ and $\mathrm{Sp}_2\mathbb{R}$, is *not a matrix group*. No proper covering group of $\mathrm{SL}_2\mathbb{R}$ has faithful finite-dimensional representations [Puk64, p. 97].

One specific covering group that will appear later is the *metaplectic group* $\mathrm{Mp}_2\mathbb{R}$, which is the unique 2-fold cover of the symplectic group $\mathrm{Sp}_2\mathbb{R} \cong \mathrm{SL}_2\mathbb{R}$, and is the 4-fold cover of $\mathrm{SO}^\uparrow(1,2)$ (see [Tay86, chapter 11]).

Remark 6.4.2 The group $\mathrm{SO}^\uparrow(1,2)$ and covering groups appear in a paper by Bojowald *et al.* [Boj+00] in an application to Schwarzschild black holes, and also in a number of papers written by Kastrup [Kas03; Kas07] where he discusses the quantization of a harmonic oscillator using action-and-angle variables. Kastrup in fact recognises in these

papers that the phase space $\{(\varphi, I) : \varphi \in \mathbb{R} \bmod 2\pi\,,\ I \in \mathbb{R}^+\}$ of action I and angle φ is diffeomorphic to $\mathbb{R}^2/\mathbb{Z}_2$ (see in particular [Kaso3, appendix A.3]). On the other hand, Kastrup parenthetically remarks in [Kaso3, section 6.3] that 'a system confined to the half plane $q \geq 0$' (realised by an infinite step potential like (6.30)) can be obtained by an identification of q with $-q$, but he explicitly *forbids* the identification of p with $-p$ (he does not justify the construction, however, nor does he give further details). Thus, he ends up with $\{(q, p) \in \mathbb{R}^2 : (q, p) \equiv (-q, p)\}$ as phase space of the half-line.

In contrast to Kastrup, we believe that the additional identification of p with $-p$ is strictly necessary to describe a particle on a half-line for several reasons: First of all, as a purely mathematical requirement the flow curves of X_{p^2} need to be connected in the correct direction to make the vector field complete (see figure 6.6) since otherwise the group-theoretical quantization method cannot be applied. This is not accomplished by Kastrup's model. Furthermore, while the symplectic 2-form $\omega = \mathrm{d}q \wedge \mathrm{d}p$ is well-defined on $\mathbb{R}^2/\mathbb{Z}_2$, we can see that ω does not survive the identification of (q, p) with $(-q, p)$, given that $\mathrm{d}(-q) \wedge \mathrm{d}p = -(\mathrm{d}q \wedge \mathrm{d}p) \neq \mathrm{d}q \wedge \mathrm{d}p$. Apart from that, the identification of p with $-p$ also makes sense from a physical perspective as the momentum in classical mechanics changes its sign when a particle gets reflected by a hard wall.

Kastrup also obtains a ground state energy of $1/2\,\hbar\omega$ for his model [Kaso3, section 6.3], but we will show in section 6.4.4 that the correct ground state energy of a harmonic oscillator on the half-line with 'hard wall' boundary conditions is $3/2\,\hbar\omega$.

6.4.2 Irreducible Unitary Representations of $\mathrm{SL}_2\mathbb{R}$

We know that the Lie groups associated to the algebra $\mathfrak{sl}(2, \mathbb{R})$ are covering groups of $\mathrm{SO}^\uparrow(1, 2)$. What remains is to find the specific covering that should be used as the Canonical Group for a particle on the half-line. However, since the representations of covering groups of $\mathrm{SO}^\uparrow(1, 2)$ are related, it is not much more work to study irreducible representations for all possibilities. Afterwards, it is possible to 'reverse-engineer' the Canonical Group by looking at which representation reproduces the operator $\hat{p}^2$ on the domain D_∞ from section 6.3. We start with representations of $\mathrm{SL}_2\mathbb{R}$.

The classic reference concerning irreducible unitary representations of $\mathrm{SO}^\uparrow(1, 2)$ and its covering groups is Bargmann's seminal paper [Bar47]. Harish-Chandra not long afterwards presented a different, more algebraic approach, using what is today known as the Harish-Chandra homomorphism [HC51; HC52]. Nevertheless, the representation theory for non-compact Lie groups was in a very early stage of development at this time so that some important concepts were not yet properly named. The notation as well often feels quite baroque (especially in Bargmann's paper).

A more accessible introduction for the contemporary reader using modern terms and notation is Lang's book [Lan85]. We will use Taylor [Tay86, chapter 8] as our main reference, who gives an excellent and succinct treatment. The modern formulation of the algebraic approach using the language of irreducible $(\mathfrak{g}, K)$-modules, on the other hand, is explained in Habib [Hab98] and Huang [Hua99, chapter 11].

A Lie algebra representation of $\mathfrak{g}$ on $\mathcal{H}$ turns $\mathcal{H}$ into a left $\mathfrak{g}$-module. A $(\mathfrak{g}, K)$-module is a $\mathfrak{g}$-module together with an action of the maximal compact subgroup $K \subset G$.

From the standpoint of representation theory, $\mathrm{SL}_2\mathbb{R}$ is the simplest non-compact, semisimple Lie group. The major difficulty to deal with is that all non-trivial irreducible unitary representations of non-compact semisimple Lie groups are infinite-dimensional. The crucial feature of a semisimple Lie group G, on the other hand, is that G contains a *maximal compact subgroup* K. This is a consequence of the ***Iwasawa decomposition*** [Iwa49], according to which any semisimple Lie group G can be written uniquely as the product $G = KAN$ of a maximal compact subgroup K, an Abelian subgroup A and a nilpotent subgroup N (see Knapp [Kna96, section VI.4] for a contemporary and readable account on the Iwasawa decomposition).

The Lie algebra $\mathfrak{g}$ as well splits into a direct sum $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}$, in agreement with the Iwasawa decomposition of G.

In our case of $G_1 = \mathrm{SL}_2\mathbb{R}$, the maximal compact subgroup is given by $K_1 = \mathrm{SO}(2)$ [Tay86, chapter 13]. The Iwasawa decomposition $G_1 = K_1 A_1 N_1$ consists of:

$$K_1 = \left\{ k_1(\vartheta) := \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix} , \ \vartheta \in [0; 2\pi) \right\} = \mathrm{SO}(2) , \tag{6.53a}$$

$$A_1 = \left\{ a_1(t) := \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} , \ t \in \mathbb{R} \right\} , \tag{6.53b}$$

$$N_1 = \left\{ n_1(\xi) := \begin{pmatrix} 1 & \xi \\ 0 & 1 \end{pmatrix} , \ \xi \in \mathbb{R} \right\} . \tag{6.53c}$$

Using the matrix representations (6.50) of the Lie algebra elements, the subgroups can be written via the exponential (exp is the usual matrix exponential in this case):

$$k_1(\vartheta) = \exp(\vartheta Z) , \quad a_1(t) = \exp(tA) , \quad n_1(\xi) = \exp(\xi N_+) , \tag{6.54}$$

with the generator Z of the maximal compact subgroup K_1 given by:

$$Z := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = (N_+ - N_-) . \tag{6.55}$$

Remark 6.4.3 (Exponential map) For compact Lie groups the exponential always maps surjectively onto the connected component of the identity. For the non-compact group $\mathrm{SL}_2\mathbb{R}$, however, the exponential $\exp : \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}_2\mathbb{R}$ is *not surjective* – the image of exp is not even dense in $\mathrm{SL}_2\mathbb{R}$ (see [Mos94]). It is not difficult to show that matrices of the form $-n_1(\xi)$ for $\xi \neq 0$ cannot be reached using an argument based on eigenvalues and the fact that $-n_1(\xi)$ for $\xi \neq 0$ cannot be diagonalised.

Nevertheless, using the Iwasawa decomposition, any group element $g \in G_1$ can be written as a product $g = kan$ and each factor taken by itself *can* be reached via the exponential. Indeed, for the problematic matrices of the form $-n_1(\xi)$ we have:

$$-n_1(\xi) = (-\mathbb{1})n_1(\xi) = k_1(\pi)n_1(\xi) = \exp(\pi Z)\exp(\xi N_+) .$$

For $\mathrm{SO}^{\uparrow}(1,2)$, though, the exponential *is* surjective because $-\mathbb{1}$ and $\mathbb{1}$ get identified by the quotient map $\mathrm{SL}_2\mathbb{R} \twoheadrightarrow \mathrm{PSL}_2\mathbb{R} \cong \mathrm{SO}^{\uparrow}(1,2)$.

Notice that Taylor includes an additional factor of $-1/2$ in his definition of A and of $1/2$ in his definition of B. We compensate for this in our definition of $X_\pm$, so that Z and $X_\pm$ are chosen like in [Tay86, section 8.1].

For the following we need to introduce some additional notation. Let us define:

$$B := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (N_+ + N_-) , \tag{6.56}$$

as well as some elements $X_\pm$ of the complexified Lie algebra $\mathfrak{sl}(2,\mathbb{C})$, which will play the role of *raising* and *lowering operators* with respect to Z (notice that X_+ and X_- will *not* be represented by skew-symmetric operators in the following):

$$X_\pm := -\tfrac{1}{2}(A \pm iB) , \tag{6.57}$$

with commutators:

$$[Z, X_\pm] = \pm 2iX_\pm , \quad [X_+, X_-] = -iZ . \tag{6.58}$$

The Casimir operator will in general *not* be quantizable. In fact, this happens to be the case for the group $\mathrm{SL}_2\mathbb{R}$, where the correspondence (6.52) yields $C \mathrel{\hat{=}} 0$ for the classical theory, but C is typically represented by a non-zero quantum operator.

In addition, we need the Casimir operator C (see [Tay86, section 8.1, eq. (1.25)]):

$$C = Z^2 - A^2 - B^2 = (N_+ - N_-)^2 - A^2 - (N_+ + N_-)^2 . \tag{6.59}$$

The Casimir operator as well is not an element of $\mathfrak{sl}(2,\mathbb{R})$ itself but lies in the centre of the universal enveloping algebra $\mathcal{U}(\mathfrak{sl}(2,\mathbb{R}))$ (see [FH91, section 25.1 and appendix C], [Tay86, section 0.3] or [Lan85, section X.1] for the Casimir operator).

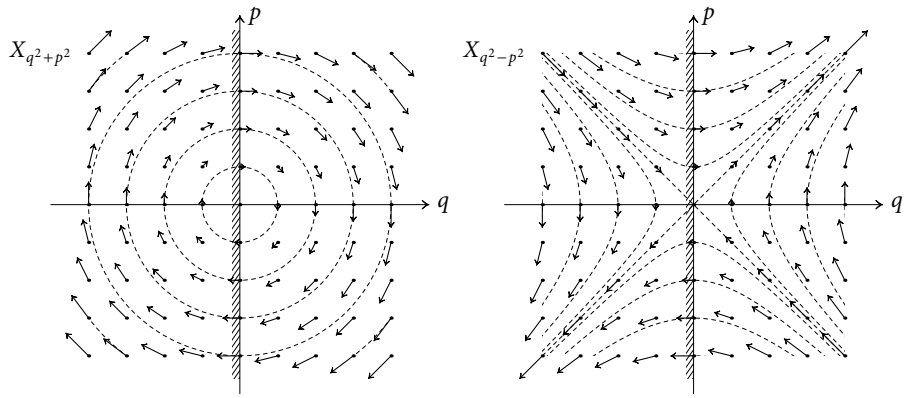


Figure 6.11 The Hamiltonian vector fields of $q^2 + p^2$ and $q^2 - p^2$ correspond (up to constants) to the Lie algebra elements Z and B, respectively.

Remark 6.4.4 (Universal enveloping algebra) The ***universal enveloping algebra*** of $\mathfrak{g}$ is the quotient $\mathcal{U}\mathfrak{g} = \mathcal{T}\mathfrak{g}/\mathcal{I}$ of unital associative algebras, of the tensor algebra $\mathcal{T}\mathfrak{g}$ over $\mathfrak{g}$ seen as complex vector space, modulo the two-sided ideal $\mathcal{I} = \langle X \otimes Y - Y \otimes X - [X, Y] : X, Y \in \mathfrak{g} \rangle$. Dividing out the ideal implements the Lie brackets of $\mathfrak{g}$ as commutators in the universal enveloping algebra $\mathcal{U}\mathfrak{g}$, i. e. $[X, Y] = X \otimes Y - Y \otimes X$ for all $X, Y \in \mathfrak{g}$.

The universal enveloping algebra $\mathcal{U}\mathfrak{g}$ with embedding $\iota : \mathfrak{g} \to \mathcal{U}\mathfrak{g}$ (the so-called *structure map*; a Lie algebra morphism) has the following *universal property*: for any unital associative algebra $\mathcal{A}$ over $\mathbb{C}$, let $L(\mathcal{A})$ be the Lie algebra we obtain from $\mathcal{A}$ if we define the Lie bracket of $L(\mathcal{A})$ as the commutator $[a, b] := ab - ba$ for $a, b \in \mathcal{A}$ using the multiplication in $\mathcal{A}$. Then, for any $\mathcal{A}$, any *Lie algebra morphism* $\varphi : \mathfrak{g} \to L(\mathcal{A})$ induces a unique *algebra morphism* $\tilde{\varphi} : \mathcal{U}\mathfrak{g} \to \mathcal{A}$ such that the following diagram commutes:

Notice that a Lie algebra is a non-associative algebra. There is the Jacobi identity instead of associativity.

$$\begin{array}{ccc} \mathcal{U}\mathfrak{g} & \overset{\tilde{\varphi}}{\cdots\cdots\cdots\cdots\rightarrow} & \mathcal{A} \\ {\scriptstyle\iota}\uparrow & \Uparrow & \downarrow{\scriptstyle L} \\ \mathfrak{g} & \xrightarrow[\varphi]{} & L(\mathcal{A}) \end{array} .$$

From a categorial point of view, L and $\mathcal{U}$ are covariant functors. An argument based on the above diagram shows that the universal algebra functor $\mathcal{U}$ is left adjoint to L (the double arrow in the diagram is a *natural transformation*; see [HS71, section VII.1]). An important consequence is that algebra representations of the universal enveloping algebra $\mathcal{U}\mathfrak{g}$ correspond in a one-to-one manner to Lie algebra representations of $\mathfrak{g}$ (the bracket is realised 'implicitly', so to say, by representations of $\mathcal{U}\mathfrak{g}$), and this is the deeper reason why the universal enveloping algebra $\mathcal{U}\mathfrak{g}$ and the Casimir operator $C \in \mathcal{Z}(\mathcal{U}\mathfrak{g})$ play such an important role in the representation theory of Lie groups.

