Complete Motion in Classical and Quantum Mechanics

Master thesis in Mathematical Physics

Author:
Ruben Stienstra
student number 3020371

Supervisor:
Prof. dr. Klaas Landsman

Mathematics Master Program
Radboud University Nijmegen

July 2014

Abstract

Classical mechanics allows for the possibility of 'incomplete motion', i.e., the motion of a particle on a geodesically incomplete configuration space Q is only defined for each time t in some bounded interval. On the other hand, the quantum-mechanical state of a particle is defined for each time $t \in \mathbb{R}$; thus the quantum-mechanical motion of that particle is complete. In this thesis, we examine different ways in which the quantum-mechanical motion can be defined by analysing the self-adjoint extensions of the Hamiltonian on some configuration space Q, our primary example being $Q = \overline{I}$, where I is some bounded open interval. Furthermore, we investigate the time evolution of particle-like states both analytically and numerically. In an attempt to explain our observations, we introduce a generalisation of the double of a manifold with boundary, and discuss when and how it can be used to define classically complete motion on the configuration space Q.

Contents

Introduction

The main topic of this thesis is *complete motion* in both classical and quantum mechanics. The notion of completeness of a motion is probably best explained from the viewpoint of classical mechanics. Suppose that we are given a particle with mass $m > 0$ on some open subset $\Omega \subseteq \mathbb{R}^n$, which we call the *configuration space* of the system. Classically, the state of a system at a time $t \in \mathbb{R}$ is given by an element of the *phase space* of the system; in the case of a single particle, the phase space is the cotangent bundle $T^*\Omega$. Since Ω is an open subset of $\mathbb{R}^n$, we can identify $T^*\Omega$ with $\mathbb{R}^n \times \Omega$, i.e., the cotangent bundle is trivial. The state of the particle at time t is now given by an element

$$(p_1(t), \dots, p_n(t), q_1(t), \dots, q_n(t)) = (p(t), q(t)) \in \mathbb{R}^n \times \Omega \cong T^*\Omega,$$

where $p(t)$ and $q(t)$ represent the momentum and the position of the particle, respectively. Classical mechanics asserts that the system obeys *Hamilton's equations of motion*:

$$\frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \quad \frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \quad j = 1, 2, \dots, n.$$

Here, H is the *classical Hamiltonian* of the system, which is the function $\mathbb{R}^n \times \Omega \to \mathbb{R}$ given by

$$(p, q) \mapsto \frac{p^2}{2m} + V(q) = \frac{1}{2m}\sum_{j=1}^{n} p_j^2 + V(q),$$

where $V\colon \Omega \to \mathbb{R}$ is a function called the *potential*. In order for Hamilton's equations to make sense, we must demand that V is differentiable. Given a potential and a set of initial conditions $(p(0), q(0)) = (p_0, q_0) \in \mathbb{R}^n \times \Omega$, one can attempt to solve Hamilton's equations. If there exists a *global* solution $t \mapsto (p(t), q(t))$ to this system of differential equations, i.e. if $(p(t), q(t))$ is defined for each $t \in \mathbb{R}$, then we say that the motion of the particle is *complete*. Otherwise, if only local solutions exist, then the motion is said to be *incomplete*.

It is very easy to find examples of both types of motions. If $n = 1$, $\Omega = \mathbb{R}$, and V vanishes everywhere, then any initial condition $(p(0), q(0)) = (p_0, q_0)$ will yield complete motion; this is the motion of a free particle on a line that is at q_0 when $t = 0$, and that moves with constant velocity p_0/m along the line. If however Ω is a proper open subset of $\mathbb{R}$, for example, if Ω is the open interval $\{x \in \mathbb{R}\colon -1 < x < 1\}$, and $(p(0), q(0)) = (p_0, q_0)$ with $p_0 > 0$ and $q_0 = 0$, then the motion is incomplete; it is only defined for $t \in \mathbb{R}$ with $-p_0/m < t < p_0/m$.

One can also obtain incomplete motion by choosing an appropriate potential. For example, setting $V(x) := -x^4 - 2x^2$, one can check that for a particle with mass 1, a solution to Hamilton's equations with initial conditions $(p(0), q(0)) = (2, 0)$ is given by $(p(t), q(t)) = (2(\tan^2(t) + 1), 2\tan(t))$. Thus the particle flies off to infinity in finite time. For a sufficient condition on the potential for the motion of a particle to be incomplete, we refer to [15, Theorem X.5].

In the realm of quantum mechanics, the situation appears to be quite different. First, let us recall that in this theory, the state of a particle is described by its wave function $\Psi(x, t)$, where x assumes values in Ω and t represents time. For a fixed time t, we require that $\Psi(\cdot, t) \in L^2(\Omega)$, and that $\|\Psi(\cdot, t)\|_{L^2(\Omega)} = 1$, so that $|\Psi(\cdot, t)|^2$ is the probability density function of some probability distribution, and the integral over some measurable subset $A \subseteq \Omega$ of $|\Psi(\cdot,t)|^2$ is the probability that the particle may be found on A upon measurement of the position of the particle at time t. The time evolution of the wave function of the particle is governed by *Schrödinger's equation*:

$$i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\Psi + V\Psi,$$

where $\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}$ is the Laplacian, V is again the potential, and $\hbar$ is a constant of nature, called the *reduced Planck constant* or the *Dirac constant*, with value $\hbar \approx 1.055 \cdot 10^{-34}$ Js. In order for this equation to have mathematical meaning, one must impose additional conditions on Ψ besides the requirement that Ψ be square integrable on Ω for fixed t. For example, $\Delta\Psi$ is not defined for each $\Psi \in L^2(\Omega)$. For the moment, though, we shall ignore these issues. The operator $-\frac{\hbar^2}{2m}\Delta + V$ is called the *Hamiltonian*, and is also denoted by H, so that Schrödinger's equation is often written more compactly as

$$i\hbar\frac{\partial \Psi}{\partial t} = H\Psi. \tag{0.1}$$

Schrödinger's equation, like Hamilton's equations, has to be supplemented with initial condition $\Psi(\cdot,0) = \psi \in L^2(\Omega)$, with $\|\psi\|_{L^2(\Omega)} = 1$.

One may solve this equation *in abstracto* using functional analytic methods. Here, we freely use some of the terminology from sections 1.4 and 1.5. First, one defines the operator H as a linear map on the space of smooth, compactly supported functions $C_0^\infty(\Omega)$ on Ω. This operator can subsequently be extended to a linear map $\widetilde{H}$ on a larger subspace $\mathcal{D}(\widetilde{H})$ of $L^2(\Omega)$, in such a way that $\widetilde{H}$ is self-adjoint. The converse of Stone's theorem (Theorem 1.5.3) now says that there exists a family of unitary operators on $L^2(\Omega)$, called a unitary evolution group $(U(t))_{t\in\mathbb{R}}$, with infinitesimal generator $\widetilde{H}$. Now, if $\psi \in \mathcal{D}(\widetilde{H})$, then it can be shown that $\Psi(\cdot,t) = U(t)\psi$ is the unique solution to Schrödinger's equation, and that $U(t)\psi \in \mathcal{D}(\widetilde{H})$ for each $t \in \mathbb{R}$.

There are three important things to notice here. First of all, the element $\Psi(\cdot,t) \in L^2(\Omega)$ is defined for each $t \in \mathbb{R}$. Secondly, $\Psi(\cdot,t)$ is contained in $\mathcal{D}(\widetilde{H})$ for each $t \in \mathbb{R}$, which means that equation (0.1) makes sense if we replace H with $\widetilde{H}$, and that $\Psi(\cdot,t)$ is a solution of this differential equation, where the time derivative is taken with respect to the norm on $L^2(\Omega)$. Thirdly, $U(t)$ is a unitary operator for each $t \in \mathbb{R}$. Since ψ was assumed to be normalised, it follows that $\Psi(\cdot,t)$ is normalised for each $t \in \mathbb{R}$, a property of the wave function that is often referred to by physicists as 'conservation of probability'. We conclude that $\Psi(\cdot,t) = U(t)\psi$ is a global solution of Schrödinger's equation, and therefore, that the quantum-mechanical motion can always be made complete, provided that there exists a self-adjoint extension $\widetilde{H}$ of H with $\psi \in \mathcal{D}(\widetilde{H})$.

This poses two problems. The first, most obvious one is the discrepancy between classical and quantum mechanics when it comes to the completeness of motion. Classical mechanics can to some extent be regarded as a limit case of quantum mechanics, by taking the limit $\hbar \to 0$ in some sense, see for example [12]. This begs the following question: does incompleteness of the motion of a particle arise upon taking the limit $\hbar \to 0$, or is something else going on here?

The second problem is more subtle. Recall that according to our method of solving Schrödinger's equation, we have to pick a self-adjoint extension $\widetilde{H}$ of H such that $\psi \in$

$\mathcal{D}(\widetilde{H})$. It can (and in this thesis, will) be shown that H always has at least one self-adjoint extension. However, it depends on the domain Ω and the potential V whether this extension is unique. It is known that H has a unique self-adjoint extension if $\Omega = \mathbb{R}^n$ and V vanishes everywhere, and more generally, for free particles on complete Riemannian manifolds (cf. [8] and [16]). On the other hand, if Ω is a bounded open interval and V vanishes everywhere, then we shall see later on that H has a family of self-adjoint extensions parametrised by the unitary group $\mathrm{U}(2)$. Though we shall see some examples of systems with Hamiltonians with a nonvanishing potential, we will mainly be concerned with the motion of the free particle on a bounded domain.

Now, Stone's theorem and its converse (Theorems 1.5.4 and 1.5.3, respectively) say that for each of these self-adjoint extensions, there is a unique unitary evolution group that has that self-adjoint extension as its infinitesimal generator. Thus different self-adjoint extensions correspond to potentially different time evolutions of the initial state, and hence to possibly different physical behaviour. This leads to the following philosophical issue: if domains like the open interval, on which H has multiple self-adjoint extensions, represent real physical systems, then which of the unitary evolution groups corresponding to the self-adjoint extension describes the system, and why does that specific unitary evolution group do so?

This thesis addresses both the problem of completeness of the motion, and that of non-uniqueness of the physics of the system.

The main body of the text is split into two parts: in the first part, consisting of sections 1,2 and 3, we investigate the self-adjoint extensions of the Hamiltonian and their corresponding unitary evolution groups.

- In section 1, we develop some of the analytical tools needed to formulate and approach the problem.
- In section 2, we discuss a general way of recognising and parametrising self-adjoint extensions of operators on Hilbert spaces, and employ this framework to classify the self-adjoint extensions of some interesting Hamiltonians.
- In section 3, we examine particle-like wave functions and their behaviour in the limit $\hbar \to 0$, employing both analytical and numerical methods. The main result is established in section 3.4, where the numerical simulations are discussed; it is observed that classical incomplete motion is most likely not a limit of quantum-mechanical complete motion.

In the second part of this thesis, consisting of sections 4 and 5, we outline a general procedure for constructing an alternative phase space of the system in an attempt to explain our observations at the end of section 3, and to solve both problems for free particles at a conceptual level.

- In section 4, we collect some of the results from differential geometry that are necessary to formulate the idea.
- In section 5, we shall perform the construction and discuss its merits and limitations. This is the most important section, since we introduce two new ideas here: first, we generalise the notion of the double of a manifold, and second, we argue why this generalisation is useful in understanding the physical behaviour associated to certain self-adjoint extensions of the Hamiltonian of the free particle.

Concerning the prerequisites, we assume that the reader is familiar with analysis and differential geometry at the level of a beginning master student of mathematics. More specifically, we assume that the reader has seen functional analysis at an introductory level, and as such, is familiar with the basic theory of Hilbert spaces and the most important example of these spaces, namely L^2-spaces. Furthermore, we assume that the reader has encountered differentiable manifolds and the flow of a vector field on these objects, and is comfortable with basic machinery such as the inverse function theorem.

Let us make some remarks on our notation and conventions:

- $\mathbb{N}$ denotes the set of positive integers, while $\mathbb{N}_0$ denotes the set of nonnegative integers.
- If X is a subset of a topological space, then X° denotes the interior of X.
- $\mathcal{N}(T)$ and $\mathcal{R}(T)$ denote the kernel and range of a linear map T, respectively.
- $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ will always be a Hilbert space. The symbol $\oplus$ will denote the orthogonal sum of subspaces.
- Sesquilinear forms such as the inner product on $\mathcal{H}$ are linear in their *second* argument.
- In order to avoid a mix-up of ordered pairs with open intervals, we shall use the symbols] and [, instead of (and), respectively, as delimiters of our intervals. For example, $]0,1[=\{x\in\mathbb{R}\colon 0<x<1\}$, and $[0,1[=\{x\in\mathbb{R}\colon 0\leq x<1\}$.
- If M is a matrix with complex-valued entries, then M^* denotes the hermitian conjugate of M.
- If $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)\in\mathbb{N}_0^n$, then $|\alpha|:=\sum_{j=1}^n\alpha_j$ is called the *length of* α. Furthermore, if f is a function on some open subset of $\mathbb{R}^n$ whose input is denoted by x, then $x^\alpha f$ is short-hand notation for the function

  $$x=(x_1,x_2,\ldots,x_n)\mapsto x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n}f(x).$$

  Similarly, $\partial^\alpha f$ is shorthand notation for the derivative

  $$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}\frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}}\cdots\frac{\partial^{\alpha_n}}{\partial x_1^{\alpha_n}}f.$$

  Finally, we define $D^\alpha f:=(-i)^{|\alpha|}\partial^\alpha f$. In particular, if f is a function on some interval, then $Df:=-if'$, and more generally, for each $m\in\mathbb{N}$, $D^m f:=(-i)^m f^{(m)}$.
- We shall sometimes refer to elements of some L^2-space as functions, even though this term is technically incorrect.

Finally, it is worth noting that throughout the rest of the text, the configuration space of the particle will be assumed to be the closure $\overline{\Omega}$ of Ω rather than Ω itself. Since the boundary of Ω typically has measure zero, we have $L^2(\overline{\Omega})=L^2(\Omega)$, so this assumption changes little about the quantum mechanical description of the problem. It *does* however have important implications for the classical mechanics of the system, as the nature of the boundary of Ω will play an essential role in section 5.

Acknowledgements

First and foremost, I would like to express my gratitude to Klaas Landsman for suggesting the subject of this thesis, for his excellent supervision, and for taking the time to read and correct my work and discuss it with me, even while he was officially on sabbatical leave during the second semester. I would also like to thank Ben Moonen for acting as the second reader of this thesis. In addition, I am indebted to Robin Reuvers for helping me with the visualisation of my MATLAB simulations, and to Hessel Posthuma for pointing me to some literature on the double of a manifold with boundary. Finally, I would like to thank my parents for their continuing support, mostly regarding nonacademical matters, throughout my studies.

1 Preliminaries from analysis

In this section we introduce some notions from various branches of analysis, such as distribution theory, Fourier analysis, and the theory of unbounded operators. These will play a major role in the next two sections.

1.1 Distribution theory

Here, we shall discuss the most basic ideas from the theory of distributions. The main purpose of this subsection is to generalise the notion of differentiation. All of the results in this section for which we do not explicitly give a reference can be found in chapters 2 and 3 of [10].

1.1.1 Definition. Let $n \in \mathbb{N}$, let $\Omega \subseteq \mathbb{R}^n$ be a subset, and let $f\colon \Omega \to \mathbb{C}$ be a function.

- The *support of* f, denoted by $\operatorname{supp}(f)$, is the closure of $f^{-1}(\mathbb{C}\backslash\{0\})$ with respect to the topology on $\mathbb{R}^n$.
- The function f is said to be *compactly supported* iff $\operatorname{supp}(f)$ is compact.

Now assume that Ω is a measurable subset of $\mathbb{R}^n$, and that f is a measurable function.

- Let X be an open subset of Ω. Then we say that f *is zero on* X iff $\{x \in X\colon f(x) \neq 0\}$ is a set of measure zero.
- Ω is a subset of $\mathbb{R}^n$, so endowed with its subspace topology, it is a second countable topological space. It follows that the union $X_{\max}$ of open sets on which f is zero, can be written as a countable union of open sets on which f is zero, which implies that $X_{\max}$ is the largest open subset of Ω on which f is zero. The set $\Omega\backslash X_{\max}$, denoted by $\operatorname{ess\,supp}(f)$, is called the *essential support of* f.
- Two functions that are equal almost everywhere on Ω have the same essential support. Thus we may define the essential support of an element of $L^1_{\text{loc}}(\Omega)$ as the essential support of one of its representatives.

1.1.2 Definition. Let $n \in \mathbb{N}$, and let $\Omega \subseteq \mathbb{R}^n$ be an open subset. The set of compactly supported, smooth functions f on Ω with $\operatorname{supp}(f) \subseteq \Omega$ is called *the space of test functions on* Ω and is denoted by $C_0^\infty(\Omega)$. It is a vector space under pointwise addition and scalar multiplication of functions.

The space of test functions on Ω can be endowed with a topology that turns this space into a topological vector space. In order to define this topology, we require the following lemma:

1.1.3 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be an open subset.*

(1) There exists a sequence $(K_j)_{j=1}^\infty$ of compact subsets of Ω such that for each $j \geq 1$, we have $K_j \subset K_{j+1}^\circ$, and $\bigcup_{j=1}^\infty K_j = \Omega$.

(2) For each $j \geq 1$ and each $k \geq 1$, define the map $p_{k,j}\colon C^\infty(\Omega) \to [0,\infty[$ by

$$p_{k,j}(f) := \sup\{|\partial^\alpha f(x)|\colon |\alpha| \leq k,\ x \in K_j\}.$$

Then $p_{k,j}$ is a seminorm on $C^\infty(\Omega)$, and the family $(p_{k,j})_{j\geq1,k\geq0}$ separates the points of $C^\infty(\Omega)$, endowing this space with a locally convex vector space topology.

Assume we are given Ω, $(K_j)_{j=1}^\infty$, and $(p_{k,j})_{j\geq 1,k\geq 0}$ as in the above lemma. For each $j \geq 1$, let $C^\infty_{K_j}(\Omega) := \{f \in C^\infty(\Omega)\colon \operatorname{supp}(f) \subseteq K_j\}$, and let $\iota_j\colon C^\infty_{K_j}(\Omega) \to C^\infty_0(\Omega)$ be the inclusion map. Endow $C^\infty_{K_j}(\Omega)$ with the subspace topology τ_j inherited from $C^\infty(\Omega)$. As to the topology on $C^\infty_0(\Omega)$, we take the strongest topology τ with the property that $\iota_j\colon (C^\infty_{K_j}, \tau_j)(\Omega) \to (C^\infty_0(\Omega), \tau)$ is continuous for each $j \geq 1$ and such that $(C^\infty_0(\Omega), \tau)$ is a locally convex topological vector space.

1.1.4 Definition. Let $\mathscr{D}'(\Omega)$ be the dual space of $(C^\infty_0(\Omega), \tau)$, endowed with the weak*-topology. Then $\mathscr{D}'(\Omega)$ is called the *space of distributions* on Ω.

1.1.5 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be an open subset, let $(K_j)_{j=1}^\infty$ be a sequence of compact subsets of Ω such that for each $j \geq 1$ we have $K_j \subset K^\circ_{j+1}$ and $\bigcup_{j=1}^\infty K_j = \Omega$, and let $(p_{k,j})_{j\geq 1,k\geq 0}$ be the corresponding family of seminorms. Then:*

(1) A sequence $(\varphi_l)_{l=1}^\infty$ in $C^\infty_0(\Omega)$ converges to an element $\varphi \in C^\infty_0(\Omega)$ if and only if there exists $j \in \mathbb{N}$ such that $\operatorname{supp}(\varphi_l) \subseteq K_j$ for each $l \geq 1$, and $\lim_{l\to\infty} p_{k,j}(\varphi_l - \varphi) = 0$ for each $k \geq 0$.

(2) A linear functional $\Lambda\colon C^\infty_0(\Omega) \to \mathbb{C}$ is a distribution on Ω if and only if for each $j \geq 1$, there exist $N_j \in \mathbb{N}_0$ and $c_j > 0$ such that for each $\varphi \in C^\infty_{K_j}(\Omega)$, we have

$$|\Lambda(\varphi)| \leq c_j \sup\{|\partial^\alpha \varphi(x)|\colon x \in K_j,\ |\alpha| \leq N_j\}.$$

1.1.6 Example. Let $\Omega \subseteq \mathbb{R}^n$ be an open subset.

(1) Let $f \in L^1_{\text{loc}}(\Omega)$, that is, f is integrable on every compact subset of Ω. Then the map $\Lambda_f\colon C^\infty_0(\Omega) \to \mathbb{C}$, given by $\varphi \mapsto \int_\Omega f(x)\varphi(x)\,dx$, is a distribution on Ω. A distribution is said to be *regular* iff it is of this form.

(2) Let $x_0 \in \Omega$. Then the map $\delta_{x_0}\colon C^\infty_0(\Omega) \to \mathbb{C}$, given by $\varphi \mapsto \varphi(x_0)$, is a distribution on Ω. We call δ_{x_0} the *Dirac* or the *delta distribution at* x_0.

We shall mainly be concerned with regular distributions. The following lemma allows us to identify $L^1_{\text{loc}}(\Omega)$ with the space of regular distributions on Ω:

1.1.7 Lemma. *Let $\Omega \subseteq \mathbb{R}^n$ be an open subset, and let $f \in L^1_{\text{loc}}(\Omega)$. If Λ_f is the zero functional, then $f = 0$.*

Finally, we define some maps on the space of distributions:

1.1.8 Definition. Let $\Omega \subseteq \mathbb{R}^n$ be an open subset, and let Λ be a distribution on Ω.

- Let $\alpha \in \mathbb{N}_0^n$. Then the map $\partial^\alpha\Lambda\colon C^\infty_0(\Omega) \to \mathbb{C}$, given by $\varphi \mapsto (-1)^{|\alpha|}\Lambda(\partial^\alpha\varphi)$, is a distribution on Ω. A distribution of the form $\partial^\alpha\Lambda$ with $\alpha \in \mathbb{N}_0^n$ is called a *distributional derivative of* Λ.
- Let $f \in C^\infty(\Omega)$. Then the map $M_f\Lambda\colon C^\infty_0(\Omega) \to \mathbb{C}$, given by $\varphi \mapsto \Lambda(f\varphi)$, is a distribution on Ω.

The above maps were defined with the intention of generalising the notions of differentiation and multiplication with a function to the space of distributions, as is demonstrated by the following proposition:

1.1.9 Proposition. *Let $\Omega \subseteq \mathbb{R}^n$ be an open subset, and let $f \in L^1_{\text{loc}}(\Omega)$.*

(1) Let $\alpha \in \mathbb{N}_0^n$. Then the map $\partial^\alpha \colon \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$, given by $\Lambda \mapsto \partial^\alpha \Lambda$, is continuous. Moreover, if g is $|\alpha|$-times continuously differentiable, then we have $\partial^\alpha \Lambda_f = \Lambda_{\partial^\alpha f}$.

(2) Let $g \in C^\infty(\Omega)$. Then the map $M_g \colon \mathscr{D}'(\Omega) \to \mathscr{D}'(\Omega)$, given by $\Lambda \mapsto M_g\Lambda$, is continuous. Moreover, we have $M_f \Lambda_g = \Lambda_{fg}$.

From here on, we shall often identify functions with their associated regular distributions.

1.2 Sobolev spaces

Even though the theory of distributions vastly expands the class of objects that we can differentiate in a sensible way, it does not guarantee that derivatives of elements of, say, $L^1_{\text{loc}}(\Omega)$, are again elements of that same space. For example, the distributional derivative $\frac{d}{dx}H$ of the Heaviside function $H \colon \mathbb{R} \to \mathbb{C}$, given by

$$x \mapsto \begin{cases} 0 & x \leq 0, \\ 1 & x > 0, \end{cases}$$

is the Dirac delta distribution at 0, which is not an element of $L^1_{\text{loc}}(\Omega)$. This motivates the following definitions:

1.2.1 Definition. Let $\Omega \subseteq \mathbb{R}^n$, and let $f \in L^1_{\text{loc}}(\Omega)$. If $\partial_{x_j} f \in L^1_{\text{loc}}(\Omega)$ for $j = 1, 2, \ldots, n$, then f is said to be *weakly differentiable*. The derivatives $\partial_{x_j} f$ are called *weak derivatives of f*.

1.2.2 Definition. Let $\Omega \subseteq \mathbb{R}^n$ be open, and let $m \in \mathbb{N}_0$.

- We define the *Sobolev space $H^m(\Omega)$ of order m on* Ω, by

  $$H^m(\Omega) := \{\phi \in L^1_{\text{loc}}(\Omega) \colon \partial^\alpha \phi \in L^2(\Omega) \text{ for each } \alpha \in \mathbb{N}_0^n \text{ such that } |\alpha| \leq m\}.$$

  It carries the structure of an inner product space, with inner product

  $$\langle \psi, \phi \rangle_{H^m(\Omega)} := \sum_{|\alpha| \leq m} \langle \partial^\alpha \psi, \partial^\alpha \phi \rangle_{L^2(\Omega)} = \sum_{|\alpha| \leq m} \int_\Omega \overline{\partial^\alpha \psi(x)} \partial^\alpha \phi(x)\, dx.$$

  In particular, we have $H^0(\Omega) = L^2(\Omega)$.

- The space $H_0^m(\Omega)$ is by definition the closure of $C_0^\infty(\Omega)$ in $(H^m(\Omega), \langle \cdot, \cdot \rangle_{H^m(\Omega)})$.

1.2.3 Proposition. *The spaces $(H^m(\Omega), \langle \cdot, \cdot \rangle_{H^m(\Omega)})$ and $(H_0^m(\Omega), \langle \cdot, \cdot \rangle_{H^m(\Omega)}|_{H_0^m(\Omega) \times H_0^m(\Omega)})$ are Hilbert spaces.*

Proof. See [10, p. 62]. ∎

We are primarily interested in open, connected subsets of $\mathbb{R}$, i.e., open intervals. Some of the properties that we shall be using are summed up in the following theorem:

1.2.4 Theorem. *Let $I \subseteq \mathbb{R}$ be an open interval (possibly unbounded) and let $m \in \mathbb{N}$.*

(1) Suppose $m > 0$. Each $\phi \in H^m(I)$ has a unique representative in $C^{m-1}(I)$ that can be (uniquely) extended to an element of $C^{m-1}(\overline{I})$ that we shall also call ϕ, slightly abusing notation. In this sense, ϕ and its derivatives of order $\leq m-1$ are bounded, and there exists a constant $C > 0$ such that

$$\sum_{j=0}^{m-1} \|\phi^{(j)}\|^2_{L^\infty(I)} \leq C \sum_{j=0}^{m-1} \|\phi^{(j)}\|^2_{L^2(I)} = C\|\phi\|^2_{H^m(I)} \text{ for each } \phi \in H^m(I).$$

(2) We have

$$H_0^m(I) = \{\phi \in H^m(I) \colon \phi^{(j)}(c) = 0 \text{ for each } c \in \partial I \text{ and } j = 0, 1, \ldots, m-1\}.$$

(3) If $I =]a, \infty[$, then for each $\phi \in H^m(I)$ and $j = 0, 1, \ldots, m-1$, the limit $\lim_{x\to\infty} \phi^{(j)}(x)$ exists and is equal to 0. Similarly, if $I =]-\infty, b[$, then for each $\phi \in H^m(I)$ and $j = 0, 1, \ldots, m-1$, the limit $\lim_{x\to-\infty} \phi^{(j)}(x)$ exists and is equal to 0. Finally, if $I = \mathbb{R}$, then for each $\phi \in H^m(I)$ and $j = 0, 1, \ldots, m-1$, both limits $\lim_{x\to\infty} \phi^{(j)}(x)$ and $\lim_{x\to-\infty} \phi^{(j)}(x)$ exist and are equal to 0.

(4) If $\phi \in H^m(\mathbb{R})$ satisfies $\operatorname{ess\,supp}(\phi) \subseteq I$, then the restriction $\phi|_I$ of ϕ to I is an element of $H^m(I)$. Conversely, if $\phi \in H^m(I)$, then its extension by zero $\widetilde{\phi}$ to $\mathbb{R}$ is an element of $H^m(\mathbb{R})$.

Proof. Parts (1), (2) and (4) can be found in [10], sections 4.2 and 4.3. To prove (3), suppose that $I =]a, \infty[$ and fix a constant $C > 0$ such that the inequality in part (3) of the theorem holds. For each $k \in \mathbb{N}$, let $I_k :=]a + k, \infty[$, let $\mathbf{1}_{I_k}$ be its characteristic function and let $\tau_k \colon I \to I_k$ be the map given by $x \mapsto x + k$. Then $\phi \mapsto \phi \circ \tau_k$ defines a unitary map $H^m(I_k) \to H^m(I)$. Let $\phi \in H^m(I)$. Then for each $k \in \mathbb{N}$, we have $\phi|_{I_k} \in H^m(I_k)$, and

$$\begin{aligned}\sum_{j=0}^{m-1} \|\phi^{(j)}|_{I_k}\|^2_{L^\infty(I_k)} &= \sum_{j=0}^{m-1} \|\phi^{(j)}|_{I_k} \circ \tau_k\|^2_{L^\infty(I)} \leq C\|\phi \circ \tau_k\|^2_{H^m(I)} \\ &= C\|\phi|_{I_k}\|^2_{H^m(I_k)} = C\|\phi \cdot \mathbf{1}_{I_k}\|^2_{H^m(I)}.\end{aligned}$$

Clearly, the functions $(\mathbf{1}_{I_k})_{k\in\mathbb{N}}$ converge to 0 pointwise, so by Lebesgue's theorem, the right-hand side of the above equation converges to 0 as $k \to \infty$. Hence, the left-hand side also converges to 0 as $k \to \infty$, and consequently, the limit $\lim_{x\to\infty} \phi^{(j)}(x)$ exists and is equal to 0 for $j = 0, 1, \ldots, m-1$.

A similar argument can be used to prove the statement for the case $I =]-\infty, b[$. To prove the statement for the case $I = \mathbb{R}$, we reduce it to the previous two cases by remarking that the restriction of any element $\phi \in H^m(\mathbb{R})$ to $]0, \infty[$ is an element of $H^m(]0, \infty[)$ and that its restriction to $]-\infty, 0[$ is contained in $H^m(]-\infty, 0[)$. ∎

Sobolev spaces are very useful in the study of partial differential equations. On the one hand, their Hilbert space structure allows one to prove existence and uniqueness of certain PDEs using methods from functional analysis, while on the other hand, they can be regarded as subspaces of spaces of functions that are differentiable up to a certain order (this is also true for Sobolev spaces on domains in dimension > 1).

Before we wrap up our discussion on Sobolev spaces, let us mention the following result:

1.2.5 Lemma. (Integration by parts) *Let $\phi, \psi \in H^1(a,b)$. Then*

$$\langle D_{\max}\phi, \psi\rangle - \langle \phi, D_{\max}\psi\rangle = i(\overline{\phi(b)}\psi(b) - \overline{\phi(a)}\psi(a)).$$

Proof. See [10, Theorem 4.14]. ∎

1.3 The Fourier transform

Here, we shall discuss two different ways to define the Fourier transform, along with its most important properties. This subsection summarises section 5.1 in [10]. For details and proofs of the statements, we refer to the aforementioned book. We begin with the easier of the two definitions of the Fourier transform:

1.3.1 Definition. Let $f \in L^1(\mathbb{R}^n)$. Then we define the *Fourier transform* $\mathcal{F}_1(f)\colon \mathbb{R}^n \to \mathbb{C}$ *of* f by

$$\mathcal{F}(f)(\xi) := \int_{\mathbb{R}^n} f(x)e^{-i\xi x}\,dx.$$

Using Lebesgue's theorem, it is readily seen that $\mathcal{F}_1(f) \in C(\mathbb{R}^n)$ for each $f \in L^1(\mathbb{R}^n)$. To define the second notion of a Fourier transform, we require the following space:

1.3.2 Definition. We define the *Schwartz space* $\mathcal{S}(\mathbb{R}^n)$ on $\mathbb{R}^n$ by

$$\mathcal{S}(\mathbb{R}^n) := \{f \in C^\infty(\mathbb{R}^n)\colon \sup_{x\in\mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty \text{ for each } \alpha, \beta \in \mathbb{N}_0^n\}.$$

The elements of $\mathcal{S}(\mathbb{R}^n)$ are called *rapidly decreasing functions.*

As their name already suggests, the elements of $\mathcal{S}(\mathbb{R}^n)$ decay rapidly at infinity. As a result, they have nice integrability properties. Furthermore, note that $C_0^\infty(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n)$.

1.3.3 Proposition.

(1) We have $\mathcal{S}(\mathbb{R}^n) \subseteq L^p(\mathbb{R}^n)$ for $p \in [1,\infty]$.

(2) The space $C_0^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $p \in [1,\infty[$. Consequently, $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $p \in [1,\infty[$ as well.

The Fourier transform behaves especially well on this space:

1.3.4 Lemma.

(1) The Gaussian $e^{-x^2/2}$ is an element of $\mathcal{S}(\mathbb{R}^n)$, and $\mathcal{F}_1(e^{-x^2/2})(\xi) = (2\pi)^{n/2}e^{-\xi^2/2}$.

(2) The map $\mathcal{F}_1|_{\mathcal{S}(\mathbb{R}^n)}\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is an isomorphism of vector spaces, with inverse

$$f \mapsto (\xi \mapsto (2\pi)^{-n}\mathcal{F}_1(f)(-\xi)).$$

(3) For each $f \in \mathcal{S}(\mathbb{R}^n)$, we have $\|f\|_{L^2(\mathbb{R}^n)} = (2\pi)^{-n/2}\|\mathcal{F}_1(f)\|_{L^2(\mathbb{R}^n)}$.

Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, the final part of the above lemma and the following lemma allow us to extend this map to $L^2(\mathbb{R}^n)$.

1.3.5 Lemma. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces, let X_1 be a dense subspace of X, and let $A\colon X \to Y$ be a bounded linear operator on X_1.*

(1) If $(Y, \|\cdot\|_Y)$is complete, then A has a unique bounded linear extension $\overline{A}$ to X, and $\|\overline{A}\| = \|A\|$.

(2) Suppose that A is an isometric isomorphism onto its image, and that the image $\mathcal{R}(A)$ is dense in Y. If both $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are complete, then the extension $\overline{A}$ is an isometric isomorphism from X onto Y.

Proof.
(1) See [18, Theorem 4.19].

(2) Since A is an isometric isomorphism onto its image, it has an inverse A^{-1}, and A and A^{-1} are both continuous linear maps with operator norm 1. By part 1 of the lemma, A has a continuous linear extension $\overline{A}$ to X, and $\|\overline{A}\| = 1$. Moreover, since $\mathcal{R}(A)$ is dense in Y and since X is complete, A^{-1} also has a continuous linear extension $B := \overline{A^{-1}}$ to Y, and $\|B\| = 1$. But then $B \circ \overline{A}$ and the identity map I_X on X are both continuous extensions of the identity map I_{X_1} on X_1. But X_1 is dense in X, so $B \circ \overline{A} = I_X$. Similarly, we have $\overline{A} \circ B = I_Y$, so $\overline{A}$ and B are mutually inverse. In particular, $\overline{A}$ is surjective. Finally, for each $x \in X$, we have

$$\|\overline{A}x\|_Y \le \|\overline{A}\| \cdot \|x\|_X = \|x\|_X = \|B \circ \overline{A}x\|_X \le \|B\| \cdot \|\overline{A}x\|_Y = \|\overline{A}x\|_Y,$$

so $\|\overline{A}x\|_Y = \|x\|_X$, which implies that $\overline{A}\colon X \to Y$ is an isometric isomorphism, as desired. ∎

1.3.6 Theorem. (Parseval-Plancherel)

(1) There exists a unique map $\mathcal{F}_2\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ that extends the map $\mathcal{F}_1|_{\mathcal{S}(\mathbb{R}^n)}\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, and has the property that $(2\pi)^{-n}\mathcal{F}_2$ is an isometric isomorphism, or equivalently, a unitary map.

(2) We have $\mathcal{F}_1(f) = \mathcal{F}_2(f)$ for each $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

By the second part of the Parseval-Plancherel theorem, we can define a map $\mathcal{F}$ on

$$L^1(\mathbb{R}^n) + L^2(\mathbb{R}^n) = \{f_1 + f_2 \colon f_1 \in L^1(\mathbb{R}^n),\ f_2 \in L^2(\mathbb{R}^n)\}.$$

that extends both $\mathcal{F}_1$ and $\mathcal{F}_2$. It is even possible to extend the two maps further to the *space of temperate distributions* $\mathcal{S}'(\mathbb{R}^n)$, which is the dual space of the Schwartz space endowed with a suitable vector space topology. However, we do not require this degree of generality, and we shall close our discussion of the Fourier transform by stating the following properties:

1.3.7 Proposition.

(1) *Let* $f \in L^2(\mathbb{R}^n)$*, let* $\alpha \in \mathbb{N}_0^n$*, and suppose that* $D^\alpha f \in L^2(\mathbb{R}^n)$*. Then* $\mathcal{F}(D^\alpha f) = \xi^\alpha \mathcal{F}(f)$*.*

(2) *Let* $f \in L^2(\mathbb{R}^n)$*, let* $\alpha \in \mathbb{N}_0^n$*, and suppose that* $x^\alpha f \in L^2(\mathbb{R}^n)$*. Then* $\mathcal{F}(x^\alpha f) = ((-D)^\alpha \mathcal{F}(f))$*.*

(3) *Let* $f, g \in L^1(\mathbb{R}^n)$*. Then the convolution* $f * g$ *of* f *and* g*, given by*

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\, dy,$$

is an element of $L^1(\mathbb{R}^n)$*, and* $\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$*.*

(4) *Let* $f, g \in L^2(\mathbb{R}^n)$*. Then* $f \cdot g \in L^1(\mathbb{R}^n)$*, and* $\mathcal{F}(f \cdot g) = (2\pi)^{-n} \mathcal{F}(f) * \mathcal{F}(g)$*.*

1.4 Unbounded operators

Let $\Omega \subseteq \mathbb{R}^n$ be an open subset. Many interesting operations, such as the differential or multiplication operators, are not well-defined maps, let alone bounded, from the Hilbert space $L^2(\Omega)$ to itself, even if we interpret them as operations on distributions. The theory of unbounded operators avoids this problem by dropping the requirement that such ill-defined operations be defined on the entire Hilbert space:

1.4.1 Definition. Let $\mathcal{H}$ be a Hilbert space.

- Let $V \subseteq \mathcal{H}$ be a subspace of $\mathcal{H}$. A linear map $T\colon V \to \mathcal{H}$ is called an *operator* on $\mathcal{H}$. The set V is called the *domain of* T, and is denoted by $\mathcal{D}(T)$.
- An operator T is said to be *densely defined* iff $\mathcal{D}(T)$ is dense in $\mathcal{H}$.
- Suppose S and T are linear operators on a Hilbert space. If $\mathcal{D}(S) \subseteq \mathcal{D}(T)$ and $T|_{\mathcal{D}}(S) = S$, then we write $S \subseteq T$.

Next, we wish to define the adjoint of a densely defined operator, for which we need the following lemma:

1.4.2 Lemma. *Let* T *be a densely defined operator on a Hilbert space* H*. Then for each* $x \in \mathcal{D}(T)$*, the linear functional* $f_x\colon \mathcal{D}(T) \to \mathbb{C}$*, given by* $z \mapsto \langle x, Tz\rangle$*, is continuous if and only if there exists a unique* $y \in \mathcal{H}$ *such that* $f_x(z) = \langle y, z\rangle$ *for each* $z \in \mathcal{D}(T)$*. Moreover, if such a* y *exists, then* f_x *has a unique continuous extension to an element of* $\mathcal{H}^*$*.*

Proof. Suppose f_x is continuous. By Lemma 1.3.5, it has a unique continuous extension g to $\mathcal{H}$. It follows from the Riesz representation theorem that there exists a unique $y \in \mathcal{H}$ such that $g(z) = \langle y, z\rangle$ for each $z \in \mathcal{H}$, so in particular, we have $f_x = \langle y, z\rangle$ for each $z \in \mathcal{D}(T)$.

Conversely, suppose that there exists a unique $y \in \mathcal{H}$ such that $f_x(z) = \langle y, z\rangle$ for each $z \in \mathcal{D}(T)$. Then f_x is continuous by the Cauchy-Schwarz inequality, and its unique continuous extension to $\mathcal{H}$ is of course the functional $z \mapsto \langle y, z\rangle$. ■

1.4.3 Definition. Let T be a densely defined operator on a Hilbert space H. For each $x \in \mathcal{D}(T)$, let $f_x \colon \mathcal{D}(T) \to \mathbb{C}$ be the linear functional given by $z \mapsto \langle x, Tz \rangle$. Then we define the *adjoint of* T as the operator T^* on $\mathcal{H}$ with domain

$$\mathcal{D}(T^*) := \{x \in \mathcal{H} \colon f_x \text{ is continuous}\},$$

that assigns to each $x \in \mathcal{D}(T^*)$ the unique element $y \in \mathcal{H}$ such that $f_x(z) = \langle y, z \rangle$ for each $z \in \mathcal{D}(T)$.

1.4.4 Remark. One readily verifies that T^* is indeed an operator on $\mathcal{H}$, i.e., it is linear. In addition, if T is a bounded operator, then T^* can be defined in two ways: either using the above definition, or as the adjoint of the bounded linear extension of T to $\mathcal{H}$. Both definitions yield the same (bounded) adjoint with domain $\mathcal{H}$, so the above definition extends the definition of adjoints of bounded operators on $\mathcal{H}$.

Next, we study the relation between an operator and its adjoint. A useful notion is the graph of an operator. Before we introduce it, let us recall that if $(V, \langle \cdot, \cdot \rangle)$, is an inner product space, then $V^2 = V \times V$ can be given the structure of an inner product space as well, with inner product $\langle \cdot, \cdot \rangle_{V^2}$ given by

$$\langle (x_1, y_1), (x_2, y_2) \rangle_{V^2} = \langle x_1, x_2 \rangle + \langle y_1, y_2 \rangle.$$

1.4.5 Definition. Let T be an operator on a Hilbert space $\mathcal{H}$.

- The set $\mathcal{G}(T) := \{(x, Tx) \in \mathcal{H}^2 \colon x \in V\}$ is called the *graph of* T.
- The operator T is said to be *closed* iff $\mathcal{G}(T)$ is a closed subspace of $\mathcal{H}^2$.
- The operator T is said to be *closable* iff the closure of $\mathcal{G}(T)$ in $\mathcal{H}^2$ is the graph of an operator on $\mathcal{H}$.
- It T is closable, then the *closure of* T, denoted by $\overline{T}$, is defined as the unique operator on $\mathcal{H}$ with graph $\overline{\mathcal{G}(T)}$.

1.4.6 Proposition. *Let T be an operator on a Hilbert space $\mathcal{H}$.*

(1) T is closable if and only if there exists a closed operator S on $\mathcal{H}$ such that $T \subseteq S$.

Now assume that T is densely defined.

(2) If S is an operator extending T, i.e., $T \subseteq S$, then $S^* \subseteq T^*$.

(3) Let $J \colon \mathcal{H}^2 \times \mathcal{H}^2$ be the unitary operator given by $(x, y) \mapsto (-y, x)$. Then we have $J(\overline{\mathcal{G}(T)}) \oplus \mathcal{G}(T^*) = \mathcal{H}^2$. Consequently, T^* is closed.

(4) The operator T is closable if and only if T^* is densely defined.

(5) If T is closable, then $T^{**} = \overline{T}$.

Proof.
(1) Obviously, if T is closable, then $\overline{T}$ is a closed operator extending T.

Conversely, suppose S is a closed operator extending T. Let

$$V := \{x \in \mathcal{D}(S) \colon (x, Sx) \in \overline{\mathcal{G}(T)}\},$$

and define the operator T' on $\mathcal{H}$ as the restriction of S to V. Since S is an operator extending T, we have $\mathcal{G}(T) \subseteq \mathcal{G}(S)$, and from the fact that S is closed, we infer that $\overline{\mathcal{G}(T)} \subseteq \mathcal{G}(S)$. This implies that $\mathcal{G}(T') = \overline{\mathcal{G}(T)}$, so T is closable, and $\overline{T} = T'$.

(2) This is readily verified from the definition of the adjoint.

(3) Let $(x, y) \in \mathcal{H}^2$. Then the following are equivalent:
- $(x, y) \in \mathcal{G}(T^*)$;
- For each $z \in \mathcal{D}(T)$, we have $\langle x, Tz \rangle = \langle y, z \rangle$;
- For each $z \in \mathcal{D}(T)$, we have $\langle J(z, Tz), (x, y) \rangle_{\mathcal{H}^2} = 0$;
- $(x, y) \in J(\mathcal{G}(T))^\perp$.

This proves the assertion.

(4) First note that a subspace $V \subseteq \mathcal{H}^2$ is the graph of an operator if and only if for each $x \in \mathcal{H}$, there exists at most one $y \in \mathcal{H}$ such that $(x, y) \in V$. Since V is a linear subspace, this is equivalent to the condition that if $(0, y) \in V$, then necessarily $y = 0$. Moreover, from the previous part of the proposition and the fact that the map J defined therein is unitary and satisfies $J^2 = -\mathrm{Id}_{\mathcal{H}^2}$, we see that

$$\overline{\mathcal{G}(T)} \oplus J(\mathcal{G}(T^*)) = \mathcal{H}^2.$$

Thus for each $y \in \mathcal{H}$, the following statements are equivalent:
- T is closable;
- If $(0, y) \in \overline{\mathcal{G}(T)}$, then $y = 0$;
- If $(0, y) \in J(\mathcal{G}(T^*))^\perp$, then $y = 0$;
- If $\langle J(x, T^*x), (0, y) \rangle_{\mathcal{H}^2} = 0$ for each $x \in \mathcal{D}(T^*)$, then $y = 0$;
- If $\langle x, y \rangle = 0$ for each $x \in \mathcal{D}(T^*)$, then $y = 0$;
- $\mathcal{D}(T^*)$ is dense in $\mathcal{H}$;
- T^* is densely defined.

(5) Suppose T is closable. By the previous part of the proposition, T^* is densely defined, so it has an adjoint. Applying part (3) of the proposition twice yields $\overline{\mathcal{G}(T)} \oplus J(\mathcal{G}(T^*)) = \mathcal{G}(T^{**}) \oplus J(\mathcal{G}(T^*))$, which implies that $\overline{\mathcal{G}(T)} = \mathcal{G}(T^{**})$, or equivalently, $\overline{T} = T^{**}$, as desired. ■

Let us introduce some important terminology:

1.4.7 Definition. Let T be a densely defined operator on a Hilbert space $\mathcal{H}$.
- The operator T is said to be *hermitian* iff $T \subseteq T^*$.
- The operator T is said to be *self-adjoint* iff $T = T^*$.
- The operator T is said to be *essentially self-adjoint* iff $\overline{T}$ is self-adjoint.

1.4.8 Proposition. *Let T be a densely defined operator on a Hilbert space $\mathcal{H}$.*

(1) T is hermitian if and only if $\langle Tx, y \rangle = \langle x, Ty \rangle$ for each $x, y \in \mathcal{D}(T)$.

(2) Suppose that T is hermitian. Then T is essentially self-adjoint if and only if T^* is hermitian.

(3) If T is essentially self-adjoint, then $\overline{T}$ is its unique self-adjoint extension.

Proof.
(1) This is an easy consequence of the definition.

(2) If T is essentially self-adjoint, then $\overline{T}$ is self-adjoint. It follows from part (4) of Proposition 1.4.6 that $\overline{T} = T^{**}$. Applying parts (3) and (4) of Proposition 1.4.6 yields

$$T^{**} = \overline{T} = \overline{T}^* = T^{***} = \overline{T^*} = T^*,$$

so T^* is self-adjoint, and in particular it is hermitian.

Conversely, suppose T^* is hermitian. Then, we have $T^* \subseteq T^{**} = \overline{T}$. On the other hand, we know that T is hermitian, so $T \subseteq T^*$, and since T^* is closed, it follows that $\overline{T} \subseteq T^*$. Thus $\overline{T} = T^*$, and hence $\overline{T}^* = T^{**} = \overline{T}$, so T is essentially self-adjoint.

(3) Let S be a self-adjoint extension of T. Then S is closed by part (3) of Proposition 1.4.6, so $\overline{T} \subseteq S$. Part (2) of that proposition now implies that $S = S^* \subseteq \overline{T}^* = \overline{T}$, hence $\overline{T} = S$. ■

The two examples of hermitian operators that we shall study are the following ones:

1.4.9 Example. Let $\Omega \subseteq \mathbb{R}^n$ be an open subset with C^1-boundary.

(1) For each $\alpha \in \mathbb{N}_0^n$, the operator D^α with domain $C_0^\infty(\Omega)$ is a hermitian operator on $L^2(\Omega)$.

(2) Let $V \in L^1_{\mathrm{loc}}(\Omega)$ be a locally integrable function that is (almost everywhere) real-valued. Slightly abusing notation, we shall use the letter V for its associated multiplication operator on $\mathscr{D}'(\Omega)$. Then the operator $H = -\Delta + V$ with domain $C_0^\infty(\Omega)$ is a hermitian operator on $L^2(\Omega)$.

In both cases, one readily verifies that the operator is hermitian by using integration by parts. Hermitian differential operators with domain $C_0^\infty(\Omega)$ like the ones above are said to be *formally self-adjoint*.

1.5 Stone's theorem and its converse

Finally, we come to our main reason for introducing the notion of self-adjointness.