Now, following [Tay86, section 8.2] and [Lan85, chapter VI], let U be an irreducible unitary representation of $\mathrm{SL}_2\mathbb{R}$ on some Hilbert space $\mathcal{H}$, and let:

$$E := \mathcal{L}U(Z) , \quad R_\pm := \mathcal{L}U(X_\pm) , \tag{6.60}$$

where $\mathcal{L}U$ is the *derived representation* $\mathcal{L}U$ of the Lie algebra ($\mathcal{L}$ is the Lie functor) [Lan85, section VI.1]. Then $\mathcal{H}$ admits an orthogonal direct sum decomposition into *weight spaces* $\mathcal{H}_k$ for the action of the maximal compact subgroup $K_1 = \mathrm{SO}(2)$:

$\mathcal{L}U(X) = \dfrac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} U\big(\exp(tX)\big)$

For the theory behind *weights*, see [Tay86, section 3.2].

$$\mathcal{H} = \bigoplus_k \mathcal{H}_k \,, \tag{6.61}$$

The domain of E is actually only a *subset* of $\mathcal{H}_k$ (remember the discussion of unbounded operators from section 1.6), but it can be shown that iE is *essentially self-adjoint*. To save some pages we will skip over this technical complication. Please consult [Tay86] or [Lan85] if you want to fill in the blanks.

so that elements in K_1 act on the weight space $\mathcal{H}_k$ according to:

$$U\big(\exp(\vartheta Z)\big)|\psi\rangle = e^{i\vartheta k}|\psi\rangle \quad \Rightarrow \quad -iE|\psi\rangle = k|\psi\rangle \,, \tag{6.62}$$

for all $|\psi\rangle \in \mathcal{H}_k$. In other words, $\mathcal{H}_k$ is the eigenspace with respect to the eigenvalue k of the self-adjoint operator $-iE$ (the operator E itself is skew-adjoint). Accordingly, we will in the following usually write $|k\rangle = |\psi\rangle$ for elements $|\psi\rangle \in \mathcal{H}_k$.

To determine which values of k may appear in the decomposition, remember that the identity e of the group must be represented by the identity operator $\mathbb{1}$ on $\mathcal{H}$. In the present case we have $\exp(2\pi Z) = \mathbb{1}$ because the exponential $\exp : \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}_2\mathbb{R}$ is just the usual matrix exponential $\exp(\vartheta Z) = e^{\vartheta Z}$. Accordingly, equation (6.62) implies that k must satisfy $e^{i2\pi k} = 1$. Therefore, the values k must be *integers*:

$$k \in I \subseteq \mathbb{Z} \,. \tag{6.63}$$

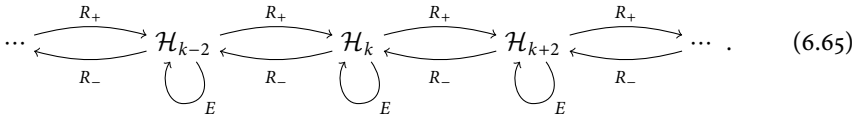
Mathematically, $R_\pm$ represent *root vectors* of the Lie algebra. The algebra $\mathfrak{sl}(2,\mathbb{R})$ has *roots* $+2$ and -2 [FH91, chapter 12].

Continuing with the analysis, we find that the operators R_+ and R_- act as raising and lowering operators on eigenvectors $|k\rangle \in \mathcal{H}_k$ of $-iE$:

$$-iE(R_\pm|k\rangle) = -i\big(R_\pm(E|k\rangle) \pm 2iR_\pm|k\rangle\big) = (k \pm 2)(R_\pm|k\rangle) \,. \tag{6.64}$$

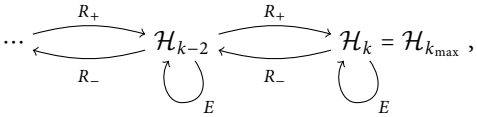
Taking into account the irreducibility, we thus get a single unbroken sequence of weight spaces $\mathcal{H}_k$ (for a reducible representation the representation space $\mathcal{H}$ could be the union of several disjoint sequences) [FH91, section 11.1]:

Given $k \in \mathbb{Z}$ there are exactly two sequences like (6.65) that continue to infinity in both directions – one for even and one for odd integers.



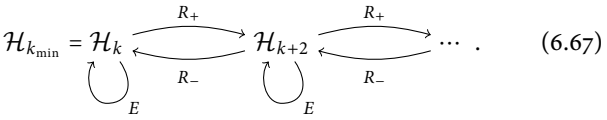
(6.65)

While the sequence is not allowed to break into disjoint sequences, it can happen that the sequence terminates in one direction or the other. This is the case if there exists a *maximal weight* $k = k_{\max}$ with $R_+|k_{\max}\rangle = 0$, which results in a sequence:



(6.66)

or a *minimal weight* $k = k_{\min}$ with $R_-|k_{\min}\rangle = 0$, which yields a sequence like:



(6.67)

On the other hand, it can be shown that any non-trivial irreducible unitary representation of a non-compact semisimple Lie group must be infinite dimensional, yet each weight space $\mathcal{H}_k$ is only one-dimensional (see [Tay86, section 8.2] for the details). Hence, unless U is the trivial representation a sequence must consist of infinitely many weight spaces. Therefore, and in contrast to unitary representations of a *compact* Lie group, the sequence cannot end in both directions.

As a result, we have four qualitatively different possibilities for the spectrum of the self-adjoint operator $-iE$. We get two possibilities for the case where the sequence in (6.65) is infinite in both directions (the superscript 'e' means 'even'; 'o' stands for 'odd'):

$$\sigma^{\mathrm{e}}(-iE) = \{2k : k \in \mathbb{Z}\} , \tag{6.68a}$$
$$\sigma^{\mathrm{o}}(-iE) = \{2k+1 : k \in \mathbb{Z}\} , \tag{6.68b}$$

and two types of spectra for half-infinite sequences, one type for each of the cases (6.66) and (6.67), indexed by the minimal (or maximal) weight k_0:

$$\sigma^{+}_{k_0}(-iE) = \{k_0 + 2k : k \in \mathbb{Z}^+_0\} , \tag{6.68c}$$
$$\sigma^{-}_{k_0}(-iE) = \{k_0 - 2k : k \in \mathbb{Z}^+_0\} . \tag{6.68d}$$

Note that [Bar47] and [Kas03] both use a different convention which amounts to an additional factor of $1/2$ in the definition of Z. Accordingly, they discuss spectra of $-\frac{1}{2}iE$. If you want to compare the results, you just have to multiply eigenvalues in our notation by $1/2$.

There are, however, further restrictions. A spectrum of the form (6.68c), for example, requires a ground state $|k_0\rangle$, such that $R_-|k_0\rangle = 0$ but $R_+|k_0\rangle \neq 0$, yet such a ground state does not exist for arbitrary values of k_0. In order to determine for which k_0 a suitable ground state exists we will calculate the operator norms of R_+ and R_-. To this end, consider that the definition (6.57) of $X_\pm$ implies:

$$X_+X_- + X_-X_+ = \tfrac{1}{2}(A^2 + B^2) ,$$

so that the Casimir operator $C = Z^2 - A^2 - B^2$ can be written as:

$$C = Z^2 - 2(X_+X_- + X_-X_+) .$$

On the other hand, the commutator $[X_+, X_-] = -iZ$ from equation (6.58) yields:

$$2(X_+X_- - X_-X_+) = -2iZ .$$

Adding and subtracting these two equations we get:

$$4X_-X_+ = Z^2 + 2iZ - C , \quad 4X_+X_- = Z^2 - 2iZ - C , \tag{6.69}$$

or, written in terms of the unitary irreducible representation U:

$$4R_-R_+ = E^2 + 2iE - \lambda\mathbb{1} , \quad 4R_+R_- = E^2 - 2iE - \lambda\mathbb{1} , \tag{6.70}$$

where the value $\lambda \in \mathbb{C}$ is the eigenvalue of the Casimir operator which, due to Schur's lemma, must be represented proportional to the identity, i. e. by $\mathcal{L}U(C) = \lambda\mathbb{1}$.

Furthermore, since $\mathcal{L}U$ represents Lie algebra elements by skew-symmetric operators the defining equation (6.57) of $X_\pm$ yields the relation $R_+^\dagger = -R_-$ between the raising and lowering operators. Taking into account that E produces the eigenvalue ik when evaluated on an eigenstate $|k\rangle \in \mathcal{H}_k$, we obtain:

$$\|R_\pm\|^2_{L(\mathcal{H}_k, \mathcal{H}_{k\pm2})} = -\langle k|R_\mp R_\pm|k\rangle = -\tfrac{1}{4}(-k^2 \mp 2k - \lambda) = \tfrac{1}{4}(k^2 \pm 2k + \lambda) ,$$

and therefore (see also [Tay86, section 8.2, equations (2.18) and (2.19)]):

$$\|R_\pm\|_{L(\mathcal{H}_k, \mathcal{H}_{k\pm2})} = \tfrac{1}{2}\big((k \pm 1)^2 + \lambda - 1\big)^{1/2} , \tag{6.71}$$

Let us now write:

$$R_+|k\rangle = \alpha_k|k+2\rangle , \quad R_-|k+2\rangle = \beta_k|k\rangle ,$$

then, given equation (6.71), α_k is a complex number with absolute value:

$$|\alpha_k| = \tfrac{1}{2}\big((k+1)^2 + \lambda - 1\big)^{1/2}\,, \tag{6.72a}$$

and $\beta_k = -\alpha_k^*$ (due to $R_- = -R_+^\dagger$; α_k^* denotes the complex conjugate of α_k), so that:

$$|\beta_{k-2}| = \tfrac{1}{2}\big((k-1)^2 + \lambda - 1\big)^{1/2}\,. \tag{6.72b}$$

Spectra of the form (6.68c) belong to representations of the type $U^+_{k_0}$ defined below.

Finally, to obtain a spectrum of the form (6.68c) we mentioned there has to exist a ground state $|k_0\rangle$ such that $R_-|k_0\rangle = 0$ but $R_+|k_0\rangle \neq 0$. The former requires $\beta_{k_0-2} = 0$ and hence fixes the value λ of the Casimir operator as $\lambda = 1 - (k_0 - 1)^2$. The second condition, $R_+|k_0\rangle \neq 0$, on the other hand, requires $|\alpha_{k_0}| > 0$ and thus $(k_0+1)^2 + \lambda - 1 > 0$. Inserting λ we get $(k_0+1)^2 + 1 - (k_0-1)^2 - 1 > 0$, which simplifies to $k_0 > 0$. As we also have $k \in \mathbb{Z}$ in the present situation, a spectrum like (6.68c) is therefore possible if and only if $k_0 \geq 1$.

A detailed case by case analysis following the lines of this example reveals that all non-trivial irreducible unitary representations of $\mathrm{SL}_2\mathbb{R}$ can be classified up to unitary equivalence by the type of spectrum of $-iE$, chosen from equation (6.68), and the value λ of the Casimir operator. We do not want to replicate the analysis for the other cases here; instead, we will just quote the results from [Tay86, section 8.2]:

The unbounded spectra σ^{e} and σ^{o} of $-iE$ give rise to the two families of *principal series* representations and the *complementary series*:

It is common to allow both signs of s, but then $U^{\mathrm{e}}_{is} \cong U^{\mathrm{e}}_{-is}$ for $s \in \mathbb{R}$. For the second principal series, this means $U^{\mathrm{o}}_{is} \cong U^{\mathrm{o}}_{-is}$ for $s \in \mathbb{R} \setminus \{0\}$, and for the complementary series $U^{\mathrm{e}}_s \cong U^{\mathrm{e}}_{-s}$ for $s \in (-1;1) \setminus \{0\}$.

First principal series U^{e}_{is} if $-iE$ has an unbounded spectrum of type $\sigma^{\mathrm{e}}(-iE)$ and the value of the Casimir operator is $\lambda = 1 + s^2$ for $s \in \mathbb{R}^+_0$ (including $s = 0$).

Second principal series U^{o}_{is} if $-iE$ has an unbounded spectrum of type $\sigma^{\mathrm{o}}(-iE)$ and the value of the Casimir operator is $\lambda = 1 + s^2$ for $s \in \mathbb{R}^+$ (excluding $s = 0$).

Complementary series U^{e}_s if $-iE$ has an unbounded spectrum of type $\sigma^{\mathrm{e}}(-iE)$ and the value of the Casimir operator is $\lambda = 1 - s^2$ for $s \in (0;1)$.

The bounded spectra $\sigma^\pm_{k_0}$, on the other hand, belong to the two *discrete series* and the *mock discrete series* representations:

The value k_0 is sometimes called *Bargmann index*. It is the 'ground state' eigenvalue of the spectrum of $-iE$.

Holomorphic (or **positive) discrete series** $U^+_{k_0}$ if $-iE$ has a bounded spectrum of type $\sigma^+_{k_0}(-iE)$ with $k_0 \in \mathbb{Z}$ and $k_0 \geq 2$ (the value λ of the Casimir operator is in this case given by $\lambda = 1 - (k_0-1)^2 = -k_0(k_0-2)$).

Antiholomorphic (or **negative) discrete series** $U^-_{k_0}$ if $-iE$ has a bounded spectrum of type $\sigma^-_{k_0}(-iE)$ with $k_0 \in \mathbb{Z}$ and $k_0 \leq -2$ (the value of the Casimir operator is in this case given by $\lambda = 1 - (k_0+1)^2 = -k_0(k_0+2)$).

Mock (or **limit of) discrete series** U^+_1 **and** U^-_{-1} if $-iE$ has a bounded spectrum of type $\sigma^+_{k_0}(-iE)$ for the special case of $k_0 = 1$, or of type $\sigma^-_{k_0}(-iE)$ for $k_0 = -1$, respectively.

The classification is completed by the (finite dimensional) **trivial representation**.

Remark 6.4.5 (Mock discrete series) The reason to put U^+_1 and U^-_{-1} in a separate class and not treat them as part of the *positive* and *negative discrete series* is that they have Plancherel measure zero and do not appear in the *Plancherel formula* for $\mathrm{SL}_2\mathbb{R}$. Apart from that, however, U^+_1 and U^-_{-1} behave a lot like other discrete series representations and it often makes sense to treat them as if they were part of the discrete series. Kastrup, for example, makes no distinction and calls $U^+_{k_0}$ for all values $k_0 \geq 1$ (in our notation) positive discrete series representations [Kas03, appendix B.3]. Bargmann speaks of the 'discrete class' representations $U^+_{k_0}$ and $U^-_{-k_0}$ for $k_0 \geq 1$ [Bar47, section 5i].

The *Plancherel formula* can be found in [Lan85, chapter VIII] but is irrelevant to our present discussion.

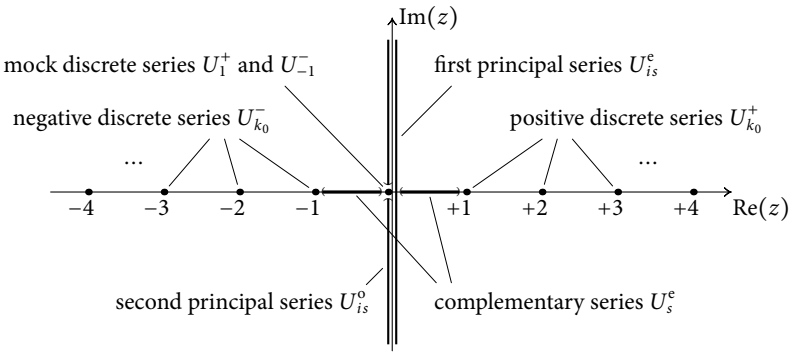


Figure 6.12 Classification of the irreducible unitary representations of $G = \mathrm{SL}_2\mathbb{R}$ and parametrisation of the unitary dual $\hat{G}$ in terms of the complex parameter z (see remark 6.4.6; figure according to [Lan85, section VI.6, figure 1]).

Remark 6.4.6 It is instructive to compare Taylor's result with the classification given in Lang's book [Lan85, section VI.6, theorem 8]. For this, recall that we found the relation $\lambda = 1 - (k_0 - 1)^2$ for representations of the positive discrete series $U_{k_0}^+$ (including the limit case $k_0 = 1$). Inserting this into equation (6.72a), we get:

$$|\alpha_k| = \tfrac{1}{2}\big((k+1)^2 - (k_0-1)^2\big)^{1/2} .$$

Accordingly, we can set α_k to:

$$\alpha_k := \tfrac{1}{2}\big((k+1) + z\big) , \tag{6.73}$$

if we define $z := k_0 - 1$ for all representations of type $U_{k_0}^+$.

The same formula (6.73) for α_k holds for representations of the negative discrete series $U_{k_0}^-$ (for $k_0 \leq -1$) provided we set $z := k_0 + 1$ in this case. Continuing, we can write $z = is$ for the first and second principal series U_{is}^e and U_{is}^o, as well as $z = s$ for the complementary series U_s^e, always resulting in the formula (6.73) for α_k. In addition, this convention is supported by the fact that there exists a formal limit $s \to 0$ for the second principal series U_s^o, and it turns out that the representation U_0^o is reducible and splits into the direct sum $U_0^o = U_1^+ \oplus U_{-1}^-$ [Tay86, section 8.2, equation (2.56)].

Looking at [Lan85, section VI.6] you will find that Lang bases the classification first of all on the parameter z (although the parameter is called s in his book), and that z allows a parametrisation of the *unitary dual* $\hat{G}$ according to figure 6.12. Notice, however, that representations of the positive discrete series (including the mock discrete series representation U_1^+) are then labelled by π_z for $z \in \mathbb{Z}_0^+$, starting with $z = 0$, and the index is no longer the 'ground state' eigenvalue k_0 of $-iE$. This is why Taylor's convention is better suited to our discussion.

As a convention, we will in the following include the mock discrete series representations when we speak of the positive or negative discrete series!

6.4.3 Representations of $\mathrm{SO}^\uparrow(1,2)$ and Covering Groups

We will see shortly that not $\mathrm{SL}_2\mathbb{R}$ itself but one of its coverings is the appropriate Canonical Group for our phase space $\mathbb{R}^2/\mathbb{Z}_2$. To identify the Canonical Group we hence

need to classify the irreducible unitary representations of $\mathrm{SO}^{\uparrow}(1,2)$ and of the other covering groups. Fortunately, most of the work is already done. As all these groups have the same Lie algebra $\mathfrak{sl}(2,\mathbb{R})$, the results from above which depend only on properties of the Lie algebra remain valid. In particular, the fundamental equation (6.64) for the raising and lowering operators $R_{\pm}$ remains correct, so that $R_{\pm}$ still maps vectors from $\mathcal{H}_k$ to vectors in $\mathcal{H}_{k\pm 2}$. Consecutive eigenvalues of $-iE$ therefore differ by a value of 2 not only for $\mathrm{SL}_2\mathbb{R}$ but for *all* the covering groups of $\mathrm{SO}^{\uparrow}(1,2)$.

What changes is the set of values that the index k may assume in the decomposition of $\mathcal{H}$ into weight spaces $\mathcal{H}_k$. This is where we used the specific group $G_1 = \mathrm{SL}_2\mathbb{R}$ to derive equation (6.63) which states that k takes on integer values. The argument was that the identity e of the group is represented by the identity operator $\mathbb{1}$ on $\mathcal{H}$. On the other hand, we had the action of the maximal compact subgroup $K_1 = \mathrm{SO}(2)$ on vectors $|k\rangle \in \mathcal{H}_k$ according to equation (6.62):

$$U\big(\exp(\vartheta Z)\big)|k\rangle = e^{i\vartheta k}|k\rangle \,.$$

Given $\exp(2\pi Z) = \mathbb{1}$ for the exponential map $\exp : \mathfrak{sl}(2,\mathbb{R}) \to \mathrm{SL}_2\mathbb{R}$, the values of k were hence determined by the requirement $e^{i2\pi k} = 1$, which implied $k \in \mathbb{Z}$.

Before we treat the general case we will first show how representations of the group $\mathrm{SO}^{\uparrow}(1,2) \cong \mathrm{PSL}_2\mathbb{R}$ are related to those of its double covering $\mathrm{SL}_2\mathbb{R}$ using an argument given in [Tay86, section 2.1]. For this, let $U : \mathrm{PSL}_2\mathbb{R} \to U(\mathcal{H})$ be a unitary representation of $\mathrm{PSL}_2\mathbb{R}$ and observe that the representation of $\mathrm{PSL}_2\mathbb{R}$ lifts to a representation $\hat{U}$ of the covering group $\mathrm{SL}_2\mathbb{R}$ via the covering morphism $p : \mathrm{SL}_2\mathbb{R} \twoheadrightarrow \mathrm{PSL}_2\mathbb{R}$:

The diagram also shows that each *faithful* representations of $\mathrm{PSL}_2\mathbb{R}$ is lifted to an only *almost faithful* representation of $\mathrm{SL}_2\mathbb{R}$.

$$\begin{array}{ccc} \mathrm{SL}_2\mathbb{R} & & \\ \Big\downarrow{\scriptstyle p} & \overset{\hat{U}}{\searrow} & \\ \mathrm{PSL}_2\mathbb{R} & \xrightarrow[U]{} & U(\mathcal{H}) \end{array} \quad . \tag{6.74}$$

We know that the kernel $\ker p = \{\pm\mathbb{1}\}$ of the covering map is a discrete normal subgroup of $\mathrm{SL}_2\mathbb{R}$, so that $\ker p$ lies in the centre $\mathcal{Z}(\mathrm{SL}_2\mathbb{R})$ [Sim96, theorem VII.6.2]. Therefore, if U is irreducible the representation $\hat{U}$ will be irreducible as well. We have, however, already classified all irreducible unitary representations of $\mathrm{SL}_2\mathbb{R}$, and the representation $\hat{U}$ must hence be one of them – the irreducible unitary representations of $\mathrm{PSL}_2\mathbb{R}$ are a *subset* of the irreducible unitary representations of $\mathrm{SL}_2\mathbb{R}$.

Now in the opposite direction, a given unitary representation $\hat{U}$ of $\mathrm{SL}_2\mathbb{R}$ yields a representation of $\mathrm{PSL}_2\mathbb{R}$ if and only if $\hat{U}$ factors through the covering map p, that is, if there exists some mapping U such that $\hat{U} = U \circ p$. We can see that $\hat{U}$ factors through p if and only if $\hat{U}$ is constant on the kernel $\ker p$. Moreover, given that $e \in \ker p$ and that the group identity e must be represented by $\hat{U}(e) = \mathbb{1}$, it follows that $\hat{U}(\ker p) = \mathbb{1}$.

Notice that the kernel $\{\pm\mathbb{1}\}$ lies in the maximal compact subgroup K_1 of $\mathrm{SL}_2\mathbb{R}$. We will see later, in theorem 6.4.7, that this is not just a coincidence.

Via the exponential, the kernel $\ker p = \{\pm\mathbb{1}\} \subset \mathrm{SL}_2\mathbb{R}$ can be parameterised as:

$$\ker p = \big\{\exp\big(\tfrac{1}{2}\vartheta Z\big) : \vartheta \in 2\pi\mathbb{Z}\big\} \,. \tag{6.75}$$

Given the usual action (6.62) of K_1 on $\mathcal{H}_k$, the condition $\hat{U}(\ker p) = \mathbb{1}$ hence requires $e^{i\pi k} = 1$. Accordingly, *an irreducible unitary representation of* $\mathrm{SL}_2\mathbb{R}$ *factors through to an irreducible unitary representation of* $\mathrm{PSL}_2\mathbb{R}$ *if and only if* $k \in 2\mathbb{Z}$.

Using the classification of irreducible unitary representations for $\mathrm{SL}_2\mathbb{R}$ from above, the classification for $\mathrm{SO}^{\uparrow}(1,2) \cong \mathrm{PSL}_2\mathbb{R}$ is hence given as follows:

- all the *first principal series* and the *complementary series* representations of $\mathrm{SL}_2\mathbb{R}$ are also representations of $\mathrm{PSL}_2\mathbb{R}$, but not the second principal series representations,
- half of the *positive* and the *negative discrete series* representations $U^{\pm}_{\pm k_0}$ of $\mathrm{SL}_2\mathbb{R}$ are as well representations of $\mathrm{PSL}_2\mathbb{R}$, namely for *even* numbers $k_0 \in 2\mathbb{Z}$ (with $k_0 > 0$).

For the positive discrete series representations of $\mathrm{SO}^{\uparrow}(1,2)$, in particular, the spectrum of the operator $-iE$ is given by:

$$\sigma^{+}_{k_0}(-iE) = \{2k + k_0 : k \in \mathbb{Z}^{+}_0\}\,, \quad k_0 \in 2\mathbb{Z}^{+}\,. \tag{6.76}$$

Finally, let us consider representations for a general covering group of $\mathrm{SL}_2\mathbb{R}$. We will see that the discussion of the double covering $\mathrm{SL}_2\mathbb{R} \twoheadrightarrow \mathrm{PSL}_2\mathbb{R}$ can be generalised appropriately. An important detail, though, necessary to derive equation (6.75) (which ultimately determined the values k) was that the kernel of the covering map p is part of the maximal compact subgroup K_1 of $\mathrm{SL}_2\mathbb{R}$. Incidentally, we already know that the kernel of a covering morphism p lies in the centre $\mathcal{Z}(G)$ for any connected Lie group G. The following theorem about the *Cartan decomposition* then does the heavy lifting and shows that the centre $\mathcal{Z}(G)$ is contained in the maximal compact subgroup of G.

Bargmann implicitly uses the Cartan decomposition in [Bar47, section 4] to construct the universal covering group of $\mathrm{SL}_2\mathbb{R}$.

Theorem 6.4.7 ([Kna96, theorem 6.31]) Let G be a semisimple Lie group, let θ be a Cartan involution of its Lie algebra $\mathfrak{g}$, let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the corresponding Cartan decomposition, and let K be the analytic subgroup of G with Lie algebra $\mathfrak{k}$. Then

(a) there exists a Lie group automorphism Θ of G (the *global Cartan involution* of G) with differential θ, and Θ has $\Theta^2 = 1$
(b) the subgroup of G fixed by Θ is K
(c) the mapping $K \times \mathfrak{p} \to G$ given by $(k, X) \mapsto k\exp(X)$ is a diffeomorphism onto G (the decomposition $G \cong K \times \mathfrak{p}$ is the *global Cartan decomposition* of G)
(d) K is closed
(e) K contains the centre $\mathcal{Z}(G)$ of G
(f) K is compact if and only if $\mathcal{Z}(G)$ is finite
(g) when $\mathcal{Z}(G)$ is finite, K is a maximal compact subgroup of G.

Definition 6.4.8 (Cartan decomposition) Let B be the Killing form of a real semisimple Lie algebra $\mathfrak{g}$. An involution θ of $\mathfrak{g}$ is called a ***Cartan involution*** if the symmetric bilinear form $B_\theta(X, Y) := -B(X, \theta Y)$ for $X, Y \in \mathfrak{g}$ is positive definite.

The Killing form B of a real semisimple Lie group is non-degenerate and indefinite. It can be shown that every real semisimple Lie algebra $\mathfrak{g}$ has a Cartan involution [Kna96, corollary 6.18].

A Cartan involution yields a ***Cartan decomposition*** $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, where $\mathfrak{g}$ splits into eigenspaces $\mathfrak{k}$ and $\mathfrak{p}$ of the involution θ with eigenvalues $+1$ and -1, respectively (so that $\theta|_{\mathfrak{k}} = +1$ and $\theta|_{\mathfrak{p}} = -1$). The eigenspace $\mathfrak{p}$ does not close under the Lie bracket, because $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}$. The eigenspace $\mathfrak{k}$, on the other hand, is a subalgebra of $\mathfrak{g}$. The direct sum in $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is a direct sum of vector spaces.

For our particular case $G = \mathrm{SL}_2\mathbb{R}$ it follows from part (g) of the theorem that the analytic subgroup K in the Cartan decomposition $G \cong K \times \mathfrak{p}$ is the maximal compact subgroup $K = \mathrm{SO}(2)$ of $\mathrm{SL}_2\mathbb{R}$ (the Iwasawa decomposition $G = KAN$ is actually a refinement of the Cartan decomposition; see [Kna96, section VI.4]). We want to remark that a Cartan involution of $\mathfrak{sl}(2,\mathbb{R})$ that yields the decomposition $\mathrm{SL}_2\mathbb{R} \cong \mathrm{SO}(2) \times \mathfrak{p}$ is given by $\theta(X) = -X^t$ for all $X \in \mathfrak{sl}(2,\mathbb{R})$ but we will not need it in the following. Instead, let us draw some conclusions from theorem 6.4.7:

Corollary 6.4.9 (Structure of covering groups) Let G be a connected semisimple Lie group, $\mathfrak{g}$ its Lie algebra. Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition of the Lie algebra $\mathfrak{g}$ and let $G \cong K \times \mathfrak{p}$ be the corresponding Cartan decomposition of G. Furthermore, let G' be a covering group of G and let $p : G' \twoheadrightarrow G$ be the covering morphism. Then:

(i) the Cartan decomposition of the covering group G' is given by $G' \cong K' \times \mathfrak{p}$, via the diffeomorphism $(k', X) \mapsto k' \exp_{G'}(X)$, where K' is the analytic subgroup of G' with Lie algebra $\mathfrak{k}$,

(ii) the kernel $\ker p$ of the covering morphism p lies in the subgroup K'. The restriction $p_K : K' \twoheadrightarrow K$ of the covering map p has kernel $\ker p_K = \ker p$. In particular, we get $K \cong K' / \ker p$.