1.5.1 Definition. Let $\mathcal{H}$ be a Hilbert space.

- A *unitary evolution group* on $\mathcal{H}$ is a group homomorphism U from $(\mathbb{R}, +)$ to the group of unitary operators on $\mathcal{H}$ (with compostion). In the rest of the text, unitary evolution groups will be denoted by $(U(t))_{t\in\mathbb{R}}$.
- Let $(U(t))_{t\in\mathbb{R}}$ be a unitary evolution group on $\mathcal{H}$. The operator T on $\mathcal{H}$ with domain

  $$\mathcal{D}(T) := \{x \in \mathcal{H}\colon \lim_{t\to 0} t^{-1}(U(t)x - x) \text{ exists.}\},$$

  on which T is given by

  $$x \mapsto i \lim_{t\to 0} t^{-1}(U(t)x - x),$$

  is called the *infinitesimal generator of* $(U(t))_{t\in\mathbb{R}}$.
- A unitary evolution group $(U(t))_{t\in\mathbb{R}}$ is said to be *strongly continuous* iff for each $x \in \mathcal{H}$, the limit $\lim_{t\to 0} U(t)x$ exists and is equal to x.

1.5.2 Lemma. *Let $(U(t))_{t\in\mathbb{R}}$ be a unitary evolution group on a Hilbert space $\mathcal{H}$ with infinitesimal generator T. Then $\mathcal{D}(T)$ is an invariant subspace of $U(t)$ for each $t \in \mathbb{R}$, and T commutes with $\mathcal{D}(T)$.*

Proof. Let $t \in \mathbb{R}$, let $x \in \mathcal{D}(T)$. For each $s \in \mathbb{R}\backslash\{0\}$, we have

$$s^{-1}(U(s) - \mathrm{Id}_{\mathcal{H}})U(t) = s^{-1}(U(s+t) - U(t)) = U(t)s^{-1}(U(s) - \mathrm{Id}_{\mathcal{H}})U(t),$$

and

$$Tx = i \lim_{s\to 0} s^{-1}(U(s) - \mathrm{Id}_{\mathcal{H}})x,$$

so by the boundedness of $U(t)$, the limit

$$i \lim_{s\to 0} s^{-1}(U(s) - \mathrm{Id}_{\mathcal{H}})U(t)x,$$

exists, and

$$TU(t)x = i \lim_{s\to 0} s^{-1}(U(s) - \mathrm{Id}_{\mathcal{H}})U(t)x = iU(t) \lim_{s\to 0} s^{-1}(U(s) - \mathrm{Id}_{\mathcal{H}})x = U(t)Tx,$$

which proves the lemma. ■

Let us first state the converse of Stone's theorem:

1.5.3 Theorem. *Let T be a self-adjoint operator on a Hilbert space H. Then there exists a unique strongly continuous unitary evolution group $(U(t))_{t\in\mathbb{R}}$ with infinitesimal generator T.*

Sketch of the proof. The proof uses some machinery from functional analysis. First, one applies the spectral theorem for unbounded self-adjoint operators to T. This yields a map E from the Borel σ-algebra of $\mathbb{R}$ to the space of bounded operators on $\mathcal{H}$ that assigns to each Borel set a projection in $\mathcal{H}$, in such a way that E is a so-called projection-valued measure, and T can be written as an integral $\int_{\mathbb{R}} \lambda\, dE(\lambda)$. The unitary evolution group is then defined as the map

$$t \mapsto \int_{\mathbb{R}} e^{-it\lambda}\, dE(\lambda),$$

and is more commonly denoted by $(e^{-itT})_{t\in\mathbb{R}}$. For details, we refer to chapters 4 and 5, and to Propositon 6.1 in [19] ■

1.5.4 Theorem. (Stone) *Let $(U(t))_{t\in\mathbb{R}}$ be a strongly continuous unitary evolution group on a Hilbert space $\mathcal{H}$, and let T be its infinitesimal generator. Then T is self-adjoint, and $U(t) = e^{-itT}$ for each $t \in \mathbb{R}$.*

Proof. See [19, Theorem 6.2] or [4, Theorem 5.3.3]. ■

1.5.5 Corollary. *Let $\mathcal{H}$ be a Hilbert space. Then there exists a bijection from the set of strongly continuous unitary evolution groups on $\mathcal{H}$ to the set of self-adjoint operators on $\mathcal{H}$. The bijection maps a unitary evolution group $(U(t))_{t\in\mathbb{R}}$ to its infinitesimal generator T. The inverse map sends a self-adjoint operator T to the unitary evolution group $(e^{-itT})_{t\in\mathbb{R}}$.*

Thus in order to study the unitary evolution groups on a Hilbert space, one can also examine the self-adjoint operators on that Hilbert space, a task that we take up in the next section.

2 Self-adjoint extensions of hermitian operators

Let $a, b \in \mathbb{R}$, $a < b$, let $I :=]a, b[$, and consider the differential operators $D = -i\frac{d}{dx}$ and $H := D^2 + V = -\frac{d^2}{dx^2} + V$ on $L^2(I)$, both with domain $C_0^\infty(I)$. We have already noted that D and H are hermitian operators. In what follows, we shall call these operators *test operators*, as their domain is the space of test functions on I. The natural question to ask now is whether they have self-adjoint extensions, and if so, how many of them. Using a method described by Everitt & Markus in [6] and [7], we shall obtain a necessary and sufficient condition on the closed extensions of these operators to be self-adjoint, involving symplectic forms (in the case of H, for suitable potentials), and we shall see how these symplectic forms can be constructed from arbitrary hermitian operators on Hilbert spaces.

2.1 First example: the operator $D = -i\frac{d}{dx}$

In this subsection, we introduce some general ideas to determine the self-adjoint extensions of the operator D on $L^2(a, b)$. Adopting the terminology in [10], we classify certain extensions of linear operators as follows:

2.1.1 Definition. Let T be a hermitian operator on $\mathcal{H}$.

- The adjoint operator T^*, denoted by $T_{\max}$, is called the *maximal realisation* of T.
- A closed extension $\widetilde{T}$ of T such that $\widetilde{T} \subseteq T_{\max}$ is called a *realisation of* T.
- The closure $\overline{T}$ of T, denoted by $T_{\min}$, is called the *minimal realisation* of T.

2.1.2 Remark. Let T be as in the previous definition.

(1) $T_{\min}$ is a realisation of T, and for any realisation $\widetilde{T}$ of T, we have $T_{\min} \subseteq \widetilde{T}$. This justifies the term 'minimal realisation'.

(2) Since T is hermitian, we have $T_{\min}^* = T^* = T_{\max}$ and $T_{\max}^* = T_{\min}$ by parts (3) and (5) of Proposition 1.4.6.

(3) Let $\Omega \subseteq \mathbb{R}^n$, and suppose T is a hermitian differential operator such as D or H on $L^2(\Omega)$ with domain $C_0^\infty(\Omega)$. Then T defines a continuous operator T_{dist} on the space of distributions $\mathscr{D}'(\Omega)$ on Ω. One readily sees from the definition of the adjoint of an operator that $\mathcal{D}(T_{\max}) = L^2(\Omega) \cap T_{\text{dist}}^{-1}(L^2(a, b))$. In other words, $\mathcal{D}(T_{\max})$ is the set of all elements of $L^2(\Omega)$ for which the differential operator is defined in the weak sense, and for which these 'weak derivatives' are again elements of $L^2(\Omega)$. This justifies the term 'maximal realisation'. In particular, $\mathcal{D}(D_{\max})$ is the set of all weakly differentiable elements of $L^2(\Omega)$, whose derivatives are also elements of $L^2(\Omega)$, so $\mathcal{D}(D_{\max}) = H^1(\Omega)$.

In order to determine the realisations of D and H, the following algebraic notion turns out to be very useful:

2.1.3 Definition. Let V be a complex vector space.

- A *complex symplectic form* ω on V is a nondegenerate sesquilinear form on V that is skew-hermitian, i.e.

$$\omega(u,v) = -\overline{\omega(v,u)} \quad \text{for all } u,v \in V.$$

  The pair (V,ω) is called a *complex symplectic vector space.*

Let $U \subseteq V$ be a linear subspace.

- The set

$$U^\omega := \{v \in V \colon \omega(u,v) = 0 \text{ for each } u \in U\}$$

  is called the *symplectic complement* of U in V.
- The subspace U is said to be *isotropic* iff $U \subseteq U^\omega$.
- The subspace U is said to be *Lagrangian* iff $U = U^\omega$.

2.1.4 Remark.

- Contrary to real symplectic forms, complex symplectic forms can exist on odd-dimensional vector spaces, e.g. $(x,y) \mapsto i\overline{x}y$ defines a complex symplectic form on $\mathbb{C}$.
- Given a finite-dimensional vector space V with basis $(e_1,\ldots,e_n)$, there is a bijective correspondence between complex symplectic forms ω on V and invertible complex $n \times n$-matrices B such that $B = -B^*$. This correspondence is given by

$$B_{jk} = \omega(e_j, e_k) \text{ for } 1 \le j,k \le n,$$

  with inverse

$$\omega\left(\sum_{j=1}^n c_j e_j, \sum_{k=1}^n d_k e_k\right) = \mathbf{c}^* B \mathbf{d},$$

  where $\mathbf{c}$ is the column vector whose j-th entry is c_j, and $\mathbf{d}$ is defined analogously. Moreover, the matrix iB is symmetric, which implies that it is diagonalisable with orthogonal eigenspaces (as subspaces of $\mathbb{C}^n$), so the same is true for B. Each eigenvalue of B is purely imaginary.

2.1.5 Proposition. *Let (V,ω) be a (possibly infinite-dimensional) complex symplectic vector space, and let $U \subseteq V$ be a linear subspace. Then:*

(1) U^ω is a linear subspace of V.

(2) Suppose $\|\cdot\|$ is a norm on V such that ω is continuous with respect to this norm. Then U^ω is a closed linear subspace of V. In particular, Lagrangian subspaces of V are closed.

(3) If V is finite dimensional, then $\dim U + \dim U^\omega = \dim V$. *In particular, U is Lagrangian if and only if U is isotropic and* $2\dim U = \dim V$.

Proof.
(1) For each $u \in U$, let $f_u \colon V \to \mathbb{C}$ be the map given by $f_u(v) = \omega(u, v)$. Then f_u is a linear functional on V, so $\mathcal{N}(f_u)$ is a linear subspace of V and hence $U^\omega = \bigcap_{u \in U} \mathcal{N}(f_u)$ is a linear subspace of V.

(2) If ω is continuous, then $|\omega(u, v)| \leq c\|u\|\|v\|$ for some $c > 0$, so the linear functionals f_u $(u \in U)$ satisfy $|f_u(v)| \leq c\|u\|$ and are therefore continuous. Hence $\mathcal{N}(f_u)$ is closed for each $u \in U$, and looking at the proof of the previous part of this propostion, this implies that U^ω is closed.

(3) Consider the map $A \colon U^\omega \to (V/U)^*$ given by $u \mapsto \tilde{f}_u$, where $\tilde{f}_u$ is defined by $\tilde{f}_u(v + U) := f_u(v) = \omega(u, v)$. Note that $\tilde{f}_u$ is well defined since $U \subseteq \mathcal{N}(f_u)$ for each $u \in U^\omega$, so A is well defined. The map A is antilinear since ω is antilinear in its first argument. Now suppose that $u \in U^\omega$ is an element such that $\tilde{f}_u(v + U) = 0$ for each $v + U \in V/U$. Then $\omega(u, v) = 0$ for each $v \in V$, so $u = 0$, since ω is nondegenerate. Thus A is injective.

Finally, suppose $g \in V/U$ is a linear functional. Then $\hat{g}(v) := g(v + U)$ defines a linear functional on V. By assumption, ω is nondegenerate, so the map $B \colon V \to V^*$ given by $u \mapsto f_u$ is an injective antilinear map. Since V is finite dimensional, we have $V \cong V^*$, so B is a bijection, and consequently, there exists a $u \in V$ such that $f_u = \hat{g}$. Because $U \subseteq \mathcal{N}(\hat{g})$, we have $u \in U^\omega$, and f_u descends to the linear funtional g on V/U. Thus A is surjective, and it follows that $U^\omega \simeq (V/U)^* \simeq V/U$, so

$$\dim U + \dim U^\omega = \dim U + \dim V/U = \dim V.$$

■

2.1.6 Lemma. *Let X and Y be topological vector spaces such that $\dim Y < \infty$, and let $S \colon X \to Y$ be a linear map. If $\mathcal{N}(S)$ is closed in X, then S is continuous.*

Proof. The range $\mathcal{R}(S)$ of S with the subspace topology inherited from Y is again a topological vector space, and the inclusion map $\iota \colon \mathcal{R}(S) \to Y$ is continuous. Now $X/\mathcal{N}(S)$ with the quotient topology is a topological vector space by [17, Theorem 1.41(a)], since $\mathcal{N}(S)$ is a closed linear subspace of X, and the canonical projection $\pi \colon X \to X/\mathcal{N}(S)$ is continuous. Finally, the map $\widetilde{S} \colon X/\mathcal{N}(S) \to \mathcal{R}(S)$, given by $\widetilde{S}(u + \mathcal{N}(S)) := S(u)$ is an isomorphism between finite dimensional vector spaces, so by [17, Theorem 1.21(a)], it is an isomorphism of topological vector spaces. In particular, $\widetilde{S}$ is continuous. Since $S = \iota \circ \widetilde{S} \circ \pi$, it follows that S is continuous. ■

2.1.7 Theorem. *Let $I :=]a, b[$.*

(1) Let $\varrho \colon \mathcal{G}(D_{\max}) \to \mathbb{C}^2$ be the map given by $(\phi, D_{\max}\phi) \mapsto (\phi(a), \phi(b))$. Then ϱ is linear, continuous and surjective, and $\mathcal{N}(\varrho) = \mathcal{G}(D_{\min})$.

(2) Let $\pi_1 \colon \mathcal{G}(D_{\max}) \to \mathcal{D}(D_{\max})$ be the projection on the first coordinate. The map

$$(*) \qquad \widetilde{D} \mapsto \varrho(\mathcal{G}(\widetilde{D})),$$

yields a bijective correspondence between realisations of D and the linear subspaces of $\mathbb{C}^2$, with inverse

$$(**) \qquad U \mapsto D_{\max}|_{\pi_1(\varrho^{-1}(U))}.$$

(3) Let $\omega : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}$ be the complex symplectic form given by $(\mathbf{c}, \mathbf{d}) \mapsto \mathbf{c}^* B\mathbf{d}$, where $\mathbf{c}$ and $\mathbf{d}$ are both column vectors, and

$$B := \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

For each linear subspace $U \subseteq \mathbb{C}^2$, let D_U be the realisation of D associated to it by the correspondence described in (2). Then $D_U^* = D_{U^\omega}$. Consequently, the hermitian realisations of D correspond to the isotropic subspaces of $\mathbb{C}^2$, and the self-adjoint realisations of D correspond to the Lagrangian subspaces of $\mathbb{C}^2$.

(4) A realisation $\widetilde{D} \subseteq D_{\max}$ of D is self-adjoint if and only if its domain is of the form

$$\mathcal{D}(\widetilde{D}) = \{\phi \in H^1(I) \colon \phi(b) = e^{i\theta}\phi(a)\},$$

for some $\theta \in [0, 2\pi[$.

Proof.

(1) The map ϱ is clearly linear. Observe that the map $H^1(I) \to L^2(I) \times L^2(I)$, given by $x \mapsto (x, D_{\max}x)$ is an isometry with image $\mathcal{G}(D_{\max})$. This, together with part (2) of Theorem 1.2.4, implies that $H_0^1(I) = \mathcal{D}(D_{\min})$ and $\mathcal{N}(\varrho) = \mathcal{G}(D_{\min})$, which is closed in $L^2(I) \times L^2(I)$ and which is therefore also closed in $\mathcal{G}(D_{\max})$. The continuity of ϱ now follows from Lemma 2.1.6.

Next, suppose $(c_1, c_2) \in \mathbb{C}^2$. Then

$$P(x) := c_1 + (c_2 - c_1)\frac{x-a}{b-a}$$

is a polynomial such that $P(a) = c_1$ and $P(b) = c_2$. P and its derivative P' are square integrable on I, so $P \in H^1(I)$, which implies that ϱ is surjective.

(2) Let $\widetilde{D}$ be some realisation of D. By linearity of ϱ, the set $\varrho(\mathcal{D}(\widetilde{D}))$ is a subspace of $\mathbb{C}^2$, so the map given in $(*)$ is well defined.

Let U be a linear subspace of $\mathbb{C}^2$. Then U is also a closed subset of $\mathbb{C}^2$. Because ϱ is continuous and linear, $\varrho^{-1}(U)$ is a closed subspace of $\mathcal{G}(D_{\max})$ containing $\mathcal{G}(D_{\min})$, and hence is a closed subspace of $L^2(I)\times L^2(I)$, since $\mathcal{G}(D_{\max})$ is closed in $L^2(I)\times L^2(I)$. Thus the operator $D_{\max}|_{\pi_1(\varrho^{-1}(U))}$ is a closed extension of $D_{\min}$, and therefore it is a realisation of D. This shows that the map given in $(**)$ is well defined.

It remains to be shown that the two maps are mutual inverses. Again, let $\widetilde{D}$ be a realisation of D. First applying $(*)$ and subsequently $(**)$ to $\widetilde{D}$ yields a realisation $\widehat{D}$ with graph $\varrho^{-1}(\varrho(\mathcal{G}(\widetilde{D}))) \subseteq \mathcal{G}(D_{\max})$. It is clear that $\mathcal{G}(\widetilde{D}) \subseteq \mathcal{G}(\widehat{D})$, or equivalently, $\widetilde{D} \subseteq \widehat{D}$. Now let $\phi \in \mathcal{D}(\widehat{D})$. Then $(\phi, D_{\max}\phi) \in \mathcal{G}(\widehat{D})$, so there exists a $\psi \in \mathcal{D}(\widetilde{D})$ such that $\varrho(\phi, D_{\max}\phi) = \varrho(\psi, D_{\max}\psi)$. But then $\varrho(\phi - \psi, D_{\max}(\phi - \psi)) = (0, 0)$, which implies $\phi - \psi \in H_0^1(I) = \mathcal{D}(D_{\min})$. $\widetilde{D}$ is a realisation of D, so $D_{\min} \subseteq \widetilde{D}$, which implies $\phi - \psi \in \mathcal{D}(\widetilde{D})$ and hence $\phi = \phi - \psi + \psi \in \mathcal{D}(\widetilde{D})$. We conclude that $\widetilde{D} = \widehat{D}$.

Finally, let $U \subseteq \mathbb{C}^2$ be a subspace. First applying $(**)$ and subsequently $(*)$ to U yields the subspace $\varrho(\varrho^{-1}(U))$, which is equal to U, since ϱ is surjective. Thus the maps defined in $(*)$ and $(**)$ are indeed inverses of each other.

(3) Let $U \in \mathbb{C}^2$ and let $\phi \in \mathcal{D}(D_{\max})$. If $\phi \in \mathcal{D}(D_U^*)$, then $D_U^*\phi = D_{\max}\phi$, so $\phi \in \mathcal{D}(D_U^*)$ if and only if

$$\langle \phi, D_{\max}\psi\rangle = \langle \phi, D_U\psi\rangle = \langle D_{\max}\phi, \psi\rangle \quad \text{for each } \psi \in \mathcal{D}(D_U).$$

By the integration by parts formula (Lemma 1.2.5), this is equivalent to

$$i(\overline{\phi(b)}\psi(b) - \overline{\phi(a)}\psi(a)) = 0 \quad \text{for each } \psi \in \mathcal{D}(D_U).$$

The expression on the left-hand side of this equation is exactly $\omega(\varrho \circ \pi_1^{-1}(\phi), \varrho \circ \pi_1^{-1}(\psi))$. In view of $\mathcal{D}(D_U) = \pi_1(\varrho^{-1}(U))$ and the fact that ϱ is a surjection and π_1 a bijection, we conclude that $\phi \in \mathcal{D}(D_U^*)$ if and only if $\varrho \circ \pi_1^{-1}(\phi) \in U^\omega$. It follows that $D_U^* = D_{U^\omega}$.

(4) From part (3) of this theorem, we know that the self-adjoint extensions D_U of D correspond to the Lagrangian subspaces U of $\mathbb{C}^2$ via

$$\mathcal{D}(D_U) = \{\phi \in H^1(I) \colon \varrho \circ \pi_1^{-1}(\phi) \in U\}.$$

Since $\mathbb{C}^2$ is two-dimensional, these are precisely the subspaces of $\mathbb{C}^2$ spanned by a single nonzero element $(c_1, c_2) \in \mathbb{C}^2$ such that $\omega((c_1, c_2), (c_1, c_2)) = 0$, which is equivalent to $|c_1|^2 = |c_2|^2$. The condition $(c_1, c_2) \neq (0, 0)$ now forces $c_1 \neq 0 \neq c_2$ and $|c_2/c_1| = 1$, so there exists a unique $\theta \in [0, 2\pi[$ such that $c_2 = e^{i\theta}c_1$. Thus, if U is the linear span of $\{c_1, c_2\}$, then

$$\mathcal{D}(D_U) = \{\phi \in H^1(I) \colon \phi(b) = e^{i\theta}\phi(a)\},$$

and conversely, if the domain D_U is as above for some $\theta \in [0, 2\pi[$, then U is Lagrangian. ■

2.2 Symplectic forms and boundary triples

Theorem 2.1.7 and its proof may seem somewhat complicated compared to the result we were after, namely part (4) of the Theorem. Indeed, we could have obtained this result in a more straightforward manner. However, the theorem can be adapted to other situations with different operators, requiring only a few modifications. In this section, we shall develop some theory that will aid us in finding the self-adjoint extensions of differential operators, in partcular of the Hamiltonian H with a suitable potential V.

2.2.1 The endpoint space of an operator

In this subsection, we will return to the more general setting of a hermitian operator T acting on an arbitrary Hilbert space $\mathcal{H}$. We will see how symplectic forms naturally arise in the study of closed extensions of hermitian operators.

2.2.1 Definition. Let T be a hermitian operator on $\mathcal{H}$, let $J \colon \mathcal{H}^2 \to \mathcal{H}^2$ given by $(x, y) \mapsto (-y, x)$, and let

$$V_T := \mathcal{G}(T_{\max}) \cap \mathcal{G}(T)^\perp = \mathcal{G}(T_{\max}) \cap J(\mathcal{G}(T_{\max})).$$

We define the sesquilinear form $\omega_T \colon V_T \times V_T \to \mathbb{C}$ by

$$\omega_T(u, v) := \langle u, Jv\rangle_{\mathcal{H}^2}, \quad u, v \in V_T.$$

The pair (V_T, ω_T) is called the *endpoint space of* T.

2.2.2 Proposition. *Let T be a hermitian operator on $\mathcal{H}$.*

(1) V_T is a closed linear subspace of $\mathcal{H}^2$, and hence it is a Hilbert space with the inner product inherited from $\mathcal{H}^2$.

(2) ω_T is a continuous complex symplectic form on V_T.

(3) Let $P\colon \mathcal{H}^2 \to V_T$ be the orthogonal projection on V_T. Then for each $x, y \in \mathcal{D}(T_{\max})$,

$$\omega_T(P((x, T_{\max}x)), P((y, T_{\max}y))) = \langle T_{\max}x, y\rangle - \langle x, T_{\max}y\rangle.$$

(4) The map

$$S \mapsto V_T \cap \mathcal{G}(S),$$

yields a bijective correspondence between the realisations S of T and the closed subspaces U of V_T, with inverse

$$U \mapsto T_U := T_{\max}|_{\pi_1(\mathcal{G}(T_{\min})\oplus U)}.$$

(5) Let $U \subseteq V_T$ be a closed subspace and let T_U be the corresponding realisation of T. Then $T_U^* = T_{U^{\omega_T}}$. Consequently, T_U is hermitian if and only if U is isotropic, and T_U is self-adjoint if and only if U is Lagrangian.

Proof.
(1) $\mathcal{G}(T_{\max})$ is a closed linear subspace of $\mathcal{H}^2$ by part (3) of Proposition 1.4.6 and $J\colon \mathcal{H}^2 \to \mathcal{H}^2$ is unitary, so $V_T = \mathcal{G}(T_{\max}) \cap J(\mathcal{G}(T_{\max}))$ is a closed linear subspace of $\mathcal{H}^2$.

(2) $\langle\cdot,\cdot\rangle_{\mathcal{H}^2}$ is sesquilinear and J is linear, so ω_T is sesquilinear. Now let

$$u = (x, T_{\max}x), v = (y, T_{\max}y) \in V_T.$$

Then

$$\begin{aligned}\omega_T(u,v) &= \langle u, Jv\rangle_{\mathcal{H}^2} = \langle T_{\max}x, y\rangle - \langle x, T_{\max}y\rangle = -\overline{\langle T_{\max}y, x\rangle - \langle y, T_{\max}x\rangle}\\ &= -\overline{\langle v, Ju\rangle_{\mathcal{H}^2}} = -\overline{\omega_T(v,u)},\end{aligned}$$

so ω_T is skew-hermitian. V_T is an invariant subspace of J, so taking $v = -Ju$, we obtain $\omega_T(u,v) = \|u\|^2_{\mathcal{H}^2}$, which shows that ω_T is nondegenerate. Finally, since J is unitary, we can apply the Cauchy-Schwarz inequality to obtain

$$|\omega_T(u,v)| = |\langle u, Jv\rangle_{\mathcal{H}^2}| \le \|u\|_{\mathcal{H}^2}\cdot\|Jv\|_{\mathcal{H}^2} = \|u\|_{\mathcal{H}^2}\cdot\|v\|_{\mathcal{H}^2},$$

which shows that ω_T is continuous.