Proof. (i) This is a direct consequence of theorem 6.4.7.

(ii) For a connected Lie group G we already mentioned in remark 6.4.1 that the kernel of the covering morphism lies in the centre $\ker p \subseteq \mathcal{Z}(G')$ of the covering group G', because $\ker p \subseteq G'$ is a discrete normal subgroup. Moreover, due to semisimplicity of G', part (e) of theorem 6.4.7 yields $\mathcal{Z}(G') \subseteq K'$, so that $\ker p \subseteq K'$ follows. Therefore, $\ker p_K = \ker p$ and we obtain $K \cong K' / \ker p_K = K' / \ker p$. ∎

In simple words, corollary 6.4.9 states that all the covering groups of a connected semisimple Lie group G are *completely determined by covering groups of the maximal compact subgroup* K. So, when we want to classify irreducible unitary representations of all the covering groups of $G = \mathrm{SL}_2\mathbb{R}$, we just have to consider coverings of the maximal compact subgroup $K = \mathrm{SO}(2) \cong S^1$.

When we write the circle group as $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, an n-fold cover of S^1 over itself is given by $p_n : S^1 \twoheadrightarrow S^1$, $z = e^{i\varphi} \mapsto z^n = e^{in\varphi}$ [Hato2, section 1.3]. The kernel of p_n consists of all nth roots of unity (see figure 6.13) and can be parametrised as:

$$\ker p_n = \left\{ e^{i(2\pi m)/n} : m = 0, \ldots, n-1 \right\}. \tag{6.77}$$

Furthermore, the universal covering group of S^1 is isomorphic to $(\mathbb{R}, +)$ with covering morphism $p_\infty : \mathbb{R} \to S^1$, $\varphi \mapsto e^{i\varphi}$ and kernel $\ker p_\infty = 2\pi\mathbb{Z}$. To parametrise the kernel of p_∞ in the fashion of equation (6.77), recall that the Lie algebra $\mathcal{L}\mathbb{R} = T_0\mathbb{R}$ can be identified with $\mathbb{R}$. Thus, the exponential map $\exp_{\mathbb{R}} : \mathcal{L}\mathbb{R} \to \mathbb{R}$ is the identity $\exp_{\mathbb{R}} = \mathrm{id}_{\mathbb{R}}$, and we obtain:

$$\ker p_\infty = 2\pi\mathbb{Z} = \left\{ \exp_{\mathbb{R}}(2\pi m) : m \in \mathbb{Z} \right\}. \tag{6.78}$$

Now, to lift this result from S^1 to the whole group let $\widetilde{\exp} : \mathfrak{sl}(2, \mathbb{R}) \to \widetilde{\mathrm{SL}_2}\mathbb{R}$ be the exponential map of the universal covering group. Then it follows from corollary 6.4.9 and equation (6.78) that the kernel of the covering map $p_\infty : \widetilde{\mathrm{SL}_2}\mathbb{R} \twoheadrightarrow \mathrm{SL}_2\mathbb{R}$ can be parametrised as (the generator $Z \in \mathfrak{k}$ is the same as before):

This result for the universal covering is mentioned near the end of [Tay86, section 8.2] but no proof is given.

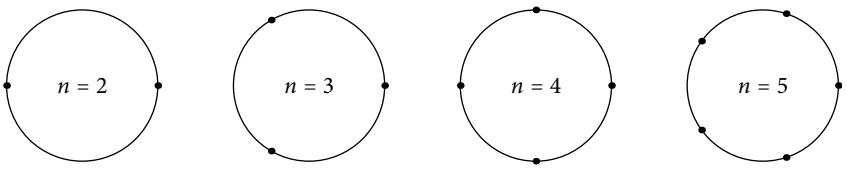


Figure 6.13 The kernel of the n-fold covering $p_n : S^1 \to S^1$, $z \mapsto z^n$ of the circle group S^1 over itself consists of all nth roots of unity.

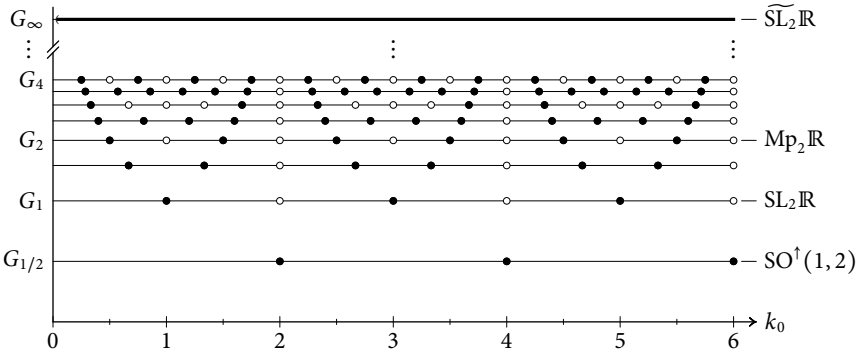


Figure 6.14 Possible *ground state eigenvalues* k_0 of $-iE$ for irreducible unitary representations of the positive discrete series $U^+_{k_0}$, pictured for covering groups of $G_{1/2} = \mathrm{SO}^\uparrow(1,2)$. Filled dots indicate *faithful* representations, which lift to unfaithful representations of the covering groups.

$$\ker p_\infty = \left\{\widetilde{\exp}(2\pi m Z) : m \in \mathbb{Z}\right\} . \tag{6.79}$$

If $p_n : G_n \twoheadrightarrow G$ is an n-fold covering of $G = \mathrm{SL}_2\mathbb{R}$, on the other hand, we have:

$$\ker p_n = \left\{\exp_{G_n}(2\pi m Z) : m = 0, \ldots, n-1\right\} . \tag{6.80}$$

The elements in $\ker p_n$ are all mapped to the identity $\mathbb{1} \in G$ under the covering map p_n but they are distinct elements of the covering group G_n, just like we have seen in the case of $\mathrm{SL}_2\mathbb{R} \twoheadrightarrow \mathrm{PSL}_2\mathbb{R}$. So, while $\exp(\vartheta Z) = \mathbb{1}$ for $\vartheta = 2\pi\mathbb{Z}$ in $\mathrm{SL}_2\mathbb{R}$, we have $\exp_{G_n}(\vartheta Z) = \mathbb{1}$ for $\vartheta = 2\pi n\mathbb{Z}$ in the covering G_n. Accordingly, the values k in the decomposition of $\mathcal{H}$ into weight spaces for an n-fold cover of $\mathrm{SL}_2\mathbb{R}$ are determined by the condition $e^{ik2\pi n} = 1$. In summary we therefore obtain the following results:

- for irreducible unitary representations of an n-fold covering G_n of $G_1 = \mathrm{SL}_2\mathbb{R}$, the index k may assume values $k \in (1/n)\mathbb{Z}$ (we can abuse the notation a little and write $G_{1/2} = \mathrm{PSL}_2\mathbb{R}$ for the '1/2-fold cover' of $\mathrm{SL}_2\mathbb{R}$),
- for irreducible unitary representations of the universal covering group $\widetilde{\mathrm{SL}_2\mathbb{R}}$ we see that arbitrary values $k \in \mathbb{R}$ may appear.

Explicit representations of the universal covering group $\widetilde{\mathrm{SL}_2\mathbb{R}}$ were first constructed in detail by Pukánszky [Puk64], later elaborated on by Sally [Sal67].

In particular for the *positive discrete series* $U^+_{k_0}$ of the covering groups we find that the ground state eigenvalue k_0 of $-iE$, the minimal weight of the representation $U^+_{k_0}$, is given by $k_0 \in (1/n)\mathbb{Z}^+$ for the n-fold covering group G_n of $\mathrm{SL}_2\mathbb{R}$. The universal cover of $\mathrm{SL}_2\mathbb{R}$ corresponds to the limit $n \to \infty$ for which $k_0 \in \mathbb{R}^+$ (figure 6.14 gives an overview). The spectrum of the operator $-iE$ is, as before:

$$\sigma^+_{k_0}(-iE) = \{2k + k_0 : k \in \mathbb{Z}\} . \tag{6.81}$$

Once the ground state eigenvalue k_0 is fixed by choice of a specific representation $U^+_{k_0}$, all higher eigenvalues of $-iE$ simply differ by a value of 2.

So, for $\mathrm{SO}^\uparrow(1,2)$ the first few possible spectra of $-iE$ are:

$$\begin{aligned}
\sigma^+_2(-iE) &= \{2,\ 4,\ 6,\ 8,\ \ldots\} , \\
\sigma^+_4(-iE) &= \{4,\ 6,\ 8,\ 10,\ \ldots\} , \\
\sigma^+_6(-iE) &= \{6,\ 8,\ 10,\ 12,\ \ldots\} , \qquad \text{etc.}
\end{aligned}$$

For the special linear group $\mathrm{SL}_2\mathbb{R}$, on the other hand, they are:

$$\begin{aligned}\sigma_1^+(-iE) &= \{1,\ 3,\ 5,\ 7,\ \ldots\}\,,\\ \sigma_2^+(-iE) &= \{2,\ 4,\ 6,\ 8,\ \ldots\}\,,\\ \sigma_3^+(-iE) &= \{3,\ 5,\ 7,\ 9,\ \ldots\}\,, \qquad \text{etc.}\,,\end{aligned}$$

and for the metaplectic group $\mathrm{Mp}_2\mathbb{R}$ the possibilities are:

$$\begin{aligned}\sigma_{1/2}^+(-iE) &= \{\tfrac{1}{2},\ \tfrac{5}{2},\ \tfrac{9}{2},\ \tfrac{13}{2},\ \ldots\}\,,\\ \sigma_1^+(-iE) &= \{1,\ 3,\ 5,\ 7,\ \ldots\}\,,\\ \sigma_{3/2}^+(-iE) &= \{\tfrac{3}{2},\ \tfrac{7}{2},\ \tfrac{11}{2},\ \tfrac{15}{2}\,,\ \ldots\}\,, \qquad \text{etc.}\end{aligned}$$

Example 6.4.10 A spectrum of $-iE$ that starts with $k_0 = 1$, for instance, belongs to the positive discrete series representation U_1^+. This representation is a faithful irreducible unitary representation of $\mathrm{SL}_2\mathbb{R}$ and a (unfaithful) representation of $\mathrm{Mp}_2\mathbb{R}$ and all higher covering groups of $\mathrm{SL}_2\mathbb{R}$, but not an irreducible unitary representation of $\mathrm{SO}^\uparrow(1,2)$.

If we find a spectrum of $-iE$ that starts with $k_0 = 3/2$, on the other hand, we can identify the associated representation $U_{3/2}^+$ as a faithful irreducible unitary representation of the metaplectic group $\mathrm{Mp}_2\mathbb{R}$. It is also a unfaithful representation of all proper covering groups of $\mathrm{Mp}_2\mathbb{R}$. Nevertheless, $U_{3/2}^+$ is no irreducible unitary representations of $\mathrm{SO}^\uparrow(1,2)$, nor of $\mathrm{SL}_2\mathbb{R}$.

Remark 6.4.11 (Casimir operator) The formula $\lambda = -k_0(k_0-2)$ for the eigenvalue λ of the Casimir operator remains correct for positive discrete series representations of the covering groups. An interesting observation, however, is that the value λ is the same for the representations $U_{k_0}^+$ and $U_{2-k_0}^+$ (when they are allowed), due to:

$$\lambda^{(2-k_0)} = -(2-k_0)\big((2-k_0)-2\big) = -(2-k_0)(-k_0) = -k_0(k_0-2) = \lambda^{(k_0)}\,.$$

The representations $U_{1/2}^+$ and $U_{3/2}^+$ of the metaplectic group provide the first non-trivial example of this behaviour.

6.4.4 Identification of the Canonical Group

Now that we have classified the irreducible unitary representations of $\mathrm{SO}^\uparrow(1,2)$ and its covering groups, recall how the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ originally appeared. To obtain the specific form (6.50) of the fundamental Lie brackets of $\mathfrak{sl}(2,\mathbb{R})$ we made the following identifications with the classical observables q^2, p^2 and qp (we assume that the units of q^2 and p^2 are chosen so that the constant α in equation (6.52) equals 1):

$$A \mathrel{\hat{=}} -qp\,, \quad N_+ \mathrel{\hat{=}} \tfrac{1}{2}q^2\,, \quad N_- \mathrel{\hat{=}} -\tfrac{1}{2}p^2\,, \tag{6.82}$$

The classical observable associated to the generator Z of the maximal compact subgroup is thus:

$$Z = N_+ - N_- \mathrel{\hat{=}} \tfrac{1}{2}(q^2+p^2)\,. \tag{6.83}$$

This classical quantity is clearly half-bounded from below (it is, after all, the energy of a harmonic oscillator). Therefore, only representations of the *positive discrete series* will play a role when we try to model a particle on a half-line.

On the other hand, section 6.3 showed that the squared momentum operator $\hat{p}^2$ with domain D_∞ according to equation (6.43) is self-adjoint on the half-line. Moreover, we know from section 6.2 that both $\hat{q}^2$ and the symmetrised combination $\frac{1}{2}(\hat{q}\hat{p}+\hat{p}\hat{q})$ (with domains chosen appropriately) are self-adjoint operators when restricted to the half-line. Thus, we have quantized versions of our fundamental observables q^2, p^2 and qp and they give a Lie algebra representation of $\mathfrak{sl}(2,\mathbb{R})$ by self-adjoint operators on the Hilbert space $\mathcal{H}^+ = L^2(\mathbb{R}^+, \mathrm{d}q)$:

$$-i\mathcal{L}U(A) = -\tfrac{1}{2}(\hat{q}\hat{p}+\hat{p}\hat{q})\,, \quad -i\mathcal{L}U(N_+) = \tfrac{1}{2}\hat{q}^2\,, \quad -i\mathcal{L}U(N_-) = -\tfrac{1}{2}\hat{p}^2\,. \tag{6.84}$$

In particular, Z is represented by:

$$-iE = -i\mathcal{L}U(Z) = -i\mathcal{L}U(N_+ - N_-) = \tfrac{1}{2}(\hat{q}^2+\hat{p}^2)\,. \tag{6.85}$$

Remark 6.4.12 Note that the whole purpose of the operator $-iE$ in equation (6.85) is to fix a specific representation when we try to identify the Canonical Group for a particle on a half-line but, in contrast to the situation in [Kas03; Kas07], the operator $-iE$ does *not* have to act as Hamiltonian. Therefore, the Canonical Group obtained below is not specific to physics of a harmonic oscillator but governs *generic* dynamics on a half-line with 'hard wall' boundary condition.