(3) Let $\pi_1 : \mathcal{H}^2 \to \mathcal{H}$ be the projection on the first coordinate. Let $x, y \in \mathcal{D}(T_{\max})$, let $x_1 := \pi_1 \circ P(x, T_{\max}x)$, let $x_2 := x - x_1$ and define y_1 and y_2 similarly. From the calculation in the previous part of the proof, we obtain

$$\omega_T(P(x, T_{\max}x), P(y, T_{\max}y)) = \langle T_{\max}x_1, y_1\rangle - \langle x_1, T_{\max}y_1\rangle.$$

On the other hand, since $T_{\min}$ is hermitian, we have

$$\begin{aligned}\langle T_{\max}x, y\rangle - \langle x, T_{\max}y\rangle &= \langle T_{\max}x_1, y_1\rangle - \langle x_1, T_{\max}y_1\rangle + \langle T_{\max}x_1, y_2\rangle - \langle x_1, T_{\max}y_2\rangle \\ &\quad + \langle T_{\max}x_2, y_1\rangle - \langle x_2, T_{\max}y_1\rangle + \langle T_{\max}x_2, y_2\rangle - \langle x_2, T_{\max}y_2\rangle \\ &= \langle T_{\max}x_1, y_1\rangle - \langle x_1, T_{\max}y_1\rangle + \langle T_{\max}x_1, y_2\rangle - \langle x_1, T_{\min}y_2\rangle \\ &\quad + \langle T_{\min}x_2, y_1\rangle - \langle x_2, T_{\max}y_1\rangle + \langle T_{\min}x_2, y_2\rangle - \langle x_2, T_{\max}y_2\rangle \\ &= \langle T_{\max}x_1, y_1\rangle - \langle x_1, T_{\max}y_1\rangle.\end{aligned}$$

This proves the identity.

(4) If S is a realisation of T, then $\mathcal{G}(S)$ is closed in $\mathcal{H}^2$, and hence closed in V_T. If U is a closed linear subspace of V_T, then it is a closed linear subspace of $\mathcal{H}^2$ because V_T is closed in $\mathcal{H}^2$. It follows that $\mathcal{G}(T_{\min}) \oplus U$ is closed in $\mathcal{H}^2$, so T_U is closed, and $\mathcal{G}(T_{\min}) \subseteq \mathcal{G}(T_{\min}) \oplus U \subseteq \mathcal{G}(T_{\min}) \oplus V_T = \mathcal{G}(T_{\max})$ implies that it is a realisation of T. Thus both maps are well defined. The two maps are inverses of each other, since $\mathcal{G}(S) = \mathcal{G}(T_{\min}) \oplus (V_T \cap \mathcal{G}(S))$ for any realisation S of T, and $U = V_T \cap (\mathcal{G}(T_{\min}) \oplus U)$ for any closed subspace $U \subseteq V_T$.

(5) Let S be a realisation of T, so that S^* is also a realisation of T. Then $\mathcal{G}(S) = \mathcal{G}(T_{\min}) \oplus (V_T \cap \mathcal{G}(S))$, and similarly, $\mathcal{G}(S^*) = \mathcal{G}(T_{\min}) \oplus (V_T \cap \mathcal{G}(S^*))$. Since J is unitary and since V_T is invariant under J, we have

$$J(\mathcal{G}(S^*)) = J(\mathcal{G}(T_{\min})) \oplus (V_T \cap J(\mathcal{G}(S^*))) = \mathcal{G}(T_{\max})^\perp \oplus (V_T \cap J(\mathcal{G}(S^*))),$$

by part (3) of Proposition 1.4.6 and Remark 2.1.2, so that, again by part (3) of 1.4.6, we obtain

$$\mathcal{H}^2 = \mathcal{G}(S) \oplus J(\mathcal{G}(S^*)) = \mathcal{G}(T_{\min}) \oplus (V_T \cap \mathcal{G}(S)) \oplus (V_T \cap J(\mathcal{G}(S^*))) \oplus \mathcal{G}(T_{\max})^\perp,$$

while on the other hand, we have $V_T = \mathcal{G}(T)^\perp \cap \mathcal{G}(T_{\max}) = \mathcal{G}(T_{\min})^\perp \cap \mathcal{G}(T_{\max})$, so that

$$\mathcal{H}^2 = \mathcal{G}(T_{\min}) \oplus V_T \oplus \mathcal{G}(T_{\max})^\perp.$$

Hence

$$V_T = (V_T \cap \mathcal{G}(S)) \oplus (V_T \cap J(\mathcal{G}(S^*))).$$

Now suppose $S = T_U$ for some closed subspace $U \subseteq V_T$. Then for $v \in V_T$, the following are equivalent:

- $v \in U^{\omega_T}$.
- $\omega_T(u, v) = 0$ for each $u \in U$.
- $\langle u, Jv\rangle_{\mathcal{H}^2} = 0$ for each $u \in U$.
- $Jv \in U^\perp$.
- $v \in \mathcal{G}(S^*)$.

We conclude that $U^{\omega_T} = V_T \cap \mathcal{G}(S^*)$, which proves that $T_U^* = T_{U^{\omega_T}}$, as desired. ■

2.2.3 Example. Consider the case where $\mathcal{H} = L^2([a, b])$ and $T = D$, and let ϱ and ω be the associated maps defined in Theorem 2.1.7. Since

$$\mathcal{G}(D_{\max}) = \mathcal{G}(D_{\min}) \oplus V_D = \mathcal{N}(\varrho) \oplus V_D,$$

the restriction $\varrho_D := \varrho|_{V_D}$ of ϱ to V_D is an isomorphism from V_D to $\mathbb{C}^2$, and by part (3) of Proposition 2.2.2, we have $\omega_D(u, v) = \omega(\varrho_D(u), \varrho_D(v))$ for each $u, v \in V_D$.

2.2.2 Boundary triples

We have shown that the self-adjoint extensions of a hermitian operator T on $\mathcal{H}$ correspond to the Lagrangian subspaces of its endpoint space (V_T, ω_T). This space is typically finite-dimensional when T is a differential operator on the space of test functions on some (possibly unbounded) interval, and later we will see that, given some 'nice' map ϱ as in the previous example, Proposition 2.2.2 provides an easy way of checking whether a given realisation of T is self-adjoint. However, it is not so easy to find *all* Lagrangian subspaces of an arbitrary complex symplectic vector space, and therefore, to find all self-adjoint extensions of T.

Fortunately, there is another way to approach this problem, namely through the method of boundary triples, and in some cases, this method *does* provide a way to generate all self-adjoint extensions of a given hermitian operator. Here, we shall make use of the endpoint space of an operator to establish some of the results concerning these objects. For different approaches, we refer to [19, chapter 14] or [4, section 7.1].

2.2.4 Proposition. *Let (V, ω) be a complex symplectic vector space. For $j = 1, 2$, let $(\mathcal{H}_j, \langle\cdot,\cdot\rangle_j)$ be Hilbert spaces and let $\sigma_j \colon V \to \mathcal{H}_i$ be linear maps. Moreover, assume that the map $\sigma_1 \oplus \sigma_2 \colon V \to \mathcal{H}_1 \times \mathcal{H}_2$, given by*

$$u \mapsto (\sigma_1(u), \sigma_2(u))$$

is surjective, and that there exists a constant $c \in \mathbb{C}\backslash\{0\}$ such that

$$\omega(u, v) = c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_2(u), \sigma_2(v)\rangle_2),$$

for each $u, v \in V$.

(1) Let $\mathcal{U} \colon \mathcal{H}_1 \to \mathcal{H}_2$ is a unitary map. Then the subspace $U \subseteq V$ given by

$$U := \{u \in V \colon \sigma_2(u) = \mathcal{U} \circ \sigma_1(u)\}$$

is a Lagrangian subspace of V.

(2) Conversely, let $U \subseteq V$ be a Lagrangian subspace. Then there exists a unique unitary map $\mathcal{U} \colon \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$(*) \qquad U = \{u \in V \colon \sigma_2(u) = \mathcal{U} \circ \sigma_1(u)\}.$$

Proof.
(1) Let $v \in U$. Then for each $u \in U$, we have

$$\begin{aligned}\omega(u, v) &= c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_2(u), \sigma_2(v)\rangle_2) \\ &= c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \mathcal{U} \circ \sigma_1(u), \mathcal{U} \circ \sigma_1(v)\rangle_2) \\ &= c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_1(u), \sigma_1(v)\rangle_1) = 0,\end{aligned}$$

since $\mathcal{U}$ is unitary. Thus U is isotropic.

To show that U is Lagrangian, we note that $\sigma_2(U) = \mathcal{H}_2$. Indeed, let $x \in \mathcal{H}_2$. Since $\sigma_1 \oplus \sigma_2$ is surjective, there exists a $u \in V$ such that $(\sigma_1(u), \sigma_2(u)) = (\mathcal{U}^{-1}x, x)$. Clearly, this implies $\mathcal{U} \circ \sigma_1(u) = \sigma_2(u)$, so $u \in U$ and hence $x \in \sigma_2(U)$.

Now let $v \in U^\omega$. Then for each $u \in U$, we have

$$\begin{aligned} 0 = \frac{1}{c}\omega(u,v) &= \langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_2(u), \sigma_2(v)\rangle_2 \\ &= \langle \mathcal{U} \circ \sigma_1(u), \mathcal{U} \circ \sigma_1(v)\rangle_2 - \langle \sigma_2(u), \sigma_2(v)\rangle_2 \\ &= \langle \sigma_2(u), \mathcal{U} \circ \sigma_1(v) - \sigma_2(v)\rangle_2, \end{aligned}$$

and since $\sigma_2(U) = \mathcal{H}_2$, it follows that $\mathcal{U} \circ \sigma_1(v) = \sigma_2(v)$. Thus $v \in U$, which implies that U is Lagrangian, as desired.

(2) Since U is Lagrangian, it is isotropic, so for each $u \in U$, we have

$$(**) \qquad 0 = \frac{1}{c}\omega(u,u) = \|\sigma_1(u)\|_1 - \|\sigma_2(u)\|_2,$$

so $\sigma_1(u) = 0$ if and only if $\sigma_2(u) = 0$, which implies $\mathcal{N}(\sigma_1|_U) = \mathcal{N}(\sigma_2|_U)$. Thus the map $\sigma_2|_U$ factors through $U/\mathcal{N}(\sigma_1|_U)$. Since $U/\mathcal{N}(\sigma_1|_U) \simeq \mathcal{R}(\sigma_1|_U) = \sigma_1(U)$, there exists a unique linear map $\widetilde{\mathcal{U}} \colon \sigma_1(U) \to \sigma_2(U)$ such that $\sigma_2|_U = \widetilde{\mathcal{U}} \circ \sigma_1|_U$. Here, $\widetilde{\mathcal{U}}$ is clearly surjective, and from equation $(**)$, it follows that $\widetilde{\mathcal{U}}$ is an isometry, so it is a unitary map.

We claim that $\sigma_1(U)$ and $\sigma_2(U)$ are dense in $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Let $x \in \sigma_1(U)^\perp$. Since $\sigma_1 \oplus \sigma_2$ is surjective, there exists a $v \in V$ such that $(\sigma_1(v), \sigma_2(v)) = (x, 0)$. Then for each $u \in U$, we have

$$\omega(u,v) = c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_2(u), \sigma_2(v)\rangle_2) = c(\langle \sigma_1(u), x\rangle_1 - \langle \sigma_2(u), 0\rangle_2) = 0,$$

so $v \in U^\omega$. The subspace U is Lagrangian, so $v \in U$, which implies that $x \in \sigma_1(U)$. We have $x \in \sigma_1(U)^\perp$ by assumption, so $x = 0$. Hence $\sigma_1(U)^\perp = \{0\}$, and it follows that $\sigma_1(U)$ is dense in $\mathcal{H}_1$. A similar argument shows that $\sigma_2(U)$ is dense in $\mathcal{H}_2$.

By Lemma 1.3.5, $\widetilde{\mathcal{U}}$ has a unique unitary extension $\mathcal{U} \colon \mathcal{H}_1 \to \mathcal{H}_2$. We show that $(*)$ holds. Note that by definition of $\mathcal{U}$, the subspace U is contained in the right-hand side. Conversely, suppose $v \in V$ satisfies $\sigma_2(v) = \mathcal{U} \circ \sigma_1(v)$. Then for each $u \in U$, we have

$$\begin{aligned} \omega(u,v) &= c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_2(u), \sigma_2(v)\rangle_2) \\ &= c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \mathcal{U} \circ \sigma_1(u), \mathcal{U} \circ \sigma_1(v)\rangle_2) \\ &= c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_1(u), \sigma_1(v)\rangle_1) = 0, \end{aligned}$$

since $\mathcal{U}$ is unitary. Hence $v \in U^\omega$, so $v \in U$ because U is Lagrangian. This proves $(*)$. Since $\sigma_1(U)$ and $\sigma_2(U)$ are dense in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, $\mathcal{U}$ is the only unitary map $\mathcal{H}_1 \to \mathcal{H}_2$ that satisfies $(*)$. ■

2.2.5 Corollary. *Let T be a hermitian operator on a hilbert space $\mathcal{H}$, let (V_T, ω_T) be its endpoint space, and for $j = 1, 2$, let $(\mathcal{H}_j, \langle \cdot, \cdot \rangle_j)$ be Hilbert spaces and $\sigma_j \colon V_T \to \mathcal{H}_j$ be linear maps. Moreover, assume that $\sigma_1 \oplus \sigma_2$ is surjective, and that there exists a constant $c \in \mathbb{C}\backslash\{0\}$ such that*

$$\omega_T(u,v) = c(\langle \sigma_1(u), \sigma_1(v)\rangle_1 - \langle \sigma_2(u), \sigma_2(v)\rangle_2),$$

for each $u, v \in V_T$. Then there exists a bijective correspondence between the self-adjoint realisations $\widetilde{T}$ of T and the unitary maps $\mathcal{U}\colon \mathcal{H}_1 \to \mathcal{H}_2$, given by

$$\mathcal{U} \mapsto T^*|_{\pi_1(\mathcal{G}(T_{\min})\oplus U)},$$

where

$$U := \{u \in V_T \colon \sigma_2(u) = \mathcal{U} \circ \sigma_1(u)\}.$$

Proof. This follows from Propositions 2.2.2 and 2.2.4. ■

2.2.6 Definition. Let T be a hermitian operator on a Hilbert space $\mathcal{H}$, let (V_T, ω_T) be its endpoint space, let $(\mathcal{H}_1, \langle\cdot,\cdot\rangle_1)$ be a Hilbert space, and for $j = 1, 2$, let $\sigma_j \colon V_T \to \mathcal{H}_1$ be linear maps. Moreover, assume that $\sigma_1 \oplus \sigma_2$ is surjective, and that there exists a constant $c \in \mathbb{C}\backslash\{0\}$ such that

$$\omega_T(u, v) = c(\langle\sigma_1(u), \sigma_1(v)\rangle_1 - \langle\sigma_2(u), \sigma_2(v)\rangle_2).$$

Then $(\mathcal{H}_1, \sigma_1, \sigma_2)$ is called a *boundary triple for T*.

2.2.7 Example. *(Continuation of Example 2.2.3)* Define the linear maps $\sigma_1, \sigma_2 \colon V_D \to \mathbb{C}$ by $\sigma_1(\phi, D_{\max}\phi) := \phi(a)$, $\sigma_2(\phi, D_{\max}\phi) := \phi(b)$. Then for each $u, v \in V_T$, we have

$$\omega_T(u, v) = -i(\langle\sigma_1(u), \sigma_1(v)\rangle_{\mathbb{C}} - \langle\sigma_2(u), \sigma_2(v)\rangle_{\mathbb{C}}).$$

(Note that $\langle z_1, z_2\rangle_{\mathbb{C}}$ is just a fancy way of writing $\overline{z_1}z_2$.) Moreover, $\sigma_1 \oplus \sigma_2 = \varrho_T$, and ϱ_T is surjective because ϱ is surjective, so $(\mathbb{C}, \sigma_1, \sigma_2)$ is a boundary triple for D. Thus the self-adjoint extensions of D are precisely the extensions $\widetilde{D}$ such that

$$\mathcal{G}(\widetilde{D}) = \mathcal{G}(D_{\min}) \oplus \{u \in V_T \colon \sigma_2(u) = \mathcal{U} \circ \sigma_1(u)\},$$

for some unitary map $\mathcal{U}\colon \mathbb{C} \to \mathbb{C}$. Now a map $\mathcal{U}\colon \mathbb{C} \to \mathbb{C}$ is unitary if and only if it is of the form $z \mapsto e^{i\theta}z$ for some $\theta \in [0, 2\pi[$. Applying the projection onto the first coordinate to both sides of the above equation, we obtain

$$\mathcal{D}(\widetilde{D}) = \{\phi \in H^1(a, b) \colon \phi(b) = e^{i\theta}\phi(a)\},$$

for some $\theta \in [0, 2\pi[$, since $\mathcal{D}(D_{\min}) = \{\phi \in H^1(]a, b[) \colon \phi(a) = 0 = \phi(b)\}$. This is exactly what we had already found in part (4) of Theorem 2.1.7.

The main reason that we did not assume that $\mathcal{H}_1 = \mathcal{H}_2$ in Proposition 2.2.4, is that it allows us to prove Theorem 2.2.9, a result by John von Neumann. We first need the following lemma, however:

2.2.8 Lemma. *Let A be a linear operator on $\mathcal{H}$. If $\mathcal{D}(A^2) = \mathcal{D}(A)$ and $A^2x = -x$ for all $x \in \mathcal{D}(A)$, then $\mathcal{G}(A) = V_+ \oplus V_-$, where*

$$V_\pm := \{(x, y) \in \mathcal{G}(A) \colon y = \pm ix\}.$$

Proof. If $v = (y, iy) \in V_+$ and $w = (z, -iz) \in V_-$, then

$$\langle v, w\rangle_{\mathcal{H}^2} = \langle y, z\rangle + \langle iy, -iz\rangle = \langle y, z\rangle - \langle y, z\rangle = 0,$$

so V_+ and V_- are mutually orthogonal linear subspaces. Now suppose $(x, Ax) \in \mathcal{G}(A)$. Define

$$y := \frac{1}{2}(x - iAx), \quad z := \frac{1}{2}(x + iAx) \in \mathcal{D}(A).$$

Then

$$Ay = \frac{1}{2}(Ax - iA^2x) = \frac{1}{2}(Ax + ix) = \frac{1}{2}i(x - iAx) = iy,$$

and similarly, we see that $Az = -iz$. Since $x = y + z$, we have

$$(x, Ax) = (y, Ay) + (z, Az) \in V_+ \oplus V_-,$$

which proves the lemma. ■

Here is the main result:

2.2.9 Theorem. *Let T be a hermitian operator on $\mathcal{H}$, let I be the identity on $\mathcal{H}$, let*

$$K_\pm := \mathcal{N}(T_{\max} \mp iI),$$

and let (V_T, ω_T) be the endpoint space of T.

(1) Let $V_\pm := \{(x, T_{\max}x) \in \mathcal{G}(T_{\max})\colon x \in K_\pm\} = \mathcal{G}(T_{\max}|_{K_\pm})$. Then we have

$$V_T = V_+ \oplus V_-.$$

In particular, V_+ and V_- are closed subspaces of the Hilbert space V_T.

(2) If $\mathcal{U}\colon K_+ \to K_-$ is a unitary map, then the operator $S \subseteq T_{\max}$ with domain

$$\mathcal{D}(S) := \{x + x' + \mathcal{U}x'\colon x \in \mathcal{D}(T_{\min}),\ x' \in K_+\},$$

is a self-adjoint realisation of T.

Conversely, if S is a self-adjoint realisation of T, then there exists a unique unitary map $\mathcal{U}\colon K_+ \to K_-$ such that

$$\mathcal{D}(S) = \{x + x' + \mathcal{U}x'\colon x \in \mathcal{D}(T_{\min}),\ x' \in K_+\}.$$

Proof.
(1) Let $\pi_1\colon \mathcal{H}^2 \to \mathcal{H}$ be the projection on the first coordinate, and let A be the restriction of $T_{\max}$ to $\mathcal{D}(A) := \pi_1(V_T)$. Recall that $\mathcal{G}(T_{\max}) = \mathcal{G}(T_{\min}) \oplus V_T$ and that $T_{\min} = T^*_{\max}$.

Next, let $x \in \mathcal{D}(A)$. Then for each $y \in \mathcal{D}(T_{\min})$, we have

$$0 = \langle (x, T_{\max}x), (y, T_{\min}y)\rangle_{\mathcal{H}^2} = \langle x, y\rangle + \langle T_{\max}x, T_{\min}y\rangle,$$

so $\langle T_{\max}x, T_{\min}y\rangle = \langle -x, y\rangle$. Thus $T_{\max}x \in \mathcal{D}(T^*_{\min}) = \mathcal{D}(T_{\max})$ and $(T_{\max})^2x = -x$. Moreover, $(T_{\max}x, -x) = -J(x, T_{\max}x) \in J(\mathcal{G}(T_{\max}))$, so $(Ax, -x) = (T_{\max}x, -x) \in V_T$.

Hence $\mathcal{D}(A^2) = \mathcal{D}(A)$, and for each $x \in \mathcal{D}(A)$, we have $A^2 x = -x$. By Lemma 2.2.8, the spaces $V'_\pm := \{(x, y) \in \mathcal{G}(A) : y = \pm ix\}$ satisfy $\mathcal{G}(A) = V'_+ \oplus V'_-$. Thus we have

$$\mathcal{G}(T_{\max}) = \mathcal{G}(T_{\min}) \oplus V'_+ \oplus V'_-.$$

It is evident that $V'_\pm \subseteq V_\pm$. The same argument that we used to prove that V'_+ and V'_- are orthogonal can be used to show that V_+ and V_- are orthogonal, so to prove the first part of the theorem, it suffices to show that the spaces $V_\pm$ are orthogonal to $\mathcal{G}(T_{\min})$. We shall show this for V_+; the argument for V_- is similar. Let $x \in \mathcal{N}(T_{\max} - iI)$. Then for each $y \in \mathcal{D}(T_{\min})$,

$$\begin{aligned}
\langle (x, T_{\max}x), (y, T_{\min}y) \rangle_{\mathcal{H}^2} &= \langle x, y \rangle + \langle T_{\max}x, T_{\min}y \rangle = \langle x, y \rangle + \langle ix, T_{\min}y \rangle \\
&= \langle x, y \rangle - i\langle T_{\max}x, y \rangle = \langle x, y \rangle - i\langle ix, y \rangle \\
&= \langle x, y \rangle - \langle x, y \rangle = 0,
\end{aligned}$$

so V_+ is indeed orthogonal to $\mathcal{G}(T_{\min})$.

(2) Let $P_\pm : V_T \to V_\pm$ be the orthogonal projection, and as before, let $\pi_1 : \mathcal{H}^2 \to \mathcal{H}$ be the projection on the first coordinate. Note that the orthogonal projections exist by part (1) of the theorem. Clearly, the maps $\pi_1 \circ P_\pm : V_T \to K_\pm$ are linear. Since

$$K_\pm = \mathcal{N}(T_{\max} \mp iI) = \mathcal{R}(T_{\min} \pm iI)^\perp,$$

the spaces $K_\pm$ are closed subspaces of $\mathcal{H}$, so they are Hilbert spaces together with the inherited inner product. The map

$$(\pi_1 \circ P_+) \oplus (\pi_1 \circ P_-) : V_T \to K_+ \times K_-,$$

is surjective, and a right inverse for this map is given by

$$(x_+, x_-) \mapsto (x_+, T_{\max}x_+) + (x_-, T_{\max}x_-) = (x_+ + x_-, i(x_+ - x_-)),$$

since $V_T = V_+ \oplus V_-$. Finally, let $x, y \in \pi_1(V_T)$, let $x_\pm := \pi_1 \circ P_\pm(x, T_{\max}x)$ and let $y_\pm := \pi_1 \circ P_\pm(y, T_{\max}y)$. Then

$$\begin{aligned}
&\omega_T((x, T_{\max}x), (y, T_{\max}y)) \\
&\quad = \langle (x, T_{\max}x), (-T_{\max}y, y) \rangle_{\mathcal{H}^2} \\
&\quad = \langle (x_+, ix_+), (-iy_+, y_+) \rangle_{\mathcal{H}^2} + \langle (x_+, ix_+), (iy_-, y_-) \rangle_{\mathcal{H}^2} \\
&\qquad + \langle (x_-, -ix_-), (-iy_+, y_+) \rangle_{\mathcal{H}^2} + \langle (x_-, -ix_-), (iy_-, y_-) \rangle_{\mathcal{H}^2} \\
&\quad = i(\langle (x_+, ix_+), (y_+, iy_+) \rangle_{\mathcal{H}^2} - \langle (x_-, -ix_-), (y_-, -iy_-) \rangle_{\mathcal{H}^2}) \\
&\quad = 2i(\langle x_+, y_+ \rangle - \langle x_-, y_- \rangle),
\end{aligned}$$

where in the third step, we used $V_+ \perp V_-$. Applying Corollary 2.2.5 with $\sigma_1 = \pi_1 \circ P_+$ and $\sigma_2 = \pi_1 \circ P_-$, we conclude that for any unitary map $\mathcal{U} : K_+ \to K_-$, the restriction of $T_{\max}$ to $\mathcal{G}(T_{\min}) \oplus U$ with

$$U := \{u \in V_T : \pi_1 \circ P_-(u) = \mathcal{U} \circ \pi_1 \circ P_+(u)\},$$

is a self-adjoint realisation of T, and conversely, that for any self-adjoint realisation S of T, there exists a unique unitary map $\mathcal{U} : K_+ \to K_-$ such that $\mathcal{G}(S) = \mathcal{G}(T_{\min}) \oplus U$, with U as above. Applying π_1 to both sides of the identity $\mathcal{G}(S) = \mathcal{G}(T_{\min}) \oplus U$ yields the result. ∎

2.2.10 Definition. The spaces $K_\pm$ in the above theorem are called the *deficiency subspaces of T*. The Hilbert dimensions of these spaces are called the *deficiency indices of T*.