Remark 6.4.13 Although equation (6.85) looks like the well-known harmonic oscillator Hamiltonian on $\mathbb{R}$, observe that it really describes (if used as Hamiltonian) a harmonic oscillator on a half-line. The difference isn't visible in form of a potential but lies hidden in the operator domain. Compared to the operator $\hat{p}^2$ on the whole line which is self-adjoint on the domain $\{\psi \in L^2(\mathbb{R}, \mathrm{d}q) : \psi \in AC^2[-\infty;+\infty]\}$, section 6.3 showed that the operator $\hat{p}^2$ restricted to the half-line requires the boundary condition $\psi(0) = 0$ in addition to $\psi \in AC^2[0;\infty]$ for all $\psi \in D_\infty$. Without the boundary condition, $\hat{p}^2$ isn't self-adjoint on $L^2(\mathbb{R}^+, \mathrm{d}q)$.

To determine the spectrum of E in equation (6.85), on the other hand, we can exploit the fact that the operators $\hat{q}^2$ and $\hat{p}^2$ restricted to $L^2(\mathbb{R}^+, \mathrm{d}q)$ act on wave functions in the same way as the original, unrestricted operators on $L^2(\mathbb{R}, \mathrm{d}q)$. *Only* the domain is different. Therefore, the spectrum $\sigma(-iE)$ consists precisely of the eigenvalues of the conventional harmonic oscillator on the whole line for which the eigenfunctions satisfy the boundary condition $\psi(0) = 0$.

Remark 6.4.14 (Eigenfunctions of the harmonic oscillator on $\mathbb{R}$) The eigenfunctions of the harmonic oscillator Hamiltonian $\frac{1}{2}(\hat{q}^2+\hat{p}^2)$ on $L^2(\mathbb{R}, \mathrm{d}q)$ are well known as (see for example [Sak94, appendix A.4]):

$$\psi_k(q) = \frac{1}{\sqrt{2^k k!}}\,\pi^{-1/4} e^{-q^2/2} H_k(q)\,. \tag{6.86}$$

and they satisfy the well-known eigenvalue equation:

$$\tfrac{1}{2}(\hat{q}^2+\hat{p}^2)|\psi_k\rangle = (k+\tfrac{1}{2})|\psi_k\rangle\,, \quad \text{for } k \in \mathbb{Z}_0^+\,. \tag{6.87}$$

The functions H_k in equation (6.86) are the Hermite polynomials:

$$H_k(q) = (-1)^k e^{q^2} (\partial_q)^k e^{-q^2}\,. \tag{6.88}$$

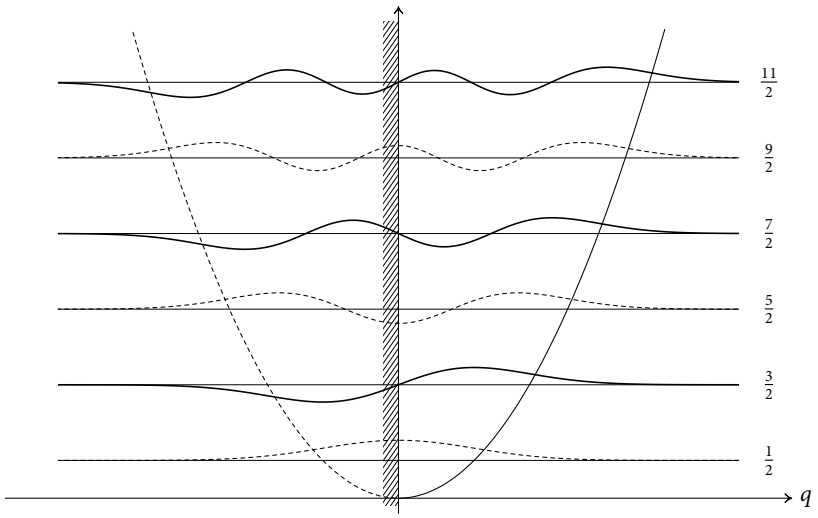


Figure 6.15 Solutions for the harmonic oscillator on a half-line $\mathbb{R}^+$ are, due to the additional boundary condition $\psi(0) = 0$, exactly the odd eigenfunctions of the regular harmonic oscillator on the whole line $\mathbb{R}$.

Explicitly, the first few Hermite polynomials are:

$$\begin{aligned} H_0(q) &= 1\,, & H_1(q) &= 2q\,,\\ H_2(q) &= 4q^2 - 2\,, & H_3(q) &= 8q^3 - 12q\,,\\ H_4(q) &= 16q^4 - 48q^2 + 12\,, & H_5(q) &= 32q^5 - 160q^3 + 120q\,. \end{aligned}$$

We see that $H_k(0) = 0$ whenever k is *odd*, whereas $H_k(0) \neq 0$ when k is *even*.

It follows from remark 6.4.14 that the eigenfunctions of the operator $-iE$ above are exactly the ψ_n from equation (6.86) for *odd* values of n (see also figure 6.15). Hence, the spectrum of $-iE$ is:

$$\sigma(-iE) = \{2k + \tfrac{3}{2} : k \in \mathbb{Z}_0^+\} = \sigma^+_{3/2}(-iE)\,. \tag{6.89}$$

Due to half-integer weights, this representation cannot arise if the Canonical Group is equal to $\mathrm{PSL}_2\mathbb{R}$ or $\mathrm{SL}_2\mathbb{R}$. Nevertheless, $U^+_{3/2}$ is an irreducible unitary representation of the metaplectic group $\mathrm{Mp}_2\mathbb{R}$. Hence we can state as result:

> The simplest Canonical Group $\mathcal{C}$ on the classical phase space $\mathbb{R}^2/\mathbb{Z}_2$ that is able to reproduce the squared momentum operator $\hat{p}^2$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ with domain D_∞ as in equation (6.43) is the **metaplectic group** $\mathcal{C} = \mathrm{Mp}_2\mathbb{R}$, the 4-fold covering group of $\mathrm{SO}^\uparrow(1,2)$ and the 2-fold covering group of $\mathrm{Sp}_2\mathbb{R}$. In particular, it is the only Canonical Group $\mathcal{C}$ for which $U^+_{3/2}$ is a *faithful* representation.

The existence of a suitable Canonical Group proves that using the group-theoretical quantization method it is possible to obtain a viable quantum theory for a particle on a half-line including the reflection at the end *as if* by a 'hard wall' potential but without need for the potential, nor for the negative part of the real line on which this potential would have be defined. Instead, this toy model of a system with contact interaction can

be quantized *directly*, over the restricted configuration space $\mathbb{R}^+$, if the boundary condition is realised on the classical level as *topological* feature of the phase space.

6.4.5 The Harmonic Oscillator and the Metaplectic Representation

To understand the relation between the one-dimensional harmonic oscillator on the whole line and the oscillator on the half-line properly, we will end this chapter with a few words on the metaplectic representation. The metaplectic representation of $\mathrm{Mp}_2\mathbb{R}$ in relation to the harmonic oscillator has been mentioned by Kastrup [Kas07, section 3.5] but an earlier and more thorough treatment can be found in [Tay86, chapters 1 and 11]. Furthermore, the metaplectic representation of $\mathrm{Mp}_2\mathbb{R}$ is closely related to the projective *Segal–Shale–Weil representation* of $\mathrm{Sp}_2\mathbb{R}$ [Wei64] which is famous for its application to modular forms and θ-series [LV80].

The problem that will lead us to the metaplectic representation reveals itself when we think about the well-known spectrum of the harmonic oscillator Hamiltonian over the whole line, $Q = \mathbb{R}$:

$$\sigma_{\mathrm{HO}} := \sigma\big(\tfrac{1}{2}(\hat{q}^2 + \hat{p}^2)\big) = \{k + \tfrac{1}{2} : k \in \mathbb{Z}_0^+\} \,. \tag{6.90}$$

The spectrum σ_{HO} looks exactly like the spectrum $\sigma_1^+(-iE) = \{2k + 1 : k \in \mathbb{Z}_0^+\}$ scaled by a factor of $1/2$. The following identification therefore seems tempting:

$$-iE \stackrel{?}{=} \hat{q}^2 + \hat{p}^2 \quad \Rightarrow \quad \sigma_{\mathrm{HO}} \stackrel{?}{=} \sigma_1^+(-\tfrac{1}{2}iE) \,. \tag{6.91}$$

Nevertheless, you will notice eventually that the unrestricted operators $\hat{q}^2$ and $\hat{p}^2$ on the Hilbert space $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$ over $\mathbb{R}$ have the same commutation relations as the restricted operators $\hat{q}^2$ and $\hat{p}^2$ on the Hilbert space $\mathcal{H}^+ = L^2(\mathbb{R}^+, \mathrm{d}q)$ over the half-line. Thus, we have to identify $-iE$ with $\frac{1}{2}(\hat{q}^2 + \hat{p}^2)$ for the regular harmonic oscillator on the whole line as well, similar to equation (6.85), which contradicts equation (6.91).

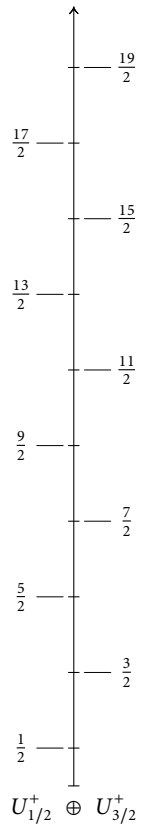
The answer to this riddle is that the spectrum σ_{HO} of the harmonic oscillator does *not* come from an irreducible unitary representation of $\mathrm{Sp}_2\mathbb{R}$ or any of the covering groups. The spectrum must rather be written as:

$$\sigma_{\mathrm{HO}} = \sigma_{1/2}^+(-iE) \cup \sigma_{3/2}^+(-iE) \,, \quad -iE = \tfrac{1}{2}(\hat{q}^2 + \hat{p}^2) \,, \tag{6.92}$$

and belongs to the *reducible* representation $U_{\mathrm{mp}} := U_{1/2}^+ \oplus U_{3/2}^+$ of $\mathrm{Mp}_2\mathbb{R}$. This is quite interesting in the light of remark 6.4.11 because the Casimir operator is represented by the same value $\lambda = 3/4$ for the representations $U_{1/2}^+$ and $U_{3/2}^+$. The representation U_{mp} of $\mathrm{Mp}_2\mathbb{R}$ on the Hilbert space $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$, with realisations as in equation (6.84) yet without restriction to the half-line, is known as ***metaplectic representation*** of the group $\mathrm{Mp}_2\mathbb{R}$ [Tay86, section 11.3]. It becomes obvious that the spectrum σ_{HO} of the harmonic oscillator must belong to a *reducible* representation of $\mathrm{Mp}_2\mathbb{R}$ once one realises that the ladder operators $R_\pm$ on $L^2(\mathbb{R}, \mathrm{d}q)$ are *quadratic* polynomials in $\hat{q}$ and $\hat{p}$ and therefore cannot change the parity of wave functions. The Hilbert space $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$ hence splits into a direct sum:

$$\mathcal{H} = \mathcal{H}^{\mathrm{e}} \oplus \mathcal{H}^{\mathrm{o}} \,, \tag{6.93}$$

for which the subspaces $\mathcal{H}^{\mathrm{e}}$ and $\mathcal{H}^{\mathrm{o}}$ of wave functions with even/odd parity are invariant under the metaplectic representation U_{mp}. The eigenfunctions ψ_k of the conventional harmonic oscillator over $\mathbb{R}$, however, alternate between even and odd parity.



The spectra show that $U^+_{1/2}$ becomes an irreducible representation on the subspace $\mathcal{H}^{\mathrm{e}} \subset \mathcal{H}$, whereas $U^+_{3/2}$ is irreducible on the subspace $\mathcal{H}^{\mathrm{o}} \subset \mathcal{H}$. We can, furthermore, restrict these representations to representations on $\mathcal{H}^+ = L^2(\mathbb{R}^+, \mathrm{d}q)$ because values $\psi^{\mathrm{e/o}}(-q)$ of even and odd wave functions on the negative half-line are determined by $\psi^{\mathrm{e/o}}(-q) = \pm\psi^{\mathrm{e/o}}(q)$ for positive $q \in \mathbb{R}^+$. The boundary condition that appears in the domain D_∞ of the operator $\hat{p}^2$ on the half-line then is a result of the fact that odd wave functions $\psi \in \mathcal{H}^{\mathrm{o}}$ satisfy $\psi(0) = 0$. This is the promised relation between the harmonic oscillator on the whole line and the oscillator on the half-line:

> The harmonic oscillator on the half-line is in a very direct sense 'half of' the oscillator on the whole line.

Remark 6.4.15 (Segal–Shale–Weil representation) The relation to the *Segal–Shale–Weil representation* of $\mathrm{Sp}_2\mathbb{R}$ is that the metaplectic representation of $\mathrm{Mp}_2\mathbb{R}$ projects down to a representation of $\mathrm{Sp}_2\mathbb{R}$ on $PU(\mathcal{H})$:

$$\begin{array}{ccc} \mathrm{Mp}_2\mathbb{R} & \xrightarrow{U^+_{1/2} \oplus U^+_{3/2}} & U(\mathcal{H}) \\ \downarrow & \Downarrow & \downarrow \\ \mathrm{Sp}_2\mathbb{R} & \longrightarrow & PU(\mathcal{H}) \end{array} .$$

Nevertheless, we have $2\pi E|k\rangle = \pi i(\hat{q}^2 + \hat{p}^2)|k\rangle = \pi ik|k\rangle$ for $k \in \{1, 3, 5, \ldots\}$, and hence $e^{2\pi E}|k\rangle = e^{\pi ik}|k\rangle = -|k\rangle$ [Tay86, section 11.3]. It follows that $\exp(2\pi Z) = e$ (in $\mathrm{Sp}_2\mathbb{R}$) is represented by $e^{2\pi E} = -\mathbb{1}$. Therefore, the Segal–Shale–Weil representation of $\mathrm{Sp}_2\mathbb{R}$ is only a *projective* representation on $PU(\mathcal{H}) = U(\mathcal{H})/\mathbb{C}^\times$ (where $-\mathbb{1} \sim +\mathbb{1}$) but does not lift to a 'true' representation of $\mathrm{Sp}_2\mathbb{R}$ on $U(\mathcal{H})$ [LV80, sections 1.6–1.7]. This matches our discussion of the possible representations of covering groups in section 6.4.3.

Finally, a different way to approach the metaplectic representation is in terms of the ladder operators a and $a^\dagger$, and the number operator $N = a^\dagger a$ of the regular harmonic oscillator over $\mathbb{R}$ [Sak94, section 2.3]:

$$a = \frac{1}{\sqrt{2}}(\hat{q} + i\hat{p}) , \quad a^\dagger = \frac{1}{\sqrt{2}}(\hat{q} - i\hat{p}) , \quad N = \frac{1}{2}(\hat{q}^2 + \hat{p}^2) - \frac{1}{2} . \tag{6.94}$$

These operators satisfy the well-known commutation relations:

$$[a, a^\dagger] = \mathbb{1} , \quad [N, a] = -a , \quad [N, a^\dagger] = a^\dagger , \tag{6.95}$$

and the harmonic oscillator Hamiltonian reads:

$$\hat{H} = \tfrac{1}{2}(\hat{q}^2 + \hat{p}^2) = N + \tfrac{1}{2} . \tag{6.96}$$

Equation (6.95) shows that the ladder operators a and $a^\dagger$ belong to to the Lie algebra of the *Heisenberg group*, not to the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$, and this is the reason why the identification (6.91) can't explain the spectrum σ_{HO}.

On the other hand, we can rewrite the Hamiltonian $\hat{H}$ in (6.96) as:

$$\hat{H} = N + \tfrac{1}{2} = a^\dagger a + \tfrac{1}{2} = \tfrac{1}{2}(a^\dagger a + aa^\dagger) , \tag{6.97}$$

and the commutation relations of $\hat{H}$ and the *squared* operators a^2 and $a^{\dagger 2}$ are:

$$[a^2, a^{\dagger 2}] = 4\hat{H} , \quad [H, a^2] = -2a^2 , \quad [H, a^{\dagger 2}] = 2a^{\dagger 2} . \tag{6.98}$$

Moreover, if we now define:

$$-iE := \hat{H}\,, \quad -iR_+ := \hat{R}_+ := \tfrac{1}{2}a^{\dagger 2}\,, \quad -iR_- := \hat{R}_- := \tfrac{1}{2}a^2\,, \tag{6.99}$$

these operators satisfy the same commutation relations as the equally named operators introduced in equation (6.60):

$$\big[E, R_\pm\big] = \pm 2iR_\pm\,, \quad \big[R_-, R_+\big] = iE\,.$$

Accordingly, the operators $R_\pm$ in equation (6.99) are precisely the raising and lowering operators of $\mathfrak{sl}(2,\mathbb{R})$ introduced before.

The relation (6.99) between the two different kinds of ladder operators is shown in figure 6.16 and explains the spectrum of the harmonic oscillator on the whole line and its relation to the spectrum of the harmonic oscillator on the half-line quite nicely:

- Over the whole line, the ladder operators a and $a^\dagger$ are available, in *addition* to the operators $\hat{R}_+$ and $\hat{R}_-$, because the phase space $T^*\mathbb{R}$ over $\mathbb{R}$ admits an action of the Heisenberg group $H(1)$. The operators a and $a^\dagger$ then 'connect' the two irreducible representations $U^+_{1/2}$ and $U^+_{3/2}$ of the metaplectic group $\mathrm{Mp}_2\mathbb{R}$ so that their direct sum becomes necessary to explain the spectrum of the harmonic oscillator over the whole line.
- The phase space $\mathbb{R}^2/\mathbb{Z}_2$, on the other hand, is incompatible with an action of the Heisenberg group, which is reasonable because the restriction to the half-line breaks the translation subgroup in the q-coordinate. The raising and lowering operators $\hat{R}_\pm$ associated to the metaplectic group are still available; however, the two irreducible components of the metaplectic representation of $\mathrm{Mp}_2\mathbb{R}$ are no longer connected. Thus, only *one* of the representations, either $U^+_{1/2}$ *or* $U^+_{3/2}$, can (and does) appear.

In this thesis, we found that the irreducible representation $U^+_{3/2}$ of $\mathrm{Mp}_2\mathbb{R}$ contains the operator $\hat{p}^2$ restricted to the half-line $\mathbb{R}^+ \subset \mathbb{R}$ with domain $D_\infty \subset L^2(\mathbb{R}^+, \mathrm{d}q)$. Since

The representation (6.99) has applications in quantum optics and is known as *single-mode 'two boson' realisation* of the Lie algebra $\mathfrak{su}(1,1)$ in this case [Per86; Ger91; Ger05; Kas03].

It is also noteworthy how well the realisation (6.99) plays with the 'mirror particle' model of motion on the half-line that we mentioned in remark 6.3.5.

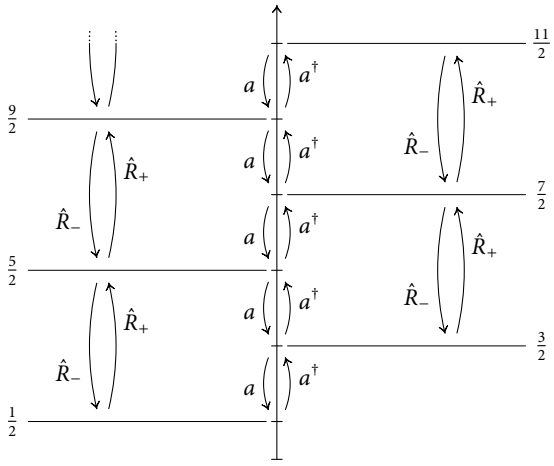


Figure 6.16 The two kinds of ladder operators. On the Hilbert space $L^2(\mathbb{R}, \mathrm{d}q)$ over the whole line, the operators a and $a^\dagger$ 'connect' the representations $U^+_{1/2}$ and $U^+_{3/2}$ of $\mathrm{Mp}_2\mathbb{R}$ but themselves belong to the Lie algebra of the Heisenberg group. Over the half-line, on the other hand, only the operators $\hat{R}_\pm$ are well-defined and $U^+_{3/2}$ is the representation that contains the operator $\hat{p}^2$ on the domain D_∞.

the domain D_∞ includes the boundary condition $\psi(0) = 0$ for wave functions $\psi \in D_\infty$, this is the representation that must be used to describe motion on the half-line where a particle gets reflected by a hard wall potential at the origin. Nevertheless, notice again that the potential *isn't needed* to arrive at correct results – the obtained quantum theory is self-consistent on the *subspace* $L^2(\mathbb{R}^+, \mathrm{d}q) \subset L^2(\mathbb{R}, \mathrm{d}q)$ and the boundary condition $\psi(0) = 0$ is *automatically* fulfilled by all Hamiltonians that can be written down using the fundamental quantum operators of this theory.

The representation $U^+_{1/2}$, on the other hand, while associated to the same Canonical Group $\mathrm{Mp}_2\mathbb{R}$, does not describe a particle on the half-line when we require this model to be equivalent to a particle confined to $\mathbb{R}^+$ by a hard wall potential. Nevertheless, the representation may still be relevant for models where $\mathbb{R}^+$ appears in a different physical context, say as a topological factor of some higher-dimensional configuration space. The half-line $\mathbb{R}^+$ might, for example, be the radius in some centrally symmetric system.

Summary and Conclusions

> Nature uses only the longest threads to weave her patterns, so each small piece of her fabric reveals the organization of the entire tapestry.
>
> *Richard P. Feynman* [Fey65, chapter 1]

In this work we studied the direct quantization of non-relativistic classical systems using a group-theoretical scheme. Although phase spaces different from $T^*\mathbb{R}^n$ pose no problem in Hamiltonian mechanics, the same cannot be said for quantum mechanics. Since global and topological aspects play a much bigger role in quantum theory than in the associated classical theory, it happens frequently that certain, sometimes even basic, classical observables have no appropriate quantum counterparts. As application of the Canonical Group Quantization method we explored the relation between classical and quantum models for a particle on a half-line with 'hard wall' boundary condition.

The first chapter started with a review of Dirac's canonical quantization programme and its formalisation in terms of the Dirac quantization map. Nevertheless, the famous theorem of Groenewold and Van Hove exposes such a 'full' quantization as impossible: the supposedly natural assumptions are inconsistent. Some of the assumptions have a physical motivation but others appear to be adopted just for the benefit of mathematical convenience. The main goal of this chapter was to understand how a physically justified quantization procedure should look like – what it can, and what it can *not* accomplish. In contrast to a widespread belief, the nature of quantization is only one of *similarity* between classical and quantum theory, *not* the inverse of a physically realised classical limit. As a consequence, there is no physical reason for a quantization method to produce a unique result, nor does it make much sense to require arbitrary classical systems, or arbitrary observables of these systems, to be quantizable.

Also in the first chapter we discussed some mathematical points in quantum theory where standard physics textbooks are insufficient to our purposes, most important the subtle difference between *symmetric* and *self-adjoint* operators. The distinction rests in the operator *domains* and domains are typically ignored in calculations. We explained why this simplification is acceptable in quantum mechanics over the trivial configuration space $\mathbb{R}^n$ but no longer adequate in case of non-trivial spaces: due to Stone's theorem, the existence of self-adjoint (but not of symmetric) operators is heavily influenced by global and topological aspects. The momentum operator $\hat{p} = -i\hbar\partial_q$, for example, can be restricted to a symmetric operator on the half-line $\mathbb{R}^+ \subset \mathbb{R}$ but there exists no domain on which the restricted operator were self-adjoint.

The proper way to handle the difficulties with domains of self-adjoint operators is in terms of corresponding continuous groups of unitary transformations. Stone's theorem asserts that any self-adjoint operator generates a strongly continuous group of unitary transformations and vice versa. This is up to a certain point analogous to the situation in classical mechanics where observables, by means of Hamiltonian vector fields, act as generators of finite transformations. There exist classical observables, however, for which the associated Hamiltonian vector fields are incomplete, so that they generate only *pseudo*groups, not groups, of symplectic transformations. Since self-adjoint operators, on the other hand, always generate *groups* of unitary transformations, such observables have no proper quantum counterparts. The Canonical Group Quantization method is a

constructive and physically transparent technique to select from the outset only classical observables that generate *groups* of transformations. A set of 'fundamental' observables then fixes a quantization map. In contrast to a quantization scheme that works only on the level of infinitesimal generators, the use of a 'quantization group' ensures that global and topological properties are preserved.

Continuing, in chapters 2 to 4 we reviewed some of the mathematical background required to understand Isham's quantization procedure. In chapter 2 we gave a review of basic differential geometry using the calculus of differential forms. Chapter 3 dealt with geometric aspects of Lie groups, the relation between Lie groups and Lie algebras via Lie functor and exponential map, and group actions on manifolds. We also included some notes on semidirect products and group extensions in this chapter. Afterwards, chapter 4 gave an introduction to symplectic manifolds and the geometric formulation of Hamiltonian mechanics. In particular, we discussed the canonical symplectic structure of cotangent bundles and two natural constructions for symplectic group actions in this case.

Chapters 5 and 6 contain the main achievements of this dissertation. In chapter 5 we studied Isham's Canonical Group Quantization technique in detail and were able to sort out some points left unclear by Isham's presentation in [Ish83]. The construction of the Canonical Group $\mathcal{C}$ via a central Lie algebra extension in the exceptional case for which the group $\mathcal{G}$ itself 'doesn't work' appears as a kind of last resort in Isham's work. The motivation given for the specific use of a central extension, in particular, is merely that a supposedly more natural form of 'enlargening' fails, yet the actual significance of the central extension from a physical point of view was left unclear. As a result, the construction of the Canonical Group seemed conceptually rather unsatisfying in this special case, suggestively called 'pathological' by Isham.

In contrast to Isham's display, a cornerstone of our analysis was a sharp distinction between *Geometric Group* $\mathcal{G}$ and *Canonical Group* $\mathcal{C}$ right from the outset. An advantage of our perspective is that both groups have a clear conceptual meaning:

- the Geometric Group $\mathcal{G}$ is a $\mathcal{G}$-space structure of the classical phase space,
- the Canonical Group $\mathcal{C}$ is associated directly to a set of preferred observables.

This allowed us to do a proper study of the *relation* between $\mathcal{C}$ and $\mathcal{G}$, and the result was a much clearer and more satisfying construction of the Canonical Group. Following our approach, the relation between the Lie algebras $\mathfrak{LC}$ and $\mathfrak{LG}$ via a central extension appears naturally and inevitably, mimicking the relation between classical observables and Hamiltonian vector fields. In fact, we found that it is *always* possible to define $\mathfrak{LC}$ by means of a central extension, also if the extension is not strictly required (although a central extension will in this case be a trivial extension, isomorphic to a direct sum of Lie algebras). On the other hand, we proved that the construction of $\mathfrak{LC}$ via a *non-trivial* central extension is strictly necessary when canonically conjugate observables are among the fundamental observables.

In chapter 6 we used the quantization procedure to examine the relation between classical and quantum models for a particle moving on a half-line with emphasis on the boundary condition. As guiding principle, an acceptable quantum theory for a particle on a half-line should be physically equivalent to the theory of a particle restricted to the spatial region $\mathbb{R}^+ \subset \mathbb{R}$ by means of a 'hard wall' potential. The value of this toy model is that it allows for a detailed study of the boundary condition at the origin under controlled circumstances – a simple example of a contact interaction. Our question was whether a consistent quantum theory over $\mathbb{R}^+$ that realises the boundary condition implicitly could

be obtained directly from a corresponding classical model with help of the quantization method. Since Isham claimed to have constructed a quantum theory over the half-line in [Ish83, section 4.5], the chapter was split roughly into two parts:

In the first half of chapter 6 we discussed Isham's earlier quantization attempt. The 'obvious' classical phase space of a particle on the half-line $\mathbb{R}^+$ is the cotangent bundle $T^*\mathbb{R}^+$. Starting with $T^*\mathbb{R}^+$, the group $\mathbb{R} \rtimes \mathbb{R}^+$ leads to a quantum theory over $\mathbb{R}^+$ based on the fundamental observables q and qp (the momentum itself is not quantizable over the half-line!). Although formally correct, some points raised doubts about the physical interpretation as quantum theory over a half-line:

- Isham's self-adjoint representation of the quantum operators constructed in analogy to the position space representation of standard quantum mechanics requires a scale-invariant measure, $\mathrm{d}q/q$, on $\mathbb{R}^+$ and this measure conflicts with the Euclidean metric induced on $\mathbb{R}^+$ via the embedding $\mathbb{R}^+ \subset \mathbb{R}$ as half of the whole line. Thus, the interpretation of $\mathbb{R}^+$ as a half-line became questionable. In fact, a much better fitting interpretation appeared to be as quantum theory for a particle moving on the whole line with an unusual choice of coordinates.
- The obtained quantum theory on $\mathbb{R}^+$ is scale-invariant in the q-coordinate so that it is impossible to fix a length scale. Physical applications are limited to systems which exhibit such a scale-invariance but we would not expect this kind of symmetry to hold in general for a particle moving on a half-line.