2.2.11 Corollary. *Let T be a hermitian operator on a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$.*

(1) T has self-adjoint extensions if and only if T has equal deficiency indices.

(2) T is essentially self-adjoint if and only if the deficiency indices of T are both equal to 0.

This corollary is relevant to our cause, because it allows us to prove the existence of a self-adjoint realisation of certain differential operators, including D and H. It is used to prove the following result, which is also due to John von Neumann, (cf. [4, Proposition 2.2.16]).

2.2.12 Proposition. *Let T be a hermitian operator on a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ and let $\mathcal{C}\colon \mathcal{H} \to \mathcal{H}$ be an antilinear isometry such that $\mathcal{C}^2 = I_{\mathcal{H}}$. If $\mathcal{C}(\mathcal{D}(T)) \subseteq \mathcal{D}(T)$ and if $\mathcal{C}$ commutes with T, then T has equal deficiency indices. Equivalently, T has a self-adjoint extension.*

2.2.13 Corollary. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set, and let T be a formally self-adjoint differential operator with domain $\mathcal{D}(T) = C_0^\infty(\Omega) \subset L^2(\Omega)$. Then T has a self-adjoint extension.*

Proof. Let $\mathcal{C}\colon L^2(\Omega) \to L^2(\Omega)$ be the map given by $\phi \mapsto \overline{\phi}$ and apply the previous proposition. ■

2.3 The Hamiltonian $H = -\frac{d^2}{dx^2} + V$

2.3.1 Hamiltonians with regular endpoints

Now we will turn our attention to the operator $H = -\frac{d^2}{dx^2} + V$ and adapt the necessary and sufficient criterion for realisations of D to be self-adjoint, obtained in part (3) of Theorem 2.1.7, to H. In addition, we will use the theory of boundary triples to obtain all self-adjoint extensions. However, we have to put some extra condition on V besides the fact that it is an element of $L^2_{\text{loc}}(]a,b[)$.

2.3.1 Definition. Let H be a Hamiltonian on $]a,b[\subset \mathbb{R}$ with $V \in L^2_{\text{loc}}(]a,b[)$.

- The endpoint a is said to be *regular for H* iff there exists a $c \in]a,b[$ such that the limit $\lim_{d\to a^+} \int_d^c |V(x)|\,dx$ exists. Similarly, the endpoint b is said to be a regular endpoint for H iff there exists a $c \in]a,b[$ such that the limit $\lim_{d\to b^-} \int_c^d |V(x)|\,dx$ exists.
- If the endpoint a is not regular, then a is said to be *singular*; likewise for b.

2.3.2 Remark. Since $L^2_{\text{loc}}(]a,b[) \subset L^1_{\text{loc}}(]a,b[)$, it follows from Levi's monotone convergence theorem that a is a regular endpoint for H if and only if there exists a $c \in]a,b[$ such that $|V| \in L^1(a,c)$. A similar statement can be formulated for the endpoint b. In particular, if both a and b are regular endpoints of V and $a,b \in \mathbb{R}$, then $|V| \in L^1(]a,b[)$, and using Lebesgue's dominated convergence theorem, one can show that $V \in L^1(]a,b[)$. Thus $V \in L^1(]a,b[)$ if (and only if) both a and b are regular endpoints for H.

In addition, let us recall the following notion:

2.3.3 Definition. Let $I \subseteq \mathbb{R}$ be a (possibly unbounded) interval. A function $f\colon I \to \mathbb{C}$ is said to be *absolutely continuous* iff there exists a $g \in L^1_{\mathrm{loc}}(I)$ such that for each $c, d \in I$ with $c < d$, we have

$$f(d) - f(c) = \int_c^d g(x)\,dx.$$

The space of all such functions is denoted by $\mathrm{AC}(I)$.

2.3.4 Remark. Using Lebesgue's dominated convergence theorem, it is easily seen that the elements of $\mathrm{AC}(I)$ are continuous; this justifies the name of the space.

In order to be able to use the same method as in Theorem 2.1.7, we require the following result, which is a summary of Proposition 2.3.20 and some of the results in section 7.2 of [4]. In what follows, let $\langle \cdot, \cdot \rangle$ be the standard inner product on $L^2(]a, b[)$.

2.3.5 Theorem. *Let $a, b \in \mathbb{R}$, $a < b$ and let H_{dist} be the operator on the space of distributions $\mathscr{D}'(]a, b[)$ on $]a, b[$ associated to H. Then:*

(1) The domain of the maximal realisation of H is given by

$$\mathcal{D}(H_{\max}) = \{\phi \in L^2(]a, b[)\colon \phi, \phi' \in \mathrm{AC}(]a, b[),\ H_{\mathrm{dist}}\phi \in L^2(]a, b[)\}.$$

This is also true whenever $a = -\infty$ or $b = \infty$.

For each $x \in]a, b[$ and $\phi, \psi \in \mathcal{D}(H_{\max})$, let $W_x(\psi, \phi) := \overline{\psi(x)}\phi'(x) - \overline{\psi'(x)}\phi(x)$.

(2) For all $c, d \in]a, b[$ and $\phi, \psi \in \mathcal{D}(H_{\max})$, we have

$$\int_c^d \overline{(H_{\max}\psi)(x)}\phi(x) - \overline{\psi(x)}(H_{\max}\phi)(x)\,dx = W_d(\psi, \phi) - W_c(\psi, \phi).$$

Moreover, the limits

$$W_a(\psi, \phi) := \lim_{c \to a^+} W_c(\psi, \phi);$$
$$W_b(\psi, \phi) := \lim_{d \to b^-} W_d(\psi, \phi),$$

exist, and

$$\langle H_{\max}\psi, \phi \rangle - \langle \psi, H_{\max}\phi \rangle = W_b(\psi, \phi) - W_a(\psi, \phi).$$

(3) The domain of the minimal realisation of H is given by

$$\mathcal{D}(H_{\min}) = \{\phi \in \mathcal{D}(H_{\max})\colon W_a(\psi, \phi) = 0 = W_b(\psi, \phi) \text{ for each } \psi \in \mathcal{D}(H_{\max})\}.$$

(4) The deficiency indices of H are both at most equal to 2.

Now suppose that both a and b are regular endpoints of H. Then:

(5) The domain of the maximal realisation of H is given by

$$\mathcal{D}(H_{\max}) = \{\phi \in L^2(]a,b[) \colon \phi, \phi' \in \mathrm{AC}[a,b],\ H_{\mathrm{dist}}\phi \in L^2(]a,b[)\},$$

i.e. the elements ϕ and ϕ' of $L^2(]a,b[)$ have absolutely continuous representatives on $]a,b[$ that can be extended to absolutely continuous functions on $[a,b]$.

(6) The domain of the minimal realisation of H is given by

$$\mathcal{D}(H_{\min}) = \{\phi \in \mathcal{D}(H_{\max}) \colon \phi(a) = \phi'(a) = 0 = \phi(b) = \phi'(b)\}.$$

(7) The deficiency indices of H are both equal to 2*.*

2.3.6 Remark. In some of the literature on quantum mechanics, the quantity

$$W_x(\psi, \phi) = \overline{\psi(x)}\phi'(x) - \overline{\psi'(x)}\phi(x),$$

defined in the previous theorem is called the *Wronskian of ψ and ϕ at x*.

Similar to Theorem 2.1.7, we have the following theorem:

2.3.7 Theorem. *Assume that H is the operator on $L^2(a,b)$ as above, and that a and b are regular endpoints for H.*

(1) Let $\varrho \colon \mathcal{G}(H_{\max}) \to \mathbb{C}^4$ be the map given by $(\phi, H_{\max}\phi) \mapsto (\phi(a), \phi'(a), \phi(b), \phi'(b))$. Then ϱ is linear, continuous and surjective, and $\mathcal{N}(\varrho) = \mathcal{G}(H_{\min})$.

(2) Let $\pi_1 \colon \mathcal{G}(H_{\max}) \to \mathcal{D}(H_{\max})$ be the projection on the first coordinate. A bijective correspondence between realisations $\widetilde{H}$ of H and linear subspaces U of $\mathbb{C}^4$ is given by

$$\widetilde{H} \mapsto \varrho(\mathcal{G}(\widetilde{H})),$$

with inverse

$$U \mapsto H_{\max}|_{\pi_1(\varrho^{-1}(U))}.$$

(3) Let ω be the complex symplectic form on $\mathbb{C}^4$ given by $\omega(\mathbf{c}, \mathbf{d}) := \mathbf{c}^* B\mathbf{d}$, where

$$B := \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix},$$

and $\mathbf{c}$ and $\mathbf{d}$ are column vectors. For each linear subspace $U \subseteq \mathbb{C}^4$, let H_U be the realisation of H associated to U by the correspondence described in (2). Then $H_U^* = H_{U^\omega}$. In particular,

- *The hermitian realisations of H correspond to the isotropic subspaces of $\mathbb{C}^4$;*
- *The selfadjoint realisations of H correspond to the Lagrangian subspaces of $\mathbb{C}^4$.*

(4) Let (V_H, ω_H) be the endpoint space of H, and let $\varrho_H \colon V_H \to \mathbb{C}^4$ be the restriction of ϱ to V_H. Then for each $u, v \in V_H$, we have

$$\omega_H(u, v) = \omega(\varrho_H(u), \varrho_H(v)).$$

Proof. We will only prove parts (1) and (4); the remaining parts of the theorem can be proved in the same way as parts (2) and (3) of Theorem 2.1.7. It is clear that ϱ is linear. By part (6) of Theorem 2.3.5, we have $\mathcal{N}(\varrho) = \mathcal{G}(H_{\min})$, which is closed, so ϱ is continuous by Lemma 2.1.6.

It remains to be shown that ϱ is surjective. By part (6) of Theorem 2.3.5, we have $\mathcal{G}(H_{\min}) = \mathcal{N}(\varrho)$, so

$$\mathcal{G}(H_{\max}) = \mathcal{G}(H_{\min}) \oplus V_H = \mathcal{N}(\varrho) \oplus V_H,$$

which implies that ϱ_H is injective. It follows from part (7) of Theorem 2.3.5 and part (1) of Theorem 2.2.9 that $\dim(V_H) = 2 + 2 = 4$, so ϱ_H is an isomorphism, and consequently, ϱ is surjective. This proves part (1).

Part (4) of the theorem follows from parts (2) and (5) of Theorem 2.3.5 and part (3) of Proposition 2.2.2. ∎

2.3.8 Example. As in the case of D, it is possible to find a boundary triple for H. Because ϱ_H is an isomorphism from V_H to $\mathbb{C}^4$, we are going to look for linear maps $\sigma_1, \sigma_2 \colon V_H \to \mathcal{H}_1$ such that $(\mathcal{H}_1, \sigma_1, \sigma_2)$ is a boundary triple for H, where the Hilbert space $\mathcal{H}_1$ is $\mathbb{C}^2$ equipped with its standard inner product.

Let ω be the symplectic form from the previous theorem and let B be its associated matrix. We note that B has two eigenvalues, namely $\pm i$, whose corresponding eigenspaces are spanned by the orthonormal vectors

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \mp i \\ 0 \\ 0 \end{pmatrix}, \quad \text{and} \quad \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 0 \\ 1 \\ \pm i \end{pmatrix}.$$

It follows that the matrix $\widetilde{A}$, given by

$$\widetilde{A} := \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 1 & 0 \\ -i & 0 & i & 0 \\ 0 & 1 & 0 & 1 \\ 0 & i & 0 & -i \end{pmatrix},$$

is unitary, and that

$$B = \widetilde{A}\Lambda\widetilde{A}^{-1} = \widetilde{A}\Lambda\widetilde{A}^*,$$

where $\Lambda = i \cdot \mathrm{diag}(1, 1, -1, -1)$. Let $A := \widetilde{A}^*$. Then

$$\omega(\mathbf{c}, \mathbf{d}) = \mathbf{c}^* A^* \Lambda A \mathbf{d} = (A\mathbf{c})^* \Lambda (A\mathbf{d}).$$

Now let $P_1, P_2 \colon \mathbb{C}^4 \to \mathbb{C}^2$ be the projections on the first two and the last two coordinates respectively. Then $P_1 \oplus P_2 \colon \mathbb{C}^4 \to \mathbb{C}^4$ is surjective, and

$$\omega(\mathbf{c}, \mathbf{d}) = (A\mathbf{c})^* \Lambda (A\mathbf{d}) = i(\langle P_1 \circ A(\mathbf{c}), P_1 \circ A(\mathbf{d})\rangle_{\mathbb{C}^2} - \langle P_2 \circ A(\mathbf{c}), P_2 \circ A(\mathbf{d})\rangle_{\mathbb{C}^2}),$$

so if we set $\sigma_j := P_j \circ A \circ \varrho_H$ for $j = 1, 2$, (identifying the matrix A with the corresponding unitary map relative to the standard basis of $\mathbb{C}^4$), then it follows from part (4) of Theorem 2.3.7 and the fact that ϱ_H and A are isomorphisms, that $(\mathbb{C}^2, \sigma_1, \sigma_2)$ is a boundary triple for H. A straightforward computation shows that

$$\sigma_1(\phi, H_{\max}\phi) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi(a) + i\phi'(a) \\ \phi(b) - i\phi'(b) \end{pmatrix} \text{ and } \sigma_2(\phi, H_{\max}\phi) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi(a) - i\phi'(a) \\ \phi(b) + i\phi'(b) \end{pmatrix},$$

so that, following the same line of reasoning as in Example 2.2.7, the self-adjoint realisations $\widetilde{H}$ of H correspond to the unitary 2×2-matrices M via

$$\mathcal{D}(\widetilde{H}) = \left\{ \phi \in \mathcal{D}(H_{\max}) \ : \ \begin{pmatrix} \phi(a) + i\phi'(a) \\ \phi(b) - i\phi'(b) \end{pmatrix} = M \begin{pmatrix} \phi(a) - i\phi'(a) \\ \phi(b) + i\phi'(b) \end{pmatrix} \right\}.$$

A complex 2×2-matrix M is unitary if and only if there exist $c, d \in \mathbb{C}$ with $|c|^2 + |d|^2 = 1$ and $\theta \in [0, 2\pi[$ such that

$$M = e^{i\theta} \begin{pmatrix} c & -\overline{d} \\ d & \overline{c} \end{pmatrix}.$$

Thus we have obtained an explicit parametrisation of the self-adjoint realisations of H.

2.3.9 Example. *(Continuation of Example 2.3.8)* For $j = 1, 2, 3, 4$, let e_j be the j-th standard basis vector of $\mathbb{C}^4$. We borrow the following examples of boundary conditions which yield selfadjoint extensions from [4]:

(1) *Dirichlet*: $\phi(a) = \phi(b) = 0$. The corresponding subspace $U \subset \mathbb{C}^4$ is spanned by e_2 and e_4, which is two-dimensional and isotropic with respect to ω, and therefore a Lagrangian subspace of $(\mathbb{C}^4, \omega)$ (apply part (3) of Proposition 2.1.5). The corresponding unitary matrix is $-I$.

(2) *Neumann*: $\phi'(a) = \phi'(b) = 0$. The corresponding Lagrangian subspace is spanned by e_1 and e_3, and the corresponding unitary matrix is I.

(3) *Periodic*: $\phi(a) = \phi(b)$, $\phi'(a) = \phi'(b)$. The corresponding Lagrangian subspace is spanned by $e_1 + e_3$ and $e_2 + e_4$, and the corresponding unitary matrix is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(4) *Antiperiodic*: $\phi(a) = -\phi(b)$, $\phi'(a) = -\phi'(b)$. The corresponding Lagrangian subspace is spanned by $e_1 - e_3$ and $e_2 - e_4$, and the corresponding unitary matrix is

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

2.3.10 Remark. It is worth noting that the self-adjoint extensions of the 'actual' Hamiltonian $H_\hbar := -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V$ are characterised by the same boundary conditions as the self-adjoint extensions of the operator $H = -\frac{d^2}{dx^2} + V$. Indeed, we have $V_{H_\hbar} = V_H$, and if we define ϱ_H and ω as in Theorem 2.3.7, and $\omega_\hbar$ is the sympectic form with the property that $\omega_{H_\hbar}(u,v) = \omega_\hbar(\varrho_H(u), \varrho_H(v))$ for each $u, v \in V_H$, then

$$\omega_\hbar = \frac{\hbar^2}{2m}\omega,$$

which means that a subspace of V_H is isotropic/Lagrangian with respect to $\omega_\hbar$ if and only if it is isotropic/Lagrangian with respect ω.

2.3.2 The free particle

Before turning to some singular examples, let us mention one specific example of a Hamiltonian on a bounded interval I with regular endpoints, namely the one with $V = 0$, so that $H = D^2$. This is the Hamiltonian of a free particle moving on a bounded interval. We will compute its self-adjoint extensions for the case that I is unbounded. However, let us first state the following fact, which can be found as Theorem 4.23 in [10].

2.3.11 Theorem. *Let $m \in \mathbb{N}_0$, and let I be an interval (possibly unbounded). Then $\mathcal{D}((D^m)_{\max}) = H^m(I)$ and $\mathcal{D}((D^m)_{\min}) = H^m_0(I)$.*

We have already found the self-adjoint realisations of $H = D^2$ on I in the case that I is a bounded interval. The case $I = \mathbb{R}$ is easy; it follows from part (2) of Theorem 1.2.4 and Theorem 2.3.11 that in this case, the operator H is essentially self-adjoint, and that its unique self-adjoint extension has domain $H^2(\mathbb{R})$. We shall briefly discuss the case $I =]a, \infty[$. Of course, the case $I =] - \infty, b[$ is quite similar.

2.3.12 Example. If $I =]a, \infty[$, then it follows from the preceding theorem that $\mathcal{D}(H_{\max}) = H^2(I)$, so that by part (1) of Theorem 1.2.4, the values $\phi(a)$ and $\phi'(a)$ are defined for each $\phi \in \mathcal{D}(H_{\max})$. Thus we can define the linear map $\varrho \colon \mathcal{G}(H_{\max}) \to \mathbb{C}^2$ that sends $(\phi, H_{\max}\phi)$ to $(\phi(a), \phi'(a))$, and from part (2) of Theorem 1.2.4 and Theorem 2.3.11, we infer that $\mathcal{N}(\varrho) = \mathcal{G}(H_{\min})$, so ϱ is continuous.

Let $(c_1, c_2) \in \mathbb{C}^2$ and let $\alpha \in C_0^\infty(\mathbb{R})$ be a cut-off function that is equal to 1 in a neighbourhood of a. Moreover, let P be a polynomial such that $(P(a), P'(a)) = (c_1, c_2)$, for example $P = c_2(x-a) + c_1$. Then $\alpha|_I P$ is an element of $H^2(I)$ such that the corresponding element in the graph of $H_{\max}$ is mapped to (c_1, c_2) under ϱ. We conclude that ϱ is surjective.

It follows from part (3) of Theorem 1.2.4 that for each $\phi, \psi \in H^2(I)$, we have $\lim_{x\to\infty} W_x(\psi, \phi) = 0$. Thus, applying part (2) of Theorem 2.3.5 to the intervals $]a, a+k[$ with $k \in \mathbb{N}$, and subsequently letting $k \to \infty$, we obtain

$$\langle H_{\max}\psi, \phi\rangle - \langle \psi, H_{\max}\phi\rangle = -W_a(\psi, \phi) = -\overline{\psi(a)}\phi'(a) + \overline{\psi'(a)}\phi(a),$$

by dominated convergence. Hence, if $\varrho_H \colon V_H \to \mathbb{C}^2$ is the restriction of ϱ to V_H, then the complex symplectic form ω on $\mathbb{C}^2$ associated to the matrix

$$B := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

via $\omega(\mathbf{c}, \mathbf{d}) = \mathbf{c}^* B \mathbf{d}$, satisfies $\omega_T(u,v) = \omega(\varrho_H(u), \varrho_H(v))$. Using the same approach as in Example 2.3.8, we obtain a boundary triple $(\mathbb{C}, \sigma_1, \sigma_2)$ for H, where

$$\sigma_1(\phi, H_{\max}\phi) := \phi(a) + i\phi'(a), \quad \text{and} \quad \sigma_2(\phi, H_{\max}\phi) := \phi(a) - i\phi'(a).$$

It follows that each self-adjoint extension $\widetilde{H}$ of H corresponds to a unique $\theta \in [0, 2\pi[$ via

$$\mathcal{D}(\widetilde{H}) = \{\phi \in \mathcal{D}(H_{\max})\colon \phi(a) - i\phi'(a) = e^{i\theta}(\phi(a) + i\phi'(a))\}.$$

In particular, $\theta = \pi$ and $\theta = 0$ correspond to Dirichlet and Neumann boundary conditions, respectively.

2.3.3 Some Hamiltonians with a singular endpoint

For Hamiltonians on an interval with an endpoint a that is singular for H, the limits $\lim_{x\to a^+} \phi(x)$ and $\lim_{x\to a^+} \phi'(x)$ need not both exist for each $\phi \in \mathcal{D}(H_{\max})$, in which case it is impossible to define the maps ϱ and ϱ_H in the way we did in Example 2.3.8. However, there are examples of potentials $V(x)$ where these limits can be replaced by different limits that do exist for all elements of $\mathcal{D}(H_{\max})$, so that, by modifying the map ϱ, we can still classify the self-adjoint extensions of the Hamiltonian associated to these potentials. We mention the following two examples, which can be found as [4, Example 7.4.1], where the existence of said limits is discussed in greater detail.

2.3.13 Example. The Hamiltonian with potential $V(x) = -\frac{1}{4x^2}$ on the interval $]0, 1[$: Here, 1 is a regular endpoint for H, so the limits $\lim_{x\to 1^-} \phi(x)$ and $\lim_{x\to 1^-} \phi'(x)$ are still well defined. The point 0, on the other hand, is a singular endpoint, and the limits $\lim_{x\to 0^+} \phi(x)$ and $\lim_{x\to 0^+} \phi'(x)$ are replaced by

$$g(\phi) := \lim_{x\to 0^+} \frac{\phi(x)}{\sqrt{x}\ln(x)} \quad \text{and} \quad f(\phi) := \lim_{x\to 0^+} \frac{\phi(x) - g(\phi)\sqrt{x}\ln(x)}{\sqrt{x}}.$$

Subsequently, the linear map $\varrho\colon \mathcal{G}(H_{\max}) \to \mathbb{C}^4$ is defined by

$$(\phi, H_{\max}\phi) \mapsto (f(\phi), g(\phi), \phi(1), \phi'(1)).$$

One can check that $\sqrt{x}$, $\sqrt{x}\ln(x)$, x^2 and x^3 are elements of $\mathcal{D}(H_{\max})$, and that the images of the corresponding elements of the graph under ϱ form a basis of $\mathbb{C}^4$, which shows that ϱ is surjective. Moreover, it turns out that the Wronskians at the endpoints satisfy

$$W_0(\psi, \phi) = \overline{f(\psi)}g(\phi) - \overline{g(\psi)}f(\phi), \quad \text{and} \quad W_1(\psi, \phi) = \overline{\psi(1)}\phi'(1) - \overline{\psi'(1)}\phi(1),$$

for each $\phi, \psi \in \mathcal{D}(H_{\max})$. From part (2) of Theorem 2.3.5, it follows that

$$\langle H_{\max}\psi, \phi\rangle - \langle \psi, H_{\max}\phi\rangle = W_1(\psi, \phi) - W_0(\psi, \phi) = \omega(\varrho(\psi, H_{\max}\psi), \varrho(\phi, H_{\max}\phi)),$$

for each $\phi, \psi \in \mathcal{D}(H_{\max})$, where ω is the same symplectic form as in Example 2.3.8. Now suppose $\phi \in \mathcal{D}(H_{\min})$. In view of $H_{\max}^* = H_{\min} \subseteq H_{\max}$ and the equation above, we have

$$\omega(\varrho(\psi, H_{\max}\psi), \varrho(\phi, H_{\max}\phi)) = 0,$$

for all $\psi \in \mathcal{D}(H_{\max})$. ϱ is surjective and ω is nondegenerate, so $\varrho(\phi, H_{\max}\phi) = 0$, and we conclude that $\mathcal{N}(\varrho) = \mathcal{G}(H_{\min})$.