We were able to sort out the first problem by an explicit construction of the proper position space representation on the Hilbert space $\mathcal{H}^+ = L^2(\mathbb{R}^+, \mathrm{d}q)$ equipped with the usual Lebesgue measure. To do this, we proved that the operators $\hat{q}$ and $\hat{\pi}$ of standard quantum mechanics on the Hilbert space $\mathcal{H} = L^2(\mathbb{R}, \mathrm{d}q)$ over the whole line that are associated to the classical observables q and $\pi = qp$ can be restricted to self-adjoint operators on $\mathcal{H}^+$ in the strict mathematical sense. Given this result, we could show that Isham's representation is unitary equivalent to the proper position space representation. Such a unitary equivalence exists because only a *G-quasi*-invariant, not a *G*-invariant, measure is required. This clarifies the relation between Isham's quantum theory on $\mathbb{R}^+$ and conventional quantum mechanics restricted to the spatial region $\mathbb{R}^+ \subset \mathbb{R}$ by means of a 'hard wall' potential. What remains, however, is that the constructed quantum theory on $\mathbb{R}^+$ is severely restricted with respect to possible applications to the half-line.

Given the limited value this prior result with respect to dynamics on a half-line, the central question of the second half of chapter 6 was whether a 'better' quantum theory over $\mathbb{R}^+$ can be obtained. To this end, we considered the problem from its quantum-mechanical side and found that the squared momentum operator $\hat{p}^2$ (but not $\hat{p}$ itself!) can be made self-adjoint on the Hilbert space $\mathcal{H}^+$ over $\mathbb{R}^+$ by choice of a suitable domain. In particular, there exists a domain $D_\infty \subset \mathcal{H}^+$ which requires the Dirichlet boundary condition $\psi(0) = 0$. The Hamiltonian $\hat{p}^2/(2m)$ on this domain describes a reflection as if by a 'hard wall' potential at the origin. Nevertheless, in contrast to standard quantum mechanics, the boundary condition is an intrinsic feature of the theory and does not have to be enforced by an explicit potential. In fact, the operator $\hat{p}^2$ is no longer self-adjoint over the half-line if the boundary condition is omitted! We showed that Isham's quantum theory on the half-line is unable to explain this dynamical aspect and hence fails with respect to our 'guiding principle': the quantization scheme applied to the cotangent bundle $T^*\mathbb{R}^+$ yields an incomplete quantum theory that is not equivalent to quantum mechanics restricted to $\mathbb{R}^+$ by a 'hard wall' potential. As Isham's main objective in [Ish83] is quantum gravity for which $\mathbb{R}^+$ can be interpreted in a different manner,

this problem did probably not occur to him. Nevertheless, the contact interaction is quite interesting in its own right, which justifies a further examination.

The problem that prevents inclusion of the contact interaction turned out to be the 'obvious' classical phase space $T^*\mathbb{R}^+$. While the squared momentum p^2 as observable on the cotangent bundle $T^*\mathbb{R}^+$ is inherently not quantizable because the Hamiltonian vector field X_{p^2} on $T^*\mathbb{R}^+$ is incomplete, we could show that this is no shortcoming but actually a *benefit* of the quantization method. Based on the negative result that X_{p^2} is incomplete on $T^*\mathbb{R}^+$ we constructed a different classical phase space for the particle on a half-line on which X_{p^2} becomes complete: the orbifold $\mathbb{R}^2/\mathbb{Z}_2$. This phase space emerges naturally upon closer examination and realises the reflection as a topological feature.

Starting with the classical observables q^2, p^2 and qp on $\mathbb{R}^2/\mathbb{Z}_2$ which generate the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ we were able to reverse-engineer the specific Canonical Group that yields the desired operator $\hat{p}^2$ on $L^2(\mathbb{R}^+, \mathrm{d}q)$ mentioned above. The appropriate group is the *metaplectic group* $\mathrm{Mp}_2\mathbb{R}$, a 2-fold cover of the symplectic group $\mathrm{Sp}_2\mathbb{R} \cong \mathrm{SL}_2\mathbb{R}$, and the specific representation of $\mathrm{Mp}_2\mathbb{R}$ that describes a particle on the half-line with reflection at the end is the irreducible unitary representation $U^+_{3/2}$ of the *positive discrete series*. Furthermore, we established the relation between our result and the metaplectic representation of $\mathrm{Mp}_2\mathbb{R}$ on the Hilbert space $L^2(\mathbb{R}, \mathrm{d}q)$ which governs the harmonic oscillator in conventional quantum mechanics over the whole line.

The results of sections 6.3 and 6.4 show that a quantum theory over the half-line that includes the reflection at the end as an intrinsic feature can be constructed by direct quantization using the group-theoretical approach. In contrast to the view advertised for example in [GK04], it is not necessary to include the rest of the line to obtain a physically satisfying description; the restriction to the half-line due to the contact interaction with a hard wall can be realised already classically. Whereas the reflection in quantum theory, however, takes on the form of a boundary condition imposed on operators, the reflection on the classical level must be realised as a topological feature of the phase space. In fact, the group-theoretical approach enforces consistency between the global behaviour of the classical and the associated quantum theory: a closer look reveals that classical mechanics on the cotangent bundle $T^*\mathbb{R}^+$ does not support the reflection we aimed for and it is therefore actually *desirable* that the method fails in this case. A publication of these results is in preparation [JHP12].

In order to generalise the observations made in the present special case it would be interesting to examine systems for which $\mathbb{R}^+$ appears as part of a higher-dimensional configuration space, for example as radius $r \in \mathbb{R}^+$ of a spherically symmetric problem. We expect an analogous construction of the phase space to be possible in these cases. Moreover, there exists a connection to the Riemann hypothesis via the Berry–Keating conjecture [BK99; SRL11; SH11] which relates quantizations of the classical observable qp to zeros of the Riemann ζ function, an unexpected link between physics and pure mathematics.

Bibliography

Quotations in this thesis usually refer to the *e-print* versions of references, where available. The online versions of [Kas03; Kas06; Kas07], in particular, are more detailed than the print versions.

[AM78] Ralph Abraham and Jerrold E. Marsden: *Foundations of Mechanics*. 2nd edition. Revised, enlarged, and reset. The Benjamin/Cummings Publishing Company, 1978.

[AMR88] Ralph Abraham, Jerrold E. Marsden and Tudor S. Ratiu: *Manifolds, Tensor Analysis, and Applications*. 2nd edition. Applied Mathematical Sciences 75. Springer, New York, 1988.

[Arn86] Vladimir I. Arnol'd: 'First Steps in Symplectic Topology'. In *Russian Mathematical Surveys* 41.6 (1986), pages 1–20.

[Ash10] John Ashmead: *Morlet wavelets in quantum mechanics*. 2010. arXiv:1001.0250 [quant-ph].

[Bar47] Valentine Bargmann: 'Irreducible Unitary Representations of the Lorentz Group'. In *Annals of Mathematics* 48.3 (1947), pages 568–640.

[Bel04] John S. Bell: *Speakable and Unspeakable in Quantum Mechanics*. 2nd edition. Cambridge University Press, 2004.

[Bel82] John S. Bell: 'On the Impossible Pilot Wave'. In *Foundations of Physics* 12.10 (1982), pages 989–999.

[BFV01] Guy Bonneau, Jacques Faraut and Galliano Valent: 'Self-adjoint extensions of operators and the teaching of quantum mechanics'. In *American Journal of Physics* 69.3 (2001), pages 322–332.

[BHJ26] Max Born, Werner Heisenberg and Pascual Jordan: 'Zur Quantenmechanik II'. In *Zeitschrift für Physik* 35.8–9 (1926), pages 557–615.

[BJ25] Max Born and Pascual Jordan: 'Zur Quantenmechanik'. In *Zeitschrift für Physik* 34.1 (1925), pages 858–888.

[BK99] Michael V. Berry and Jonathan P. Keating: '$H = xp$ and the Riemann Zeros'. In. *Supersymmetry and Trace Formulae: Chaos and Disorder*. Edited by Igor V. Lerner, Jonathan P. Keating and David E. Khmelnitskii. NATO Science Series B: Physics 370. Springer, 1999.

[Bla92] Matthias Blau: 'Symplectic Geometry and Geometric Quantization'. Lecture Notes. 1992. http://www.blau.itp.unibe.ch/Lecturenotes.html.

[Boh52a] David Bohm: 'A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I'. In *Physical Review* 85.2 (1952), pages 166–179.

[Boh52b] David Bohm: 'A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. II'. In *Physical Review* 85.2 (1952), pages 180–193.

[Boj+00] Martin Bojowald et al.: 'Group Theoretical Quantization of a Phase Space $S^1 \times \mathbb{R}^+$ and the Mass Spectrum of Schwarzschild Black Holes in D space–time dimensions'. In *Physical Review D* 62.4 (2000), page 044026. arXiv:gr-qc/9906105.

[BRL10] Carlos Benavides and Andrés F. Reyes-Lega: 'Canonical Group Quantization, Rotation Generators and Quantum Indistinguishability'. In. *Geometric and Topological Methods for Quantum Field Theory*. Edited by Hernán Ocampo, Eddy Pariguan and Paycha Sylvie. Cambridge University Press, 2010, pages 355–367. arXiv:0806.2449 [math-ph].

[Bro87] Ronald Brown: 'From Groups to Groupoids: a Brief Survey'. In *Bulletin of the London Mathematical Society* 19.2 (1987), pages 113–134.

[BS00] Martin Bojowald and Thomas Strobl: 'Group Theoretical Quantization and the Example of a Phase Space $S^1 \times \mathbb{R}^+$'. In *Journal of Mathematical Physics* 41.5 (2000), pages 2537–2567. arXiv:quant-ph/9908079.

[BW10] Borislav Belchev and Mark Walton: 'On Robin Boundary Conditions and the Morse Potential in Quantum Mechanics'. In *Journal of Physics A* 43.8 (2010), page 085301. arXiv:1002.2139v1 [quant-ph].

[Böh78] Arno Böhm: *The Rigged Hilbert Space and Quantum Mechanics*. Lecture Notes in Physics 78. Springer, 1978.

[CDD77] Yvonne Choquet-Bruhat, Cécile DeWitt-Morette and Margaret Dillard-Bleick: *Analysis, Manifolds and Physics*. North-Holland Publishing Company, Amsterdam, 1977.

[CDL99] Claude Cohen-Tannoudji, Bernard Diu and Franck Laloë: *Quantenmechanik*. 2nd edition. Volume 1. de Gruyter, Berlin, 1999.

[DHS00] Daniel A. Dubin, Mark A. Hennings and Tom B. Smith: *Mathematical Aspects of Weyl Quantization and Phase*. World Scientific, Singapore, 2000.

[Dir25] Paul A. M. Dirac: 'The Fundamental Equations of Quantum Mechanics'. In *Proceedings of the Royal Society of London*. A 109.752 (1925), pages 642–653.

[Dir38] Paul A. M. Dirac: 'The Relation between Mathematics and Physics'. In *Proceedings of the Royal Society of Edinburgh* 59 (1938), pages 122–129. Lecture delivered on February 6, 1939.

[Dir58] Paul A. M. Dirac: *The Principles of Quantum Mechanics*. 4th edition. The International Series of Monographs on Physics. Oxford University Press, 1958.

[DK00] Johannes J. Duistermaat and Johan A. C. Kolk: *Lie Groups*. Universitext. Springer, Berlin, 2000.

[DL03] Chris Doran and Anthony Lasenby: *Geometric Algebra for Physicists*. Cambridge University Press, 2003.

[DR99] Michael A. Doncheski and Richard W. Robinett: 'Anatomy of a Quantum "Bounce"'. In *European Journal of Physics* 20.1 (1999), pages 29–37. arXiv:quant-ph/0307010.

[EG73] Henri Epstein and Vladimir J. Glaser: 'The Role of Locality in Perturbation Theory'. In *Annales de l'institut Henri Poincaré A* 19.3 (1973), pages 211–295.

[Els11] Jürgen Elstrodt: *Maß- und Integrationstheorie*. 7th edition. Springer, Berlin, 2011.

[FCT02] Tamás Fülöp, Taksu Cheon and Izumi Tsutsui: 'Classical Aspects of Quantum Walls in One Dimension'. In *Physical Review A* 66.5 (2002), page 052102. arXiv:quant-ph/0111057.

[Fey65] Richard P. Feynman: *The Character of Physical Law*. MIT Press, 1965.

[FH91] William Fulton and Joe Harris: *Representation Theory – A First Course*. Graduate Texts in Mathematics 129. Springer, New York, 1991.

[Fuc01] Christopher A. Fuchs: 'Quantum Foundations in the Light of Quantum Information'. In *Decoherence and its Implications in Quantum Computation and Information Transfer*. Edited by Tony Gonis and Patrice E. A. Turchi. NATO Science Series. IOS Press, 2001. arXiv:quant-ph/0106166.

[Ger05] Christopher C. Gerry: 'Infinite Statistics and the SU(1, 1) phase operator'. In *Journal of Physics A* 38.14 (2005), pages L213–L217.

[Ger07] Dieter Gernet: 'Ockham's Razor and Its Improper Use'. In *Journal of Scientific Exploration* 21.1 (2007), pages 135–140.

[Ger91] Christopher C. Gerry: 'Correlated Two-Mode SU(1, 1) Coherent States: Nonclassical Properties'. In *Journal of the Optical Society of America B* 8.3 (1991), pages 685–690.

[Gie00] François Gieres: 'Mathematical surprises and Dirac's formalism in quantum mechanics'. In *Reports on Progress in Physics* 63.12 (2000), pages 1893–1931. arXiv:quant-ph/9907069v2.

[Giu08] Gian F. Giucide: *Naturally Speaking: The Naturalness Criterion and Physics at the LHC*. 2008. arXiv:0801.2562v2 [hep-ph].

[GK04] Piotr Garbaczewski and Witold Karwowski: 'Impenetrable Barriers and Canonical Quantization'. In *American Journal of Physics* 72.7 (2004), pages 924–933. arXiv:math-ph/0310023.

[Gol87] Sheldon Goldstein: 'Stochastic Mechanics and Quantum Theory'. In *Journal of Statistical Physics* 47.5–6 (1987), pages 645–667.

[Gro46] Hilbrand J. Groenewold: 'On the Principles of Elementary Quantum Mechanics'. In *Physica* 12.7 (1946), pages 405–460.

[Gro88] Siegfried Großmann: *Funktionalanalysis im Hinblick auf Anwendungen in der Physik*. 4th edition. AULA-Verlag, Wiesbaden, 1988.

[GV64] Israïl M. Gel'fand and Naum Ya. Vilenkin: *Generalized Functions. Volume 4: Applications of Harmonic Analysis*. Academic Press, 1964.

[Hab98] Amber Habib: 'Representations of Reductive Lie Groups'. Lecture Notes. 1998. http://www.oocities.org/a_habib/Math/MRI/.

[Hat02] Allen Hatcher: *Algebraic Topology*. Cambridge University Press, Cambridge, 2002. http://www.math.cornell.edu/~hatcher/AT/ATpage.html.

[HC51] Harish-Chandra: 'On Some Applications of the Universal Enveloping Algebra of a Semisimple Lie Algebra'. In *Transactions of the American Mathematical Society* 70.1 (1951), pages 28–96.