Thus the map ϱ is continuous, and its restriction $\varrho_H\colon V_H \to \mathbb{C}^4$ to V_H satisfies

$$\omega_H(u, v) = \omega(\varrho_H(u), \varrho_H(v))$$

for each $u, v \in \mathcal{G}(H_{\max})$. We can now proceed by the exact same way as in Example 2.3.8 to show that the self-adjoint realisations $\widetilde{H}$ of H correspond to the unitary 2×2-matrices M via

$$\mathcal{D}(\widetilde{H}) = \left\{ \phi \in \mathcal{D}(H_{\max}) \;:\; \begin{pmatrix} f(\phi) + ig(\phi) \\ \phi(1) - i\phi'(1) \end{pmatrix} = M \begin{pmatrix} f(\phi) - ig(\phi) \\ \phi(1) + i\phi'(1) \end{pmatrix} \right\}.$$

For the next example, we require the following lemma:

2.3.14 Lemma. *Let $I_1, I_2, \ldots, I_m \subseteq \mathbb{R}$ be open intervals that are pairwise disjoint, let $I := \bigcup_{j=1}^m I_j$, let H be a Hamiltonian on I (with domain $\mathcal{D}(H) = C_0^\infty(I)$) and let H_j be the corresponding Hamiltonian on I_j for $j = 1, 2 \ldots, m$, so that $\mathcal{D}(H_j) = C_0^\infty(I_j)$ and $H_j(\phi|_{I_j}) = (H\phi)|_{I_j}$ for each $\phi \in C_0^\infty(I)$. Then we have*

$$\mathcal{D}(H_{\min}) = \bigoplus_{j=1}^m \mathcal{D}(H_{j,\min}), \quad \mathcal{D}(H_{\max}) = \bigoplus_{j=1}^m \mathcal{D}(H_{j,\max}), \quad V_H = \bigoplus_{j=1}^m V_{H_j},$$

and

$$\omega_H(u, v) = \sum_{j=1}^m \omega_{H_j}(P_j u, P_j v) \text{ for all } u, v \in V_H,$$

where $P_j \colon V_H \to V_{H_j}$ is the map that restricts an element of $(L^2(I))^2$ to the corresponding element of $(L^2(I_j))^2$.

Proof. We identify elements of $L^2(I_j)$ with elements of $L^2(I)$ by extension by zero outside I_j. In that way, we have

$$C_0^\infty(I) = \bigoplus_{j=1}^m C_0^\infty(I_j), \quad \text{and} \quad \mathcal{G}(H) = \bigoplus_{j=1}^m \mathcal{G}(H_j),$$

so

$$\mathcal{G}(H_{\min}) = \overline{\mathcal{G}(H)} = \bigoplus_{j=1}^m \overline{\mathcal{G}(H_j)} = \bigoplus_{j=1}^m \mathcal{G}(H_{j,\min}),$$

which implies $\mathcal{D}(H_{\min}) = \bigoplus_{j=1}^m \mathcal{D}(H_{j,\min})$. Using part (3) of Proposition 1.4.6 and the fact that $H_{\max}^* = H_{\min}$, we obtain

$$\begin{aligned}
\mathcal{G}(H_{\max}) = J(\mathcal{G}(H_{\min}))^\perp &= \left(\bigoplus_{j=1}^m J(\mathcal{G}(H_{j,\min})) \right)^\perp \\
&= \bigcap_{j=1}^m J(\mathcal{G}(H_{j,\min}))^\perp = \bigcap_{j=1}^m \left(\mathcal{G}(H_{j,\max}) \oplus \bigoplus_{k \neq j} (L^2(I_k))^2 \right) \\
&= \bigoplus_{j=1}^m \mathcal{G}(H_{j,\max}),
\end{aligned}$$

which implies $\mathcal{D}(H_{\max}) = \bigoplus_{j=1}^{m} \mathcal{D}(H_{j,\max})$. Next, we note that

$$\begin{aligned}
V_H = \mathcal{G}(H_{\max}) \cap J(\mathcal{G}(H_{\max})) &= \left(\bigoplus_{j=1}^{m} \mathcal{G}(H_{j,\max})\right) \cap \left(\bigoplus_{k=1}^{m} J(\mathcal{G}(H_{k,\max}))\right) \\
&= \bigoplus_{j=1}^{m} \mathcal{G}(H_{j,\max}) \cap J(\mathcal{G}(H_{j,\max})) = \bigoplus_{j=1}^{m} V_{H_j}.
\end{aligned}$$

Finally, let $\phi, \psi \in \mathcal{D}(H_{\max})$. By part (3) of Proposition 2.2.2, we have

$$\begin{aligned}
\omega_H((\psi, H_{\max}\psi), (\phi, H_{\max}\phi)) &= \langle H_{\max}\psi, \phi\rangle_{L^2(I)} - \langle \psi, H_{\max}\phi\rangle_{L^2(I)} \\
&= \sum_{j=1}^{m} \langle H_{j,\max}\psi|_{I_j}, \phi|_{I_j}\rangle_{L^2(I_j)} - \langle \psi|_{I_j}, H_{j,\max}\phi|_{I_j}\rangle_{L^2(I_j)} \\
&= \sum_{j=1}^{m} \omega_{H_j}(P_j(\psi, H_{\max}\psi), P_j(\phi, H_{\max}\phi)),
\end{aligned}$$

as desired. ■

2.3.15 Example. *(The one-dimensional hydrogen atom)*
The Hamiltonian with potential $V(x) = -\frac{\kappa}{|x|}$ on $\mathbb{R}\backslash\{0\}$ ($\kappa \neq 0$):
We have $\mathbb{R}\backslash\{0\} =]-\infty, 0[\cup]0, \infty[$, so if H_- and H_+ are the Hamiltonians with the same potential on $]-\infty, 0[$ and $]0, \infty[$ respectively, then by the above lemma, we can determine $H_{\max}$ by determining $H_{-,\max}$ and $H_{+,\max}$. We will determine $H_{+,\max}$; the realisation $H_{-,\max}$ is obtained analogously.

Part (1) of Theorem 2.3.5 immediately gives us the maximal realisation of $H_{+,\max}$:

$$\mathcal{D}(H_{+,\max}) = \{\phi \in L^2(]0, \infty[) \colon \phi, \phi' \in \mathrm{AC}(]0, \infty[),\ D^2\phi + V\psi \in L^2(]0, \infty[)\}.$$

In view of the fact that V is continuous and bounded on $(1, \infty)$, we have $D^2\phi|_{]1,\infty[} \in L^2(]1, \infty[)$ for each $\phi \in \mathcal{D}(H_{+,\max})$, so that by Theorem 2.3.11, we have $\phi|_{]1,\infty[} \in H^2(]1, \infty[)$. By part (4) of the same theorem, $\phi(x), \phi'(x) \to 0$ as $x \to \infty$, so we can argue as in Example 2.3.12 that

$$\langle H_{+,\max}\psi, \phi\rangle_{L^2(]0,\infty[)} - \langle \psi, H_{+,\max}\phi\rangle_{L^2(]0,\infty[)} = -W_{0^+}(\psi, \phi),$$

for each $\phi, \psi \in \mathcal{D}(H_{+,\max})$, where

$$W_{0^+}(\psi, \phi) = \lim_{x \to 0^+} \overline{\psi(x)}\phi'(x) - \overline{\psi'(x)}\phi(x).$$

Defining $W_{0^-}(\psi, \phi)$ in a similar way for $\phi, \psi \in \mathcal{D}(H_{-,\max})$, we obtain

$$\langle H_{\max}\psi, \phi\rangle_{L^2(\mathbb{R}\backslash\{0\})} - \langle \psi, H_{\max}\phi\rangle_{L^2(\mathbb{R}\backslash\{0\})} = W_{0^-}(\psi, \phi) - W_{0^+}(\psi, \phi),$$

for each $\phi, \psi \in \mathcal{D}(H_{\max})$. Because $H_{\min} = H_{\max}^* \subseteq H_{\max}$, the above equation implies

$$\mathcal{D}(H_{\min}) = \{\phi \in \mathcal{D}(H_{\max}) \colon W_{0^-}(\psi, \phi) - W_{0^+}(\psi, \phi) = 0 \text{ for each } \psi \in \mathcal{D}(H_{\max})\}.$$

It is shown in [4, Section 7.4.1] that for each $\phi \in \mathcal{D}(H_{\max})$, the limits

$$f_\pm(\phi) := \lim_{x \to 0^\pm} \phi(x), \quad \text{and} \quad g_\pm(\phi) := \lim_{x \to 0^\pm} \phi'(x) \pm \kappa\phi(x) \ln|x|,$$

exist, that

$$W_{0^\pm}(\psi, \phi) = \overline{f_\pm(\psi)} g_\pm(\phi) - \overline{g_\pm(\psi)} f_\pm(\phi),$$

and that the deficiency indices of H are equal to 2. Let $\varrho \colon \mathcal{G}(H_{\max}) \to \mathbb{C}^4$ be the map given by

$$(\phi, H_{\max}\phi) \mapsto (f_-(\phi), g_-(\phi), f_+(\phi), g_+(\phi)),$$

let ϱ_H be its restriction to V_H, and let ω be the same symplectic form as in Examples 2.3.8 and 2.3.12. Then by part (3) of Proposition 2.2.2, we have

$$\begin{aligned}
\omega_H((\psi, H_{\max}\psi), (\phi, H_{\max}\phi)) &= \langle H_{\max}\psi, \phi\rangle_{L^2(\mathbb{R}\setminus\{0\})} - \langle \psi, H_{\max}\phi\rangle_{L^2(\mathbb{R}\setminus\{0\})} \\
&= W_{0^-}(\psi, \phi) - W_{0^+}(\psi, \phi) \\
&= \overline{f_-(\psi)} g_-(\phi) - \overline{g_-(\psi)} f_-(\phi) - \overline{f_+(\psi)} g_+(\phi) + \overline{g_+(\psi)} f_+(\phi) \\
&= \omega(\varrho_H(\psi, H_{\max}\psi), \varrho_H(\phi, H_{\max}\phi)),
\end{aligned}$$

for each $(\phi, H_{\max}\phi), (\psi, H_{\max}\psi) \in V_H$. Suppose that $\varrho_H(u) = 0$ for some $u \in V_H$. Then from the above equation, it follows that $\omega_H(u, v) = 0$ for each $v \in V_H$, so $u = 0$ by part (1) of Proposition 2.2.2. Thus ϱ_H is injective. Since the deficiency indices of H are both equal to 2, we have $\dim V_H = 4$, so ϱ_H is an isomorphism and consequently, ϱ is surjective.

Earlier, we found $\mathcal{D}(H_{\min})$, and from it we infer that $\mathcal{N}(\varrho) \subseteq \mathcal{G}(H_{\min})$. On the other hand, we have

$$\dim(\mathcal{G}(H_{\max})/\mathcal{G}(H_{\min})) = \dim V_H = 4 = \dim \mathcal{R}(\varrho) = \dim(\mathcal{G}(H_{\max})/\mathcal{N}(\varrho)),$$

so $\mathcal{N}(\varrho) = \mathcal{G}(H_{\min})$ and ϱ is continuous. The rest of this example is completely analogous to Examples 2.3.8 and 2.3.12. One finds that the self-adjoint realisations $\widetilde{H}$ of H correspond to the unitary 2×2-matrices M via

$$\mathcal{D}(\widetilde{H}) = \left\{ \phi \in \mathcal{D}(H_{\max}) \; : \; \begin{pmatrix} f_+(\phi) + ig_+(\phi) \\ f_-(\phi) - ig_-(\phi) \end{pmatrix} = M \begin{pmatrix} f_+(\phi) - ig_+(\phi) \\ f_-(\phi) + ig_-(\phi) \end{pmatrix} \right\}.$$

2.4 Higher dimensions

So far, we have only looked at self-adjoint extensions of the Hamiltonian on open intervals. What about extensions of the Hamiltonian on open, connected subsets Ω of $\mathbb{R}^n$ with $n \geq 2$? Some of the results we gave have an analogue in higher dimensions. For example, if I is an interval, then using part (1) of Theorem 1.2.4, one can define the restriction of an element $\phi \in H^m(I)$ to the boundary ∂I in a meaningful way. In higher dimensions, we have the following statement, which can be found in [2, pp. 315-316]:

2.4.1 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be an open set with a bounded, C^1 boundary. Then there exists a unique bounded, surjective linear map $H^2(\Omega) \to H^{3/2}(\partial\Omega) \times H^{1/2}(\partial\Omega)$ that sends* *each $\phi \in C^\infty(\Omega)$ to $(\phi|_{\partial\Omega}, \frac{\partial\phi}{\partial n})$, where $\frac{\partial\phi}{\partial n}$ is the normal derivative of ϕ at the boundary. Furthermore, we have*

$$\langle -\Delta\psi, \phi\rangle_{L^2(\Omega)} - \langle \psi, -\Delta\phi\rangle_{L^2(\Omega)} = \langle \frac{\partial\psi}{\partial n}, \phi\rangle_{L^2(\partial\Omega)} - \langle \psi, \frac{\partial\phi}{\partial n}\rangle_{L^2(\partial\Omega)}.$$

In this expression, $\partial\Omega$ is equipped with its surface measure.

The spaces $H^{3/2}(\partial\Omega)$ and $H^{1/2}(\partial\Omega)$ are *fractional Sobolev spaces*; we shall not bother to define them here, and instead refer the reader to [2, p. 314]. Theorems such as the one above, relating elements of Sobolev spaces on a certain domain to elements of Sobolev spaces on the boundary of that domain, are called *trace theorems*, and the integration by parts formula in the above theorem is a rigorous version of what is commonly known as *Green's second identity*. We have already encountered the one-dimensional version of this identity in part (2) of Theorem 2.3.5.

Theorem 2.3.11, on the other hand, carries over only partially to higher dimensions for $m = 2$; we have $\mathcal{D}(-\Delta_{\min}) = H_0^2(\Omega)$ (cf. [10, Theorem 6.24]), but in general, $H_0^2(\Omega)$ is a proper subspace of $\mathcal{D}(-\Delta_{\max})$ (cf. [10, p. 143]). This of course also means that the above trace theorem does not apply to all elements of $\mathcal{D}(-\Delta_{\max})$. Finally, because $H^{3/2}(\partial\Omega)$ and $H^{1/2}(\partial\Omega)$ are infinite dimensional spaces, the endpoint space of $-\Delta$ has infinite dimension as well.

Despite the fact that a general classification of self-adjoint extensions of the Laplacian in higher dimensions seems to be out of reach, we do have the following positive result:

2.4.2 Theorem. *Let $\Omega \subseteq \mathbb{R}^n$ be an open, bounded subset with C^2-boundary, let $-\Delta_{\max}$ be the maximal realisation of the Laplacian on Ω. Then*

$$\begin{aligned} -\Delta_D &:= -\Delta_{\max}|_{H^2(\Omega)\cap H_0^1(\Omega)}, \quad \textit{and} \\ -\Delta_N &:= -\Delta_{\max}|_{H^2(\Omega)\cap V}, \end{aligned}$$

are self-adjoint realisations of $-\Delta$, where

$$V := \{\phi \in H^2(\Omega) \colon \frac{\partial\phi}{\partial n} = 0\},$$

and the normal derivative $\frac{\partial\phi}{\partial n}$ is defined as in the trace theorem.

Proof. See Theorems 10.19 and 10.20 in [19]. ■

The realisations $-\Delta_D$ and $-\Delta_N$ are associated to Dirichlet and Neumann boundary conditions respectively. Thus these boundary conditions have self-adjoint extensions for a reasonably large class of subsets of $\mathbb{R}^n$.

3 Coherent states and the classical limit

We will now study so-called *coherent states*, which are useful for a couple of reasons. Firstly, they provide us with a way of quantizing classical observables, i.e. associating operators on some Hilbert space to these classical observables. Secondly, coherent states can be regarded as quantum-mechanical approximations to the classical moving point particle. They depend on the parameter $\hbar \in]0, \infty[$, and the approximation to the point particle becomes better and better as $\hbar \to 0$, enabling us to study classical motion as a limit of the time evolution of the coherent state as determined by the Schrödinger equation. Let us give the formal definition of a coherent state as it is found in [12].

3.0.1 Definition. Let M be the phase space of a system with Liouville measure μ_L let $X \subseteq]0, \infty[$ be a set such that $0 \in \overline{X}$ (the closure is taken with respect to the topology on $\mathbb{R}$), and let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. A family $(\Psi_z^\hbar)_{(\hbar,z)\in X\times M}$ of elements of $\mathcal{H}$ is called a *family of coherent states* if it satisfies the following conditions:

(1) For each $\hbar \in X$, the map $M \to \mathcal{H}$, $z \mapsto \Psi_z^\hbar$ is an injection.

(2) For each $z \in M$ and each $\hbar \in X$, we have $\|\Psi_z^\hbar\| = 1$.

(3) For each $\hbar \in X$, there exists a constant $c_\hbar > 0$ such that for each normalized $\Phi \in \mathcal{H}$, we have
$$c_\hbar \int_M d\mu_L(z)\, |\langle \Psi_z^\hbar, \Phi \rangle|^2 = 1.$$

(4) For each $z \in M$ and each $f \in C_c^0(M)$ ($C_c^0(M)$ being the space of compactly supported continuous functions on M), we have
$$\lim_{\hbar\to 0} c_\hbar \int_M d\mu_L(w)\, f(w) |\langle \Psi_w^\hbar, \Psi_z^\hbar \rangle|^2 = f(z).$$

The third property in the definition can be interpreted as follows: For a given value of $\hbar$ and a (normalized) state Φ, the map $N \mapsto c_\hbar \int_N d\mu_L(z)\, |\langle \Psi_z^\hbar, \Phi \rangle|^2$ from the μ_L-measurable sets to $[0, 1]$ defines a probability measure on the phase space M. Its probability density function $z \mapsto c_\hbar |\langle \Psi_z^\hbar, \Phi \rangle|^2$ is called the *Husimi function*.

The fourth property then implies that for $\Phi = \Psi_z^\hbar$, this measure, viewed as a functional on $C_c^0(M)$, converges weakly to the Dirac measure at z, which justifies the statement that $\Psi_z^\hbar$ is an approximation to a point particle which is located at point z in the phase space M.

Next, we show how to define an associated quantization map:

3.0.2 Definition. Let f be a measurable, essentially bounded function on M and let $\hbar \in X$. Then we define $\mathcal{Q}_\hbar^B(f)\colon \mathcal{H} \to \mathcal{H}$ by
$$\mathcal{Q}_\hbar^B(f)(\Phi) := c_\hbar \int_M d\mu_L(z)\, f(z) \langle \Psi_z^\hbar, \Phi \rangle \Psi_z^\hbar.$$
for each $\Phi \in \mathcal{H}$. This type of quantization is called *Berezin quantization*.

We will mainly be concerned with the following family of coherent states:

3.0.3 Example. Let $M := \mathbb{R}^{2n} \cong T^*\mathbb{R}^n$ be the phase space of the configuration space $\mathbb{R}^n$, with measure $d\mu_L = (2\pi)^{-n}\prod_{i=1}^n dq_i dp_i$, let $X :=]0,\infty[$ and let $\mathcal{H} := L^2(\mathbb{R}^n)$. Then the states given by

$$\Psi^\hbar_{(p,q)} := (\pi\hbar)^{-n/4} e^{-ip\cdot q/2\hbar} e^{ip\cdot x/\hbar} e^{-(x-q)^2/2\hbar},$$

are known as *Schrödinger's coherent states*.

We shall give a very rough sketch of the proof that Schrödinger's coherent states do indeed form a family of coherent states. It is straightforward to show that the states are normalized. To prove that they possess the third property in Definition 3.0.1, let $c_\hbar := \hbar^{-n}$ and use the Parseval-Plancherel theorem, performing the integral over p before the one over q. To show that they have the first and the fourth property, one can show as an intermediate step that

$$|\langle \Psi^\hbar_{(p,q)}, \Psi^\hbar_{(q',p')}\rangle|^2 = e^{-(|q-q'|^2+|p-p'|^2)/2\hbar}.$$

3.1 Modifying Schrödinger's states

The above family of states is defined on the phase space associated to the configuration space $\mathbb{R}^n$. We would like to construct a family of coherent states $(\phi^\hbar_z)$ associated to the phase space $T^*\Omega \cong \mathbb{R}^n \times \Omega$ of an arbitrary domain $\Omega \subseteq \mathbb{R}^n$.

Since we are interested in the dynamics of the system, we want to be able to study the time evolution of these states, so we must demand that $\phi^\hbar_z \in \mathcal{D}(\widetilde{H})$, where $\widetilde{H}$ is a self-adjoint extension of the Hamiltonian $H = -\Delta$ with domain $\mathcal{D}(H) = C^\infty_0(\Omega)$. In the case of $\Omega = \mathbb{R}$, the Hamiltonian is essentially self-adjoint, and its unique self-adjoint extension has domain $H^2(\mathbb{R})$, which clearly contains the family $(\Psi^\hbar_z)_{z\in M,\hbar\in X}$ of smooth, rapidly decreasing functions.

However, if Ω is not all of $\mathbb{R}$, for example if $\Omega =]0,1[$, then the domains of the self-adjoint extensions of H consist of functions satisfying certain boundary conditions, and Schrödinger's states (after normalization with respect to $\|\cdot\|_{L^2(\Omega)}$) will in general *not* meet these boundary conditions. Nevertheless, these states can still be of use to us, although we will have to modify them slightly, and weaken the definition of a coherent state as stated in 3.0.1. In particular, property (3) listed in this definition will not be satisfied for a fixed $\hbar$, but will be recovered in the limit $\hbar \to 0$.

Let us proceed by constructing the modified states for the case $\Omega = I :=]a,b[$, with $a, b \in \mathbb{R}$, $a < b$, and we identify its phase space M with $\mathbb{R} \times I$. First, choose a bump function $\chi \in C^\infty_0(\mathbb{R})$ with the following properties:

$$\chi(x) = \begin{cases} 1 & |x| \le 1, \\ \in]0,1[& |x| \in]1,2[, \\ 0 & |x| \ge 2. \end{cases}$$

Next, let $\alpha\colon X \times M \to]0,\infty[$ be a continuous function such that:

(1) For each $z \in M$, we have $\alpha(\hbar, z) \uparrow \infty$ as $\hbar \downarrow 0$;

(2) For each $z \in M$, we have

$$h^{-1} e^{-\hbar^{-1}\alpha(\hbar,\cdot)^{-2}/2} \downarrow 0 \text{ as } \hbar \downarrow 0;$$

(3) For each $\hbar \in X$ and each polynomial $P \in \mathbb{C}[q,p]$, we have

$$P(\cdot)e^{-\hbar^{-1}\alpha(\hbar,\cdot)^{-2}/2} \in L^1(M, \mu_L).$$

In view of the second and the third property of α and Lebesgue's dominated convergence theorem, we have

$$\lim_{\hbar\to 0} \hbar^{-1} \int_M d\mu_L(z)\, P(z)e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2} = 0.$$

An example of such a function α is given by $\alpha(\hbar, z) = (|p|+1)^{-1}\hbar^{-1/4}$, where the set X is of the form $]0,\delta[$ for some small number $\delta > 0$. For each $\hbar \in]0,1[$ and for each $(p,q) \in \mathbb{R} \times I$, define the function $\widetilde{\phi^{\hbar}_{(p,q)}} \colon \mathbb{R} \mapsto \mathbb{C}$,

$$x \mapsto \chi(\alpha(\hbar,(p,q)) \cdot (x-q)) \cdot \Psi^{\hbar}_{(p,q)}(x),$$

and note that this a compactly supported, smooth function, hence the function is square integrable. Moreover, $\Psi^{\hbar}_{(p,q)}$ is a nowhere vanishing smooth function, so on every compact subset of $\mathbb{R}$, its absolute value is bounded from below by a positive constant. Now note that the function $\widetilde{\phi^{\hbar}_{(p,q)}}$ is equal to $\Psi^{\hbar}_{(p,q)}$ on a compact neighbourhood of q, so $\|\widetilde{\phi^{\hbar}_{(p,q)}}\|_{L^2(\mathbb{R})} > 0$ and consequently, we may normalize the function:

$$\phi^{\hbar}_{(p,q)} := \|\widetilde{\phi^{\hbar}_{(p,q)}}\|^{-1}_{L^2(\mathbb{R})} \widetilde{\phi^{\hbar}_{(p,q)}}.$$

Thus we have constructed a family of compactly supported, smooth, normalized functions. However, we only want to keep the functions whose support is a subset of I, so we define the set

$$S := \{(\hbar, z) \in]0,1[\times M \colon \operatorname{supp}(\phi^{\hbar}_z) \subset I\},$$

and for each $\hbar \in]0,1[$, we define

$$M_\hbar := \{z \in M \colon (\hbar, z) \in S\}.$$

The next lemma establishes some properties of these functions:

3.1.1 Lemma. *The family of functions $(\phi^{\hbar}_z)_{(\hbar,z)\in S}$ defined above has the following properties:*

(1) For each $(\hbar,z) \in S$ and each self-adjoint extension $\widetilde{H}$ of H, we have $\|\phi^{\hbar}_z\|_{L^2(I)} = 1$ and $\phi^{\hbar}_z \in \mathcal{D}(\widetilde{H})$.

(2) $M_\hbar \uparrow M$ as $\hbar \downarrow 0$. Moreover, $M_\hbar$ is open in M for each $\hbar \in]0,1[$.

(3) For each $(\hbar,z) \in X \times M$, we have $\|\phi^{\hbar}_z - \Psi^{\hbar}_z\|_{L^2(\mathbb{R})} < 2e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2}$.

Proof.
(1) For each $(\hbar,z) \in S$, the function $\phi^{\hbar}_z$ is compactly supported and by definition of S, its support is a subset of I, so $\|\phi^{\hbar}_z\|_{L^2(I)} = \|\phi^{\hbar}_z\|_{L^2(\mathbb{R})}$ and $\phi^{\hbar}_z \in C^\infty_0(I) = \mathcal{D}(H)$, which implies that $\phi^{\hbar}_z$ lies in the domain of every self-adjoint extension of H.

(2) Fix $z = (p, q) \in M$. Then $\alpha(\hbar, z) \uparrow \infty$ as $\hbar \downarrow 0$, so there exists an $\hbar > 0$ such that $\alpha(\hbar, z) \cdot (a - q) < -2$ and $\alpha(\hbar, z) \cdot (b - q) > 2$, which implies that $\mathrm{supp}(\phi_z^\hbar) \subset I$, so $z \in M_\hbar$. In addition, if $0 < \hbar' < \hbar$, then $\alpha(\hbar', z) > \alpha(\hbar, z)$ and consequently, $z \in M_{\hbar'}$. We conclude that $M_\hbar \uparrow M$ as $\hbar \downarrow 0$.

To prove the second assertion, fix $(\hbar, (p, q)) \in S$ and let d be the distance of $\mathrm{supp}(\phi^\hbar_{(p,q)})$ to $\{a, b\}$. These two sets are compact and disjoint, so $d > 0$. The function α is continuous, so the function

$$M \to \mathbb{R}, \quad (p', q') \mapsto \alpha(\hbar, (p', q'))^{-1},$$

is continuous at (p, q). Hence there exists a $\delta > 0$ such that

$$|\alpha(\hbar, (p, q))^{-1} - \alpha(\hbar, (p', q'))^{-1}| < d/4,$$

whenever $\|(p, q) - (p', q')\| < \delta$. Since $\mathrm{supp}(\chi) = [-2, 2]$, we have

$$2\alpha(\hbar, (p, q))^{-1} + d = \min(q - a, b - q).$$

But then for each $(p', q') \in M$ such that $|(p, q) - (p', q')| < \min(\delta, d/2)$, we have

$$\alpha(\hbar, (p', q'))^{-1} \leq \alpha(\hbar, (p, q))^{-1} + |\alpha(\hbar, (p, q))^{-1} - \alpha(\hbar, (p', q'))^{-1}| < \alpha(\hbar, (p, q))^{-1} + d/4,$$

and $|q - q'| < d/2$, so

$$2\alpha(\hbar, (p', q'))^{-1} < 2\alpha(\hbar, (p, q))^{-1} + d/2 = \min(q - a, b - q) - d/2,$$

and hence

$$2\alpha(\hbar, (p', q'))^{-1} < \min(q' - a, b - q'),$$

which implies that $\mathrm{supp}(\phi^\hbar_{(p',q')}) \subset I$ and hence $(p', q') \in M_\hbar$. Thus $M_\hbar$ is open.

(3) Let $(\hbar, z) \in S$, and write $(p, q) = z$. First, note that

$$\begin{aligned}
\|\psi_z^\hbar - \widetilde{\phi}_z^\hbar\|^2_{L^2(\mathbb{R})} &= \int_{\mathbb{R}} |\psi_z^\hbar(x)|^2 \cdot (1 - \chi(\alpha(\hbar, z) \cdot (x - q)))^2 \, dx \\
&= (\pi\hbar)^{-1/2} \int_{\mathbb{R}} e^{-(x-q)^2/\hbar} (1 - \chi(\alpha(\hbar, z) \cdot (x - q)))^2 \, dx \\
&\leq (\pi\hbar)^{-1/2} \left(\int_{-\infty}^{-\alpha(\hbar,z)^{-1}} + \int_{\alpha(\hbar,z)^{-1}}^{\infty} e^{-x^2/\hbar} \, dx \right) \\
&= \pi^{-1/2} \left(\int_{-\infty}^{-\hbar^{-1/2}\alpha(\hbar,z)^{-1}} + \int_{\hbar^{-1/2}\alpha(\hbar,z)^{-1}}^{\infty} e^{-x^2} \, dx \right) \\
&= \pi^{-1/2} \left(\int_{-\infty}^{0} e^{-(x - \hbar^{-1/2}\alpha(\hbar,z)^{-1})^2} \, dx + \int_0^\infty e^{-(x + \hbar^{-1/2}\alpha(\hbar,z)^{-1})^2} \, dx \right) \\
&= \pi^{-1/2} e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}} \int_{\mathbb{R}} e^{-x^2} e^{-2|x|\hbar^{-1/2}\alpha(\hbar,z)^{-1}} \, dx \\
&\leq \pi^{-1/2} e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}} \int_{\mathbb{R}} e^{-x^2} \, dx \\
&= e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}}.
\end{aligned}$$

Next, we remark that

$$\|\widetilde{\phi_z^\hbar}\|_{L^2(\mathbb{R})} \leq \|\psi_z^\hbar\|_{L^2(\mathbb{R})} \cdot \|\chi\|_{L^\infty(\mathbb{R})} = \|\psi_z^\hbar\|_{L^2(\mathbb{R})} = 1,$$

so

$$\begin{aligned} 0 \leq 1 - \|\widetilde{\phi_z^\hbar}\|^2_{L^2(\mathbb{R})} &= \|\psi_z^\hbar\|^2_{L^2(\mathbb{R})} - \|\widetilde{\phi_z^\hbar}\|^2_{L^2(\mathbb{R})} \\ &= \int_{\mathbb{R}} |\psi_z^\hbar(x)|^2 \cdot (1 - \chi(\alpha(\hbar, z) \cdot (x - q))^2)\, dx. \end{aligned}$$

The final integral in the expression above is very similar to the one in the first step of the estimate for $\|\psi_z^\hbar - \widetilde{\phi_z^\hbar}\|^2_{L^2(\mathbb{R})}$. The functions

$$(1 - \chi(\alpha(\hbar, z) \cdot (x - q)))^2, \text{ and } 1 - \chi(\alpha(\hbar, z) \cdot (x - q))^2,$$

are both smooth functions of x that vanish on the same compact neighbourhood of q and which are bounded with sup-norm 1, so it is not hard to show that

$$1 - \|\widetilde{\phi_z^\hbar}\|^2_{L^2(\mathbb{R})} \leq e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}}.$$

This yields

$$\begin{aligned} \|\phi_z^\hbar - \widetilde{\phi_z^\hbar}\|_{L^2(\mathbb{R})} = 1 - \|\widetilde{\phi_z^\hbar}\|_{L^2(\mathbb{R})} &= \frac{1 - \|\widetilde{\phi_z^\hbar}\|^2_{L^2(\mathbb{R})}}{1 + \|\widetilde{\phi_z^\hbar}\|_{L^2(\mathbb{R})}} < 1 - \|\widetilde{\phi_z^\hbar}\|^2_{L^2(\mathbb{R})} \\ &\leq e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}}, \end{aligned}$$

so by the triangle inequality, we have

$$\begin{aligned} \|\phi_z^\hbar - \psi_z^\hbar\|_{L^2(\mathbb{R})} &\leq \|\phi_z^\hbar - \widetilde{\phi_z^\hbar}\|_{L^2(\mathbb{R})} + \|\psi_z^\hbar - \widetilde{\phi_z^\hbar}\|_{L^2(\mathbb{R})} \\ &< e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}} + e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2} \\ &< 2e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2}, \end{aligned}$$

as desired. ■

Now we will see to what extent the family of states $(\phi_z^\hbar)_{(\hbar,z)\in S}$ satisfies the properties listed in Definition 3.0.1.

3.1.2 Proposition. *The family of states $(\phi_z^\hbar)_{(\hbar,z)\in S}$ has the following properties:*

(1) For each $\hbar \in X$, the map $M_\hbar \to L^2(I)$, $z \mapsto \phi_z^\hbar$ is an injection.

(2) For each $z \in M$ and each $\hbar \in X$, we have $\|\phi_z^\hbar\|_{L^2(I)} = 1$.

(3) For each normalized vector $\Phi \in L^2(I)$, we have

$$\lim_{\hbar \to 0} h^{-1} \int_M d\mu_L(z)\, |\langle \phi_z^\hbar, \Phi \rangle_{L^2(I)}|^2 = 1.$$

(4) For each $z \in M$ and each $f \in C_c^0(M)$, we have

$$\lim_{\hbar \to 0} h^{-1} \int_{M_\hbar} d\mu_L(w)\, f(w) |\langle \phi_w^\hbar, \phi_z^\hbar \rangle_{L^2(I)}|^2 = f(z).$$

Proof.
(1) Let $\hbar \in X$ let $(p,q),(p',q'), \in M_\hbar$, and suppose $\phi_{(p,q)}^\hbar$ is equal to $\phi_{(p',q')}^\hbar$ almost everywhere. Then $\phi_{(p,q)}^\hbar = \phi_{(p',q')}^\hbar$, because the two functions are continuous. Now $|\phi_{(p,q)}^\hbar|^2$ attains its maximum at q, and similarly, $|\phi_{(p',q')}^\hbar|^2$ attains its maximum at q', so $q = q'$. This also implies

$$\frac{d}{dx} (\overline{\phi_{(p',q')}^\hbar(x)} \phi_{(p,q)}^\hbar(x))|_{x=q} = \frac{d}{dx} \left(|\phi_{(p,q)}^\hbar(x)|^2 \right) |_{x=q} = 0,$$

while on the other hand,

$$\begin{aligned} \frac{d}{dx} (\overline{\phi_{(p',q')}^\hbar(x)} \phi_{(p,q)}^\hbar(x))|_{x=q} &= \frac{d}{dx} (\overline{\Psi_{(p',q)}^\hbar(x)} \Psi_{(p,q)}^\hbar(x))|_{x=q} \\ &= \frac{i}{h}(p-p') e^{i(p-p')q/(2\hbar)}, \end{aligned}$$

so $p = p'$. We conclude that the map $M_\hbar \ni z \mapsto \phi_z^\hbar \in L^2(I)$ is injective.

(2) This is part (1) of Lemma 3.1.1.

(3) We will divide the proof of this part into three steps:
(I) Let $\Phi \in L^2(I)$ and assume that Φ is normalised. By abuse of notation, we will use Φ to denote (i) an equivalence class of functions, (ii) its extension by 0 to $L^2(\mathbb{R})$, and (iii) a representative $\mathbb{R} \to \mathbb{C}$ of this class. Then the function $\mathbb{R}^2 \to \mathbb{R}$ given by $(q,x) \mapsto e^{-q^2}|\Phi(x)|^2$ is an element of $L^1(\mathbb{R}^2)$, and by Fubini's theorem,

$$\pi^{-1/2} \int_{\mathbb{R}^2} e^{-q^2} |\Phi(x)|^2 \, d\lambda = \pi^{-1/2} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-q^2} |\Phi(x)|^2 \, dq \, dx = 1,$$

where λ is the Lebesgue measure on $\mathbb{R}^2$. For each $\hbar \in X$ and $x \in \mathbb{R}$, let

$$\mathbf{1}_{[h^{-1/2}(x-b), h^{-1/2}(x-a)]} \colon \mathbb{R} \to \mathbb{R},$$

be the characteristic function of the interval $[h^{-1/2}(x-b), h^{-1/2}(x-a)]$. Then the function from $\mathbb{R}^2$ to $\mathbb{R}$,

$$(q,x) \mapsto \mathbf{1}_{[h^{-1/2}(x-b), h^{-1/2}(x-a)]} e^{-q^2} |\Phi(x)|^2,$$

is integrable, and by Lebesgue's dominated convergence theorem, we have

$$\lim_{\hbar \to 0} \pi^{-1/2} \int_{\mathbb{R}^2} \mathbf{1}_{[h^{-1/2}(x-b), h^{-1/2}(x-a)]} e^{-q^2} |\Phi(x)|^2 \, d\lambda = 1,$$

since the essential support of Φ is a subset of $]a,b[$. Also note that

$$
\begin{aligned}
\pi^{-1/2} & \int_{\mathbb{R}^2} \mathbf{1}_{[h^{-1/2}(x-b),h^{-1/2}(x-a)]} e^{-q^2} |\Phi(x)|^2 \, d\lambda \\
&= \pi^{-1/2} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{[h^{-1/2}(x-b),h^{-1/2}(x-a)]} e^{-q^2} |\Phi(x)|^2 \, dq \, dx \\
&= \pi^{-1/2} \int_{\mathbb{R}} \int_{h^{-1/2}(x-b)}^{h^{-1/2}(x-a)} e^{-q^2} |\Phi(x)|^2 \, dq \, dx \\
&= (\pi\hbar)^{-1/2} \int_{\mathbb{R}} \int_{(x-b)}^{(x-a)} e^{-q^2/\hbar} |\Phi(x)|^2 \, dq \, dx \\
&= (\pi\hbar)^{-1/2} \int_{\mathbb{R}} \int_{a}^{b} e^{-(x-q)^2/\hbar} |\Phi(x)|^2 \, dq \, dx \\
&= (\pi\hbar)^{-1/2} \int_{a}^{b} \int_{\mathbb{R}} e^{-(x-q)^2/\hbar} |\Phi(x)|^2 \, dx \, dq.
\end{aligned}
$$

Defining the function $f_q^\hbar \colon \mathbb{R} \to \mathbb{C}$ by

$$f_q^\hbar(x) := (\pi\hbar)^{-1/4} e^{-(x-q)^2/2\hbar},$$

for each $q \in I$ and $\hbar \in X$, we conclude that

$$\lim_{\hbar \to 0} \int_a^b \|f_q^\hbar\|_{L^2(\mathbb{R})} \, dq = 1.$$

(II) Next, consider the inner product $\langle \Psi^\hbar_{(p,q)}, \Phi \rangle_{L^2(I)}$, which is equal to $\langle \Psi^\hbar_{(p,q)}, \Phi \rangle_{L^2(\mathbb{R})}$, again because the essential support of Φ is a subset of I. Writing out the inner product explicitly reveals that (up to some constant factors) it is the Fourier transform of the function $f_q^\hbar$ with respect to x, i.e.

$$\langle \Psi^\hbar_{(p,q)}, \Phi \rangle_{L^2(\mathbb{R})} = e^{ipq/2\hbar} \mathcal{F}(f_q^\hbar)(p/\hbar),$$

where $\mathcal{F}(f_q^\hbar)(p/\hbar)$ is the Fourier transform of $f_q^\hbar$ with respect to x evaluated at $p/\hbar$. We have $f_q^\hbar \in L^2(\mathbb{R})$, so by the Parseval-Plancherel theorem, the function $\mathbb{R} \to \mathbb{R}$, $p \mapsto |\langle \Psi^\hbar_{(p,q)}, \Phi \rangle_{L^2(\mathbb{R})}|^2$ is square integrable (for fixed q), and

$$
\begin{aligned}
(2\pi\hbar)^{-1} \int_{\mathbb{R}} |\langle \Psi^\hbar_{(p,q)}, \Phi \rangle_{L^2(I)}|^2 \, dp &= (2\pi\hbar)^{-1} \int_{\mathbb{R}} |\mathcal{F}(f_q^\hbar)(p/\hbar)|^2 \, dp \\
&= (2\pi)^{-1} \int_{\mathbb{R}} |\mathcal{F}(f_q^\hbar)(p)|^2 \, dp \\
&= \|f_q^\hbar\|_{L^2(\mathbb{R})}.
\end{aligned}
$$

From (I), we know that the function $q \mapsto \|f_q^\hbar\|_{L^2(\mathbb{R})}$ is integrable, so by Tonelli's theorem, the function $\mathbb{R}^2 \to \mathbb{R}$, defined by $(p,q) \mapsto |\langle \Psi^\hbar_{(p,q)}, \Phi \rangle_{L^2(\mathbb{R})}|^2$ is an element of $L^1(M, \mu_L)$, and by Fubini's theorem, we have

$$\int_M |\langle \Psi^\hbar_z, \Phi \rangle_{L^2(I)}|^2 \, d\mu_L(z) = \int_a^b \|f_q^\hbar\|_{L^2(\mathbb{R})} \, dq,$$

from which we infer that

$$\lim_{\hbar\to 0}\int_M |\langle \Psi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\, d\mu_L(z) = 1.$$

(III) All that remains is to replace the functions $\Psi_z^\hbar$ in the previous equation with the functions $\phi_z^\hbar$. First apply the triangle inequality and the Cauchy-Schwarz inequality together with part (3) of Lemma 3.1.1 to obtain

$$\begin{aligned}
&\big||\langle \Psi_z^\hbar, \Phi\rangle_{L^2(I)}|^2 - |\langle \phi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\big| \\
&\quad = \big|\langle \Phi, \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \Psi_z^\hbar, \Phi\rangle_{L^2(\mathbb{R})} - \langle \Phi, \phi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \phi_z^\hbar, \Phi\rangle_{L^2(\mathbb{R})}\big| \\
&\quad \le \big|\langle \Phi, \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \Psi_z^\hbar - \phi_z^\hbar, \Phi\rangle_{L^2(\mathbb{R})}\big| + \big|\langle \Phi, \Psi_z^\hbar - \phi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \phi_z^\hbar, \Phi\rangle_{L^2(\mathbb{R})}\big| \\
&\quad \le 2\|\Psi_z^\hbar - \phi_z^\hbar\|_{L^2(\mathbb{R})} < 4e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2}.
\end{aligned}$$

From this, we deduce that the function $z \mapsto |\langle \phi_z^\hbar, \Phi\rangle_{L^2(I)}|^2$ is an element of $L^1(M, \mu_L)$, and that

$$\begin{aligned}
&\left|\hbar^{-1}\int_M |\langle \Psi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\, d\mu_L(z) - \hbar^{-1}\int_M |\langle \phi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\, d\mu_L(z)\right| \\
&\quad \le \int_M \hbar^{-1}\big||\langle \Psi_z^\hbar, \Phi\rangle_{L^2(I)}|^2 - |\langle \phi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\big|\, d\mu_L(z) \\
&\quad \le 4\hbar^{-1}\int_M e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2}\, d\mu_L(z).
\end{aligned}$$

The expression in the last line converges to 0 as $\hbar \to 0$, so

$$\lim_{\hbar\to 0}\left(\hbar^{-1}\int_M |\langle \Psi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\, d\mu_L(z) - \hbar^{-1}\int_M |\langle \phi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\, d\mu_L(z)\right) = 0.$$

Together with our result from (II), this implies

$$\lim_{\hbar\to 0}\hbar^{-1}\int_M |\langle \phi_z^\hbar, \Phi\rangle_{L^2(I)}|^2\, d\mu_L(z) = 1,$$

as desired.

(4) Let $f \in C_c^0(I)$ and let $z \in M$. f is continuous and has compact support, so f is bounded, and by part (2) of Lemma 3.1.1, there exists an $\hbar_0 \in X$ such that $z \in M_{\hbar_0}$ and $\mathrm{supp}(f) \subset M_{\hbar_0}$. Using part (3) of the same lemma, the triangle inequality and the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
&\big||\langle \phi_w^\hbar, \phi_z^\hbar\rangle_{L^2(I)}|^2 - |\langle \Psi_w^\hbar, \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}|^2\big| \\
&\quad = \big||\langle \phi_w^\hbar, \phi_z^\hbar\rangle_{L^2(\mathbb{R})}|^2 - |\langle \Psi_w^\hbar, \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}|^2\big| \\
&\quad = \big|\langle \phi_w^\hbar, \phi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \phi_z^\hbar, \phi_w^\hbar\rangle_{L^2(\mathbb{R})} - \langle \Psi_w^\hbar, \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \Psi_z^\hbar, \Psi_w^\hbar\rangle_{L^2(\mathbb{R})}\big| \\
&\quad \le |\langle \phi_w^\hbar, \phi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \phi_z^\hbar, \phi_w^\hbar - \Psi_w^\hbar\rangle_{L^2(\mathbb{R})}| + |\langle \phi_w^\hbar, \phi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \phi_z^\hbar - \Psi_z^\hbar, \Psi_w^\hbar\rangle_{L^2(\mathbb{R})}| \\
&\qquad + |\langle \phi_w^\hbar, \phi_z^\hbar - \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \Psi_z^\hbar, \Psi_w^\hbar\rangle_{L^2(\mathbb{R})}| + |\langle \phi_w^\hbar - \Psi_w^\hbar, \Psi_z^\hbar\rangle_{L^2(\mathbb{R})}\langle \Psi_z^\hbar, \Psi_w^\hbar\rangle_{L^2(\mathbb{R})}| \\
&\quad \le 4(e^{-\hbar^{-1}\alpha(\hbar,w)^{-2}/2} + e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2}),
\end{aligned}$$

for each $\hbar$ such that $z, w \in M_\hbar$. Thus for each $\hbar < \hbar_0$ such that $\hbar \in X$, we have

$$\begin{aligned}
&\left|\hbar^{-1} \int_{M_\hbar} d\mu_L(w)\, f(w) |\langle \phi_w^\hbar, \phi_z^\hbar \rangle_{L^2(I)}|^2 - \hbar^{-1} \int_{\mathbb{R}^2} d\mu_L(w)\, f(w) |\langle \Psi_w^\hbar, \Psi_z^\hbar \rangle_{L^2(\mathbb{R})}|^2 \right| \\
&\quad \leq \hbar^{-1} \int_{\operatorname{supp}(f)} d\mu_L(w)\, |f(w)| \cdot \left| |\langle \phi_w^\hbar, \phi_z^\hbar \rangle_{L^2(\mathbb{R})}|^2 - |\langle \Psi_w^\hbar, \Psi_z^\hbar \rangle_{L^2(\mathbb{R})}|^2 \right| \\
&\quad \leq 4\hbar^{-1} \|f\|_{L^\infty(M)} \left(\int_M e^{-\hbar^{-1}\alpha(\hbar,w)^{-2}/2}\, d\mu_L(w) + \mu_L(\operatorname{supp}(f)) \cdot e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2} \right).
\end{aligned}$$

The last line converges to 0 as $\hbar \to 0$, so

$$\lim_{\hbar \to 0} \left(\hbar^{-1} \int_{M_\hbar} d\mu_L(w)\, f(w) |\langle \phi_w^\hbar, \phi_z^\hbar \rangle_{L^2(I)}|^2 - \hbar^{-1} \int_{\mathbb{R}^2} d\mu_L(w)\, f(w) |\langle \Psi_w^\hbar, \Psi_z^\hbar \rangle_{L^2(\mathbb{R})}|^2 \right) = 0.$$

One of the properties of Schrödinger's coherent states is that

$$\lim_{\hbar \to 0} \hbar^{-1} \int_{\mathbb{R}^2} d\mu_L(w)\, f(w) |\langle \Psi_w^\hbar, \Psi_z^\hbar \rangle_{L^2(\mathbb{R})}|^2 = f(z),$$

hence

$$\lim_{\hbar \to 0} \hbar^{-1} \int_{M_\hbar} d\mu_L(w)\, f(w) |\langle \phi_w^\hbar, \phi_z^\hbar \rangle_{L^2(I)}|^2 = f(z),$$

which is what we wanted to show. ■

3.2 Expectation values of position and momentum

Proposition 3.1.2 suggests a quantization procedure $Q_\hbar$ on M much like Berezin quantization on $T^*\mathbb{R}^n$. To a function (the classical observable) $f\colon M \to \mathbb{R}$, we associate a linear operator $Q_\hbar(f)$ on some Hilbert space $\mathcal{H}$, where the operator is uniquely determined by

$$\langle \Phi, Q_\hbar(f)\Phi \rangle_{L^2(I)} = h^{-1} \int_M d\mu_L(z)\, f(z) |\langle \phi_z^\hbar, \Phi \rangle_{L^2(I)}|^2.$$

Since we are interested in the dynamics of the system, a question that naturally poses itself is what operators this new quantization map associates to the classical observables of position and momentum. We will see that the expectation values of these observables are well-behaved in the limit $\hbar \to 0$ in the sense that we recover the 'usual' expectation values.