[HC52] Harish-Chandra: 'Plancherel Formula for the 2 × 2 Real Unimodular Group'. In *Proceedings of the National Academy of Sciences of the United States of America* 38.4 (1952), pages 337–342.

[Hei25] Werner Heisenberg: 'Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen'. In *Zeitschrift für Physik* 33.1 (1925), pages 879–893.

[Hei+87] Alfons Heil et al.: 'Scalar Matter Field in a Fixed-Point Compactified Five-Dimensional Kaluza-Klein Theory'. In *Nuclear Physics B* 281.3–4 (1987), pages 426–444.

[Hei+89] Alfons Heil et al.: 'Anomalies From the Point of View of G-Theory'. In *Journal of Geometry and Physics* 6.2 (1989), pages 237–270.

[Hep09] Richard Hepworth: 'Morse Inequalities for Orbifold Cohomology'. In *Algebraic & Geometric Topology* 9.2 (2009), pages 1105–1175. arXiv:0712.2432 [math.AT].

[HS71] Peter J. Hilton and Urs Stammbach: *A Course in Homological Algebra*. Graduate Texts in Mathematics 4. Springer, New York, 1971.

[Hua99] Jing-Song Huang: *Lectures on Representation Theory*. World Scientific, Singapore, 1999.

[Hun10] Christian Hundt: 'Quantisierung mit der kanonischen Gruppe'. Diplomarbeit. Johannes Gutenberg-Universität Mainz, 2010. http://wwwthep.physik.uni-mainz.de.

[Ish83] Christopher J. Isham: 'Topological and global aspects of quantum theory'. In *Les Houches, Session XL: Relativity, Groups and Topology II*. Edited by Bryce S. DeWitt and Raymond Stora. North-Holland Publishing Company, Amsterdam, 1983, pages 1060–1290.

[Iwa49] Kenkichi Iwasawa: 'On Some Types of Topological Groups'. In *Annals of Mathematics* 50.3 (1949), pages 507–558.

[JHP12] Florian Jung, Christian Hundt and Nikolaos Papadopoulos: 'Canonical Group Quantization and the Topological Realisation of "Hard Wall" Boundary Conditions'. preprint to be submitted. 2012.

[Jun06] Florian Jung: 'Geometrische Algebra und die Rolle des Clifford-Produkts in der Klassischen und Quantenmechanik'. Diplomarbeit. Johannes Gutenberg-Universität Mainz, 2006. http://wwwthep.physik.uni-mainz.de.

[Jän01a] Klaus Jänich: *Topologie*. 7th edition. Springer, Berlin, 2001.

[Jän01b] Klaus Jänich: *Vector Analysis*. 1st edition. Springer, Berlin, 2001.

[Jän01c] Klaus Jänich: *Vektoranalysis*. 3rd edition. Springer, Berlin, 2001.

[Kas03] Hans A. Kastrup: 'Quantization of the Optical Phase Space $\mathcal{S}^2 = \{\varphi \bmod 2\pi, I > 0\}$ in Terms of the Group $\mathrm{SO}^{\uparrow}(1,2)$'. In *Fortschritte der Physik* 51.10–11 (2003), pages 975–1134. arXiv:quant-ph/0307069.

[Kas06] Hans A. Kastrup: 'Quantization of the Canonically Conjugate Pair Angle and Orbital Angular Momentum'. In *Physical Review A* 73.5 (2006), page 052104.

[Kas07] Hans A. Kastrup: 'A New Look at the Quantum Mechanics of the Hamonic Oscillator'. In *Annalen der Physik* 16.7–8 (2007), pages 439–528. arXiv:quant-ph/0612032.

[Kel06] Kai Johannes Keller: 'Über die Rolle der Projektiven Geometrie in der Quantenmechanik'. Diplomarbeit. Johannes Gutenberg-Universität Mainz, 2006.

[Kna96] Anthony W. Knapp: *Lie Groups Beyond an Introduction*. Progress in Mathematics 140. Birkhäuser, Boston, 1996.

[KPRL08] Kai Johannes Keller, Nikolaos A. Papadopoulos and Andrés F. Reyes-Lega: 'On the Realization of Symmetries in Quantum Mechanics'. In *Mathematische Semesterberichte* 55.2 (2008), pages 149–160.

[Lan02] Serge Lang: *Algebra*. revised 3rd ed. Graduate Texts in Mathematics 211. Springer, New York, 2002.

[Lan85] Serge Lang: $\mathrm{SL}_2(\mathbb{R})$. Graduate Texts in Mathematics 105. Springer, New York, 1985.

[Lan98] Nicolaas P. Landsman: *Mathematical Topics Between Classical and Quantum Mechanics*. Springer Monographs in Mathematical Physics. Springer, 1998.

[LDM71] Michael G. G. Laidlaw and Cécile DeWitt-Morette: 'Feynman Functional Integrals for Systems of Indistinguishable Particles'. In *Physical Review D* 3.6 (1971), pages 1375–1378.

[Lee03] John M. Lee: *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics 218. Springer, New York, 2003.

[LM77] Jon M. Leinaas and Jan Myrheim: 'On the Theory of Identical Particles'. In *Il Nuovo Cimento B* 37.1 (1977), pages 1–23.

[Lol90] Renate Loll: 'Noncommutativity of Constraining and Quantizing: A U(1) Model'. In *Physical Review D* 41.12 (1990), pages 3785–3791.

[LV80] Gérard Lion and Michèle Vergne: *The Weil Representation, Maslov Index and Theta Series*. Progress in Mathematics 6. Birkhäuser, Boston, 1980.

[Mac49] George W. Mackey: 'A Theorem of Stone and von Neumann'. In *Duke Mathematical Journal* 16.2 (1949), pages 313–326.

[Mac63] George W. Mackey: *The Mathematical Foundations of Quantum Mechanics. A Lecture-Note Volume*. W. A. Benjamin, Inc., New York, 1963.

[Mac72] Maciej J. Mączyński: 'Hilbert Space Formalism of Quantum Mechanics Without the Hilbert Space Axiom'. In *Reports on Mathematical Physics* 3.3 (1972), pages 209–219.

[Mac76] George W. Mackey: *The Theory of Unitary Group Representations*. Chicago Lectures in Mathematics. University of Chicago Press, 1976.

[Mac95] Saunders Mac Lane: *Homology*. Classics in Mathematics. Springer, Berlin, 1995.

[Mac98] Saunders Mac Lane: *Categories for the Working Mathematician*. 2nd edition. Graduate Texts in Mathematics 5. Springer, New York, 1998.

[Mic08] Peter W. Michor: *Topics in Differential Geometry*. Graduate Studies in Mathematics 93. American Mathematical Society, Providence, 2008. Draft version available at http://www.mat.univie.ac.at/~michor/dgbook.pdf.

[MM03] Ieke Moerdijk and Janez Mrčun: *Introduction to Foliations and Lie Groupoids*. Cambridge Studies in Advanced Mathematics 91. Cambridge University Press, Cambridge, 2003.

[Mos94] Martin Moskowitz: 'The Surjectivity of the Exponential Map for Certain Lie Groups'. In *Annali di Matematica Pura ed Applicata* 166.1 (1994), pages 129–143.

[MR99] Jerrold E. Marsden and Tudor S. Ratiu: *Introduction to Mechanics and Symmetry*. 2nd edition. Texts in Applied Mathematics 17. Springer, New York, 1999.

[Nak03] Mikio Nakahara: *Geometry, Topology and Physics*. 2nd edition. Graduate Student Series in Physics. Taylor & Francis, New York and London, 2003.

[Neu30] Johann von Neumann: 'Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren'. In *Mathematische Annalen* 102 (1930), pages 49–131.

[Neu31] Johann von Neumann: 'Die Eindeutigkeit der Schrödingerschen Operatoren'. In *Mathematische Annalen* 104.1 (1931), pages 570–578.

[Neu68] Johann von Neumann: *Mathematische Grundlagen der Quantenmechanik*. Unveränderter Nachdruck der ersten Auflage von 1932. Springer, Berlin, 1968.

[Omn99] Roland Omnès: *Understanding Quantum Mechanics*. Princeton University Press, 1999.

[Pen05] Roger Penrose: *The Road to Reality. A Complete Guide to the Laws of the Universe*. Knopf, New York, NY, 2005.

[Per86] Askold Perelomov: *Generalized Coherent States and Their Applications*. Texts and Monographs in Physics. Springer, 1986.

[PRL10] Nikolaos Papadopoulos and Andrés F. Reyes-Lega: 'On the Geometry of the Berry–Robbins Approach to Spin–Statistics'. In *Foundations of Physics* 40.7 (2010), pages 829–851.

[Puk64] Lajos Pukánszky: 'The Plancherel Formula for the Universal Covering Group of $\mathrm{SL}(R, 2)$'. In *Mathematische Annalen* 156.2 (1964), pages 96–143.

[Rey06] Andrés Reyes: 'On the Geometry of the Spin–Statistics Connection in Quantum Mechanics'. PhD thesis. Johannes Gutenberg-Universität Mainz, 2006.

[Ros04] Jonathan M. Rosenberg: 'A Selective History of the Stone–von Neumann Theorem'. In *Operator Algebras, Quantization, and Noncommutative Geometry*. Edited by Robert S. Doran and Richard V. Kadison. Contemporary Mathematics 365. American Mathematical Society, Providence, RI, 2004, pages 123–158. http://www-users.math.umd.edu/~jmr/.

[Rot02] Joseph J. Rotman: *Advanced Modern Algebra*. Pearson Education, Inc., New Jersey, 2002.

[Rot97] Gian-Carlo Rota: *Indiscrete Thoughts*. Birkhäuser, Boston, 1997.

[Rov96] Carlo Rovelli: 'Relational Quantum Mechanics'. In *International Journal of Theoretical Physics* 35.8 (1996), pages 1637–1678.

[RS.I] Michael Reed and Barry Simon: *Methods of Modern Mathematical Physics. I: Functional Analysis*. revised and enlarged ed. Academic Press, 1980.

[RS.II] Michael Reed and Barry Simon: *Methods of Modern Mathematical Physics. II: Fourier Analysis, Self-Adjointness*. Academic Press, 1975.

[RS.III] Michael Reed and Barry Simon: *Methods of Modern Mathematical Physics. III: Scattering Theory*. Academic Press, 1979.

[RS.IV] Michael Reed and Barry Simon: *Methods of Modern Mathematical Physics. IV: Analysis of Operators*. Academic Press, 1978.

[Sak94] Jun John Sakurai: *Modern Quantum Mechanics*. Edited by San Fu Tuan. Revised Edition. Addison-Wesley, Reading, Massachusetts, 1994.

[Sal67] Paul J. Sally Jr.: *Analytic Continuation of the Irreducible Unitary Representations of the Universal Covering Group of* $\mathrm{SL}(2, R)$. Memoirs of the American Mathematical Society 69. American Mathematical Society, Providence, RI, 1967.

[SC87] Chandra S. Sharma and T. J. Coulson: 'Quantum Theory in Real Hilbert Space'. In *Il Nuovo Cimento B* 100.3 (1987), pages 417–420.

[Sch04] Maximilian Schlosshauer: 'Decoherence, the Measurement Problem, and Interpretations of Quantum Mechanics'. In *Reviews of Modern Physics* 76.4 (2004), pages 1267–1305. arXiv:quant-ph/0312059v4.

[Sch07] Florian Scheck: *Quantum Physics*. Springer, Berlin, 2007.

[Sch08] Martin Schottenloher: *A Mathematical Introduction to Conformal Field Theory*. 2nd edition. Lecture Notes In Physics 759. Springer, Berlin, 2008.

[Sch26] Erwin Schrödinger: 'Quantisierung als Eigenwertproblem'. In *Annalen der Physik* 384.4 (1926), pages 361–376.

[SH11] Daniel Schumayer and David A. W. Hutchinson: 'Physics of the Riemann Hypothesis'. In *Reviews of Modern Physics* 83.2 (2011), pages 307–330. arXiv:1101.3116 [math-ph].

[Sha88] Chandra S. Sharma: 'Quantum Theory in Complex Hilbert Space'. In *Il Nuovo Cimento B* 102.3 (1988), pages 325–329.

[Sim96] Barry Simon: *Representations of Finite and Compact Groups*. Graduate Studies in Mathematics 10. American Mathematical Society, Providence, RI, 1996.

[SRL11] Germán Sierra and Javier Rodríguez-Laguna: 'The $H = xp$ Model Revisited and the Riemann Zeros'. In *Physical Review Letters* 106.20 (2011), page 200201. arXiv:1102.5356 [math-ph].

[Sto30] Marshall H. Stone: 'Linear Transformations in Hilbert Space: III. Operational Methods and Group Theory'. In *Proceedings of the National Academy of Sciences of the United States of America* 16.2 (1930), pages 172–175.

[Stu60] Ernst C. G. Stueckelberg: 'Quantum Theory in Real Hilbert Space'. In *Helvetica Physica Acta* 33.7 (1960), pages 727–752.

[Swa62] Richard G. Swan: 'Vector Bundles and Projective Modules'. In *Transactions of the American Mathematical Society* 105.2 (1962), pages 264–277.

[Sła91] Jan Sładkowski: 'Generalized G-Theory'. In *International Journal of Theoretical Physics* 30.4 (1991), pages 517–520.

[TAE05] Syed Twareque Ali and Miroslav Engliš: 'Quantization Methods: A Guide for Physicists and Analysts'. In *Reviews in Mathematical Physics* 17 (2005), pages 391–490. arXiv:math-ph/0405065v1.

[Tay86] Michael E. Taylor: *Noncommutative Harmonic Analysis*. Mathematical Surveys and Monographs 22. American Mathematical Society, Providence, RI, 1986.

[VH00] Léon Van Hove: 'Sur certaines représentations unitaires d'un groupe infini de transformations'. In *The Legacy of Léon Van Hove*. Edited by Alberto Giovannini. World Scientific, Singapore, 2000, pages 5–10.

[Wae67] Bartel L. Van der Waerden, editor: *Sources of Quantum Mechanics*. North-Holland Publishing Company, Amsterdam, 1967.

[Wal07] Stefan Waldmann: *Poisson-Geometrie und Deformations-Quantisierung. Eine Einführung*. Springer, Berlin, 2007.

[Wei64] André Weil: 'Sur Certains Groupes d'Opérateurs Unitaires'. In *Acta Mathematica* 111.1 (1964), pages 143–211.

[Woo09] Richard P. Woodard: 'How Far Are We From the Quantum Theory of Gravity?' In *Reports on Progress in Physics* 72.12 (2009), page 126001. arXiv:0907.4238 [gr-qc].

[Woo97] Nick M. J. Woodhouse: *Geometric Quantization*. 2nd edition. Oxford Mathematical Monographs. Clarendon Press, Oxford, 1997.