Let us start with the classical observable position, given by the function

$$f\colon M = \mathbb{R} \times I \to \mathbb{R}, \quad (p, q) \mapsto q.$$

Since $I =]a, b[$ is a bounded interval, the function f is bounded as well, and $\|f\|_{L^\infty(I)} = \max(|a|, |b|)$. This is a very convenient property, as it implies that the equation above makes sense for each $\Phi \in L^2(I)$, unlike the case $I = \mathbb{R}$. Recall that in quantum mechanics, the expectation value $\langle x \rangle$ of the position is given by

$$\langle x \rangle = \int_a^b x |\Phi(x)|^2\, dx.$$

3.2.1 Proposition. *Let $\Phi \in L^2(I)$. Then for each $\hbar \in X$, the function $M \to \mathbb{R}$, given by*

$$(p,q) \mapsto q|\langle \phi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2,$$

is an element of $L^1(M, \mu_L)$, and we have

$$\lim_{\hbar\to 0} \hbar^{-1} \int_M d\mu_L(p,q)\, q|\langle \phi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2 = \int_a^b x|\Phi(x)|^2\, dx.$$

Proof. We already argued in the proof of part (3) of Proposition 3.1.2 that the functions

$$(p,q) \mapsto |\langle \Psi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2, \text{ and } (p,q) \mapsto |\langle \phi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2,$$

are elements of $L^1(M, \mu_L)$. The function $M \ni (p,q) \mapsto q \in I$ is continuous and bounded, so the functions

$$(p,q) \mapsto q|\langle \Psi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2, \text{ and } (p,q) \mapsto q|\langle \phi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2,$$

are also elements of $L^1(M, \mu_L)$. One can show in the same way as in said proof that

$$\begin{aligned}
\hbar^{-1} \int_M d\mu_L(p,q)\, q|\langle \Psi^{\hbar}_{(p,q)}, \Phi\rangle_{L^2(I)}|^2 &= (\pi\hbar)^{-1/2} \int_a^b \int_{\mathbb{R}} q e^{-(x-q)^2/\hbar} |\Phi(x)|^2\, dx\, dq \\
&= (\pi\hbar)^{-1/2} \int_a^b \int_a^b q e^{-(x-q)^2/\hbar} |\Phi(x)|^2\, dx\, dq
\end{aligned}$$

Reversing the order of integration and performing the substitution $s = q - x$ yields

$$\begin{aligned}
&(\pi\hbar)^{-1/2} \int_a^b \int_a^b q e^{-(x-q)^2/\hbar} |\Phi(x)|^2\, dx\, dq \\
&= (\pi\hbar)^{-1/2} \int_a^b \int_{a-x}^{b-x} (x+s) e^{-s^2/\hbar} |\Phi(x)|^2\, ds\, dx \\
&= \pi^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} (x + \hbar^{1/2} s) e^{-s^2} |\Phi(x)|^2\, ds\, dx.
\end{aligned}$$

Now note that

$$\pi^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} x e^{-s^2} |\Phi(x)|^2\, ds\, dx = \pi^{-1/2} \int_a^b x|\Phi(x)|^2 \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} e^{-s^2}\, ds\, dx.$$

Taking the limit $\hbar \to 0$ yields

$$\lim_{\hbar\to 0} \pi^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} x e^{-s^2} |\Phi(x)|^2\, ds\, dx = \int_a^b x|\Phi(x)|^2\, dx,$$

by dominated convergence and the fact that $\int_{\mathbb{R}} e^{-s^2}\, ds = \sqrt{\pi}$. Furthermore, note that

$$\begin{aligned}
&\left| \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} \hbar^{1/2} s e^{-s^2} |\Phi(x)|^2 \, ds \, dx \right| \\
&\quad = \hbar^{1/2} \left| \int_a^b -\frac{1}{2} \left[e^{-s^2} \right]_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} |\Phi(x)|^2 \, dx \right| \\
&\quad = \hbar^{1/2} \left| \int_a^b \frac{1}{2} \left(e^{-(a-x)^2/\hbar} - e^{-(b-x)^2/\hbar} \right) |\Phi(x)|^2 \, dx \right| \\
&\quad \le \hbar^{1/2} \int_a^b \frac{1}{2} \left(|e^{-(a-x)^2/\hbar}| + |e^{-(b-x)^2/\hbar}| \right) |\Phi(x)|^2 \, dx \\
&\quad \le \hbar^{1/2} \|\Phi\|^2_{L^2(I)},
\end{aligned}$$

hence

$$\lim_{\hbar \to 0} \pi^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} \hbar^{1/2} s e^{-s^2} |\Phi(x)|^2 \, ds \, dx = 0.$$

Thus

$$\lim_{\hbar \to 0} \pi^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} (x + \hbar^{1/2} s) e^{-s^2} |\Phi(x)|^2 \, ds \, dx = \int_a^b x |\Phi(x)|^2 \, dx,$$

and therefore,

$$\lim_{\hbar \to 0} \hbar^{-1} \int_M d\mu_L(p,q) \, q |\langle \Psi^{\hbar}_{(p,q)}, \Phi \rangle_{L^2(I)}|^2 = \int_a^b x |\Phi(x)|^2 \, dx.$$

One can now apply part (3) of Lemma 3.1.1 in the same way as in the proof of part (3) of Proposition 3.1.2 to deduce that

$$\lim_{\hbar \to 0} \hbar^{-1} \int_M d\mu_L(p,q) \, q |\langle \phi^{\hbar}_{(p,q)}, \Phi \rangle_{L^2(I)}|^2 = \int_a^b x |\Phi(x)|^2 \, dx.$$

■

Note that the limit $\hbar \to 0$ is not uniform in Φ. We do, however, have the following result:

3.2.2 Proposition. *Let* $(p_0, q_0) \in M$. *Then*

$$\lim_{\hbar \to 0} \hbar^{-1} \int_M d\mu_L(p,q) \, q |\langle \phi^{\hbar}_{(p,q)}, \phi^{\hbar}_{(p_0,q_0)} \rangle_{L^2(I)}|^2 = q_0.$$

Proof. Slightly modifying the proof of Proposition 3.2.1, one can show that

$$\lim_{\hbar \to 0} \hbar^{-1} \int_M d\mu_L(p,q) \, q |\langle \phi^{\hbar}_{(p,q)}, \phi^{\hbar}_{(p_0,q_0)} \rangle_{L^2(I)}|^2 - \int_a^b x |\phi^{\hbar}_{(p_0,q_0)}|^2 \, dx = 0.$$

In particular, we remark that

$$\lim_{\hbar \to 0} \pi^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} x e^{-s^2} |\phi^{\hbar}_{(p_0,q_0)}|^2 \, ds \, dx - \int_a^b x |\phi^{\hbar}_{(p_0,q_0)}|^2 \, dx = 0,$$

since the support of $\phi^\hbar_{(p_0,q_0)}$ gets smaller and smaller as $\hbar \to 0$. Moreover, by the Cauchy-Schwarz inequality and part (3) of Lemma 3.1.1 we have

$$\begin{aligned}
&\left| \int_a^b x|\Psi^\hbar_{(p_0,q_0)}|^2\,dx - \int_a^b x|\phi^\hbar_{(p_0,q_0)}|^2\,dx \right| \\
&\quad \le \max(|a|,|b|) \int_a^b ||\Psi^\hbar_{(p_0,q_0)}|^2 - |\phi^\hbar_{(p_0,q_0)}|^2\,dx \\
&\quad \le \max(|a|,|b|) \| |\Psi^\hbar_{(p_0,q_0)}| + |\phi^\hbar_{(p_0,q_0)}| \|_{L^2(I)} \|\Psi^\hbar_{(p_0,q_0)} - \phi^\hbar_{(p_0,q_0)}\|_{L^2(I)} \\
&\quad \le \max(|a|,|b|) (\|\Psi^\hbar_{(p_0,q_0)}\|_{L^2(I)} + \|\phi^\hbar_{(p_0,q_0)}| \|_{L^2(I)}) \|\Psi^\hbar_{(p_0,q_0)} - \phi^\hbar_{(p_0,q_0)}\|_{L^2(I)} \\
&\quad 4\max(|a|,|b|) e^{-\hbar^{-1}\alpha(\hbar,z)^{-2}/2}.
\end{aligned}$$

The last line converges to 0 as $\hbar \to 0$, so

$$\lim_{\hbar\to 0} \left(\int_a^b x|\Psi^\hbar_{(p_0,q_0)}|^2\,dx - \int_a^b x|\phi^\hbar_{(p_0,q_0)}|^2\,dx \right) = 0.$$

Finally, we have

$$\begin{aligned}
\int_a^b x|\Psi^\hbar_{(p_0,q_0)}|^2\,dx &= (\pi\hbar)^{-1/2} \int_a^b x e^{-(x-q_0)^2/\hbar}\,dx \\
&= \pi^{-1/2} \int_a^b (\hbar^{1/2}s + q_0) e^{-s^2}\,dx \\
&= \pi^{-1/2} q_0 \int_{\hbar^{-1/2}(a-q_0)}^{\hbar^{-1/2}(b-q_0)} e^{-s^2}\,ds + \frac{1}{2}\pi^{-1/2}\hbar^{1/2}(e^{-(a-q_0)^2/\hbar} - e^{-(b-q_0)^2/\hbar}).
\end{aligned}$$

The last line converges to q_0 as $\hbar \to 0$. Putting these results together, we obtain

$$\lim_{\hbar\to 0} \hbar^{-1} \int_M d\mu_L(p,q)\, q |\langle \phi^\hbar_{(p,q)}, \phi^\hbar_{(p_0,q_0)} \rangle_{L^2(I)}|^2 = q_0,$$

which is what we wanted to show. ∎

The other important observable of our system is of course momentum, which corresponds to the function $(p,q) \mapsto p$. Unlike position, momentum is not a bounded function, which forces us to restrict the class of functions among which we can consider the expectation value of the momentum. Recall that in quantum mechanics the expectation value $\langle p \rangle$ of the momentum is given by the formula

$$\langle p \rangle := -i\hbar \int_a^b \Phi'(x)\overline{\Phi(x)}\,dx.$$

3.2.3 Proposition. *Let $\hbar \in X$, let $\Phi \in H^1_0(I)$. Then the function $M \to \mathbb{R}$, given by $(p,q) \mapsto p|\langle \phi^\hbar_{(p,q)}, \Phi \rangle_{L^2(I)}|^2$, is an element of $L^1(M,\mu_L)$, and we have*

$$\lim_{\hbar\to 0} h^{-2} \int_M d\mu_L(p,q)\, p |\langle \phi^\hbar_{(p,q)}, \Phi \rangle_{L^2(I)}|^2 - \int_a^b -i\Phi'(x)\overline{\Phi(x)}\,dx = 0.$$

Proof. First, we consider the function $f_\Phi^\hbar \colon M \to \mathbb{R}$, given by

$$(p,q) \mapsto p|\langle \Psi^\hbar_{(p,q)}, \Phi\rangle_{L^2(I)}|^2 = p|\langle \Psi^\hbar_{(p,q)}, \Phi\rangle_{L^2(\mathbb{R})}|^2.$$

Here we remark that by part (4) of 1.2.4, the extension of Φ by zero is still weakly differentiable, and that its weak derivative is square integrable. Now note that

$$\begin{aligned}
&p|\langle \Psi^\hbar_{(p,q)}, \Phi\rangle_{L^2(\mathbb{R})}|^2 \\
&\quad = (\pi\hbar)^{-1/2} p\mathcal{F}(e^{-(x-q)^2/2\hbar}\Phi)(p/\hbar)\overline{\mathcal{F}(e^{-(x-q)^2/2\hbar}\Phi)(p/\hbar)} \\
&\quad = (\pi\hbar)^{-1/2} \mathcal{F}\left(-i\hbar\frac{d}{dx}(e^{-(x-q)^2/2\hbar}\Phi)\right)(p/\hbar)\overline{\mathcal{F}(e^{-(x-q)^2/2\hbar}\Phi)(p/\hbar)} \\
&\quad = (\pi\hbar)^{-1/2} \mathcal{F}\left(-i\hbar e^{-(x-q)^2/2\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\right)(p/\hbar)\overline{\mathcal{F}(e^{-(x-q)^2/2\hbar}\Phi)(p/\hbar)},
\end{aligned}$$

where $\mathcal{F}$ is again the Fourier transform with respect to x. But the last line is a product of two Fourier transforms of square-integrable functions, so the function $p \mapsto p|\langle \Psi^\hbar_{(p,q)}, \Phi\rangle_{L^2(\mathbb{R})}|^2$ is integrable. Let $B \geq \hbar$ be a constant such that $|(q-x)e^{-(x-q)^2/2\hbar}| \leq B$ for each $x \in \mathbb{R}$ and each $q \in]a,b[$. Then by the Cauchy-Schwarz inequality, the Parseval-Plancherel theorem and the triangle inequality, we have

$$\begin{aligned}
&(2\pi\hbar)^{-1}\int_{\mathbb{R}} \left|p|\langle \Psi^\hbar_{(p,q)}, \Phi\rangle_{L^2(\mathbb{R})}|^2\right| \, dp \\
&\quad = (2\pi\hbar)^{-1}(\pi\hbar)^{-1/2}\int_{\mathbb{R}} \left| \begin{array}{c} \mathcal{F}\left(-i\hbar e^{-(x-q)^2/2\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\right)(p/\hbar) \\ \cdot\overline{\mathcal{F}(e^{-(x-q)^2/2\hbar}\Phi)(p/\hbar)} \end{array} \right| dp \\
&\quad \leq (\pi\hbar)^{-1/2}B(\|\Phi\|_{L^2(\mathbb{R})} + \|\Phi'\|_{L^2(\mathbb{R})})\|\Phi\|_{L^2(\mathbb{R})} \\
&\quad \leq 2(\pi\hbar)^{-1/2}B\|\Phi\|_{H^1(I)}\|\Phi\|_{L^2(I)}.
\end{aligned}$$

Hence

$$(2\pi\hbar)^{-1}\int_a^b\int_{\mathbb{R}} |f_\Phi^\hbar| \, dp \, dq \leq 2(\pi\hbar)^{-1/2}(b-a)B\|\Phi\|_{H^1(I)}\|\Phi\|_{L^2(I)},$$

so by Tonelli's theorem, $f_\Phi^\hbar \in L^1(M,\mu_L)$, and by Fubini's theorem,

$$\begin{aligned}
&\hbar^{-2}\int_M d\mu_L(p,q)\, p|\langle \Psi^\hbar_{(p,q)}, \Phi\rangle_{L^2(\mathbb{R})}|^2, \\
&\quad = (2\pi\hbar)^{-1}(\pi\hbar)^{-1/2}\int_a^b\int_{\mathbb{R}} \begin{array}{c} \mathcal{F}\left(-ie^{-(x-q)^2/2\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\right)(p/\hbar) \\ \cdot\overline{\mathcal{F}(e^{-(x-q)^2/2\hbar}\Phi)(p/\hbar)} \end{array} \, dp\, dq \\
&\quad = (2\pi)^{-1}(\pi\hbar)^{-1/2}\int_a^b \begin{array}{c} \mathcal{F}\left(-ie^{-(x-q)^2/2\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\right) \\ * \mathcal{F}(e^{-(x-q)^2/2\hbar}\overline{\Phi})(0) \end{array} \, dq \\
&\quad = (\pi\hbar)^{-1/2}\int_a^b \mathcal{F}\left(-ie^{-(x-q)^2/\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\overline{\Phi}\right)(0)\, dq \\
&\quad = (\pi\hbar)^{-1/2}\int_a^b\int_{\mathbb{R}} -ie^{-(x-q)^2/\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\overline{\Phi}\, dx\, dq \\
&\quad = (\pi\hbar)^{-1/2}\int_a^b\int_a^b -ie^{-(x-q)^2/\hbar}\left(\frac{q-x}{\hbar}\Phi + \Phi'\right)\overline{\Phi}\, dx\, dq.
\end{aligned}$$

Reversing the order of integration and substituting $s = \hbar^{-1/2}(q-x)$, we find

$$\begin{aligned}
&(\pi\hbar)^{-1/2} \int_a^b \int_a^b -ie^{-(x-q)^2/\hbar} \frac{q-x}{\hbar} |\Phi(x)|^2 \, dx \, dq \\
&\quad = -i(\pi\hbar)^{-1/2} \int_a^b \int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} se^{-s^2} \, ds |\Phi(x)|^2 \, dx \\
&\quad = -i(\pi\hbar)^{-1/2} \int_a^b \frac{1}{2} \left(e^{-(a-x)^2/\hbar} - e^{-(b-x)^2/\hbar} \right) |\Phi(x)|^2 \, dx.
\end{aligned}$$

Because $\Phi \in H_0^1(I)$, we have $\Phi(a) = \Phi(b) = 0$, and from the Cauchy-Schwarz inequality, it follows that

$$|\Phi(x)|^2 = \left| \int_a^x \Phi'(s) \, ds \right|^2 = \left| \int_a^b \Phi'(s) \mathbf{1}_{[a,x]}(s) \, ds \right|^2 \leq \|\Phi'\|_{L^2(I)}^2 (x-a),$$

for each $x \in I$, where $\mathbf{1}_{[a,x]}$ denotes the characteristic function of the set $[a,x]$. Similarly, we have $|\Phi(x)|^2 \leq \|\Phi'\|_{L^2(I)}^2 (b-x)$. Hence

$$\begin{aligned}
\int_a^b e^{-(a-x)^2/\hbar} |\Phi(x)|^2 \, dx &\leq \|\Phi'\|_{L^2(I)}^2 \int_a^b e^{-(a-x)^2/\hbar} (x-a) \, dx \\
&= -\frac{1}{2} \hbar \|\Phi'\|_{L^2(I)}^2 \left[e^{-(x-a)^2/\hbar} \right]_a^b = \frac{1}{2} \hbar \|\Phi'\|_{L^2(I)}^2 \left(1 - e^{-(b-a)^2/\hbar} \right).
\end{aligned}$$

Similarly, we have

$$\int_a^b e^{-(b-x)^2/\hbar} |\Phi(x)|^2 \, dx \leq \frac{1}{2} \hbar \|\Phi'\|_{L^2(I)}^2 \left(1 - e^{-(b-a)^2/\hbar} \right),$$

so

$$(\pi\hbar)^{-1/2} \left| \int_a^b \int_a^b -ie^{-(x-q)^2/\hbar} \frac{q-x}{\hbar} |\Phi(x)|^2 \, dx \, dq \right| \leq \pi^{-1/2} \hbar^{1/2} \|\Phi'\|_{L^2(I)}^2 \left(1 - e^{-(b-a)^2/\hbar} \right).$$

The right-hand side of this inequality converges to 0 as $\hbar \to 0$, so

$$\lim_{\hbar \to 0} (\pi\hbar)^{-1/2} \int_a^b \int_a^b -ie^{-(x-q)^2/\hbar} \frac{q-x}{\hbar} |\Phi(x)|^2 \, dx \, dq = 0,$$

and therefore

$$\lim_{\hbar \to 0} \left(h^{-2} \int_M p |\langle \Psi_{(p,q)}^{\hbar}, \Phi \rangle_{L^2(I)}|^2 \, d\mu_L(p,q) - (\pi\hbar)^{-1/2} \int_a^b \int_a^b -ie^{-(x-q)^2/\hbar} \, dq \, \Phi'(x) \overline{\Phi(x)} \, dx \right) = 0.$$

One can use the same approximation as in the proof of part (3) of Proposition 3.1.2 to show that

$$\lim_{\hbar \to 0} \left(h^{-2} \int_M p |\langle \phi_{(p,q)}^{\hbar}, \Phi \rangle_{L^2(I)}|^2 \, d\mu_L(p,q) - (\pi\hbar)^{-1/2} \int_a^b \int_a^b -ie^{-(x-q)^2/\hbar} \, dq \, \Phi'(x) \overline{\Phi(x)} \, dx \right) = 0.$$

The second integral in the above expression can be written as

$$\pi^{-1/2}\int_a^b\int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} -ie^{-s^2}\, ds\, \Phi'(x)\overline{\Phi(x)}\, dx.$$

Thus it remains to be shown that

$$\lim_{\hbar\to 0}\pi^{-1/2}\int_a^b\int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} -ie^{-s^2}\, ds\, \Phi'(x)\overline{\Phi(x)}\, dx - \int_a^b -i\Phi'(x)\overline{\Phi(x)}\, dx = 0.$$

Now fix $\varepsilon > 0$. Then by Lebesgue's dominated convergence theorem, there exists an $r > 0$ such that

$$\int_{I\setminus[a+r,b-r]} |\Phi'(x)\overline{\Phi(x)}|\, dx < \varepsilon/2,$$

and by the same theorem, we can find a $\delta \in]0,1]$ such that for each $\hbar \in X$ with $\hbar < \delta$, we have

$$\left(1 - \pi^{-1/2}\int_{-\hbar^{-1/2}r}^{\hbar^{-1/2}r} e^{-s^2}\, ds\right)\int_{a+r}^{b-r} |\Phi'(x)\overline{\Phi(x)}|\, dx < \varepsilon/2,$$

so

$$\begin{aligned}
&\left|\pi^{-1/2}\int_a^b\int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} -ie^{-s^2}\, ds\, \Phi'(x)\overline{\Phi(x)}\, dx - \int_a^b -i\Phi'(x)\overline{\Phi(x)}\, dx\right| \\
&\leq \left|\int_a^b\left(\pi^{-1/2}\int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} e^{-s^2}\, ds - 1\right)\Phi'(x)\overline{\Phi(x)}\, dx\right| \\
&\leq \left(1 - \pi^{-1/2}\int_{-\hbar^{-1/2}r}^{\hbar^{-1/2}r} e^{-s^2}\, ds\right)\left(\int_a^b |\Phi'(x)\overline{\Phi(x)}|\, dx\right) \\
&\leq \int_{I\setminus[a+r,b-r]} |\Phi'(x)\overline{\Phi(x)}|\, dx \\
&\quad + \left(1 - \pi^{-1/2}\int_{-\hbar^{-1/2}r}^{\hbar^{-1/2}r} e^{-s^2}\, ds\right)\int_{a+r}^{b-r} |\Phi'(x)\overline{\Phi(x)}|\, dx \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon,
\end{aligned}$$

We conclude that

$$\lim_{\hbar\to 0}\pi^{-1/2}\int_a^b\int_{\hbar^{-1/2}(a-x)}^{\hbar^{-1/2}(b-x)} -ie^{-s^2}\, ds\, \Phi'(x)\overline{\Phi(x)}\, dx - \int_a^b -i\Phi'(x)\overline{\Phi(x)}\, dx = 0,$$

and this completes the proof of the proposition. ■

3.2.4 Remark. When the domain Ω is the real line or more generally, $\Omega = \mathbb{R}^n$, rather than a bounded open interval, then the theory of coherent states yields much stronger results. For instance, if $f \in \mathcal{S}(\mathbb{R}^{2n})$ is a function on the phase space $\mathbb{R}^{2n} \cong T^*\mathbb{R}^n$, and $t \mapsto f_t$ is the classical time evolution of that observable, with $f_0 = f$, then for each $(p_0, q_0) \in \mathbb{R}^{2n}$, we have [12, p. 477]

$$\lim_{\hbar\to 0}\hbar^{-n}\int_{\mathbb{R}^{2n}} d\mu_L(p,q)\, f(p,q)|\langle\Psi^\hbar_{(p,q)}, U(t)\Psi^\hbar_{(p_0,q_0)}\rangle|^2 = f_t(p_0,q_0).$$

The reason that the convergence properties of Schr̆odinger's states on $\mathbb{R}^n$ are much nicer, is probably due to the fact coherent states typically exist on locally compact topological groups (cf. [1]). In the present case, Schrödinger's states are defined on the locally compact group $(\mathbb{R}^n, +)$. Upon restriction to a proper open subset of $\mathbb{R}^n$, as we did above, one loses this group structure.

3.3 Time evolution of the coherent states

Having constructed good approximations $(\phi^{\hbar}_{(p,q)})$ to classical states, we wish to study their dynamical behaviour, that is, we want to find the solution Ψ to the Schrödinger equation with initial condition $\Psi(x,0) = \phi^{\hbar}_{(p,q)}(x)$ for each $x \in I$. To do so, we will follow the usual method: first, we solve the time-independent Schrödinger equation, i.e. the eigenvalue equation for a specific self-adjoint realisation $\widetilde{H}$ of the test Hamiltonian H with domain $C_0^\infty(]a,b[)$. This will allow us to explicitly compute the unitary evolution group $(U(t))_{t\in\mathbb{R}}$ associated to $\widetilde{H}$. Subsequently, we apply the unitary evolution group to the initial state $\phi^{\hbar}_{(p,q)}(x)$ to obtain the solution to the Schrödinger equation.

Our first objective is to show that the normalised eigenfunctions of any self-adjoint realisation $\widetilde{H}$ constitute an orthonormal basis of $L^2(]a,b[)$. It turns out that the essential spectrum is a useful tool.

3.3.1 Definition. Let $\mathcal{H}$ be a Hilbert space, and let T be an operator on $\mathcal{H}$.

- The *resolvent set* $\rho(T)$ of T is the collection of complex numbers $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is an isomorphism from $\mathcal{D}(T)$ onto $\mathcal{H}$ and $(T-\lambda I)^{-1}$ is a bounded operator on $\mathcal{H}$.
- The *spectrum* $\sigma(T)$ of T is the set $\mathbb{C}\backslash\rho(T)$.
- The *essential spectrum* $\sigma_{\text{ess}}(T)$ is the set consisting of all $\lambda \in \mathbb{C}$ such that either λ is an eigenvalue of T with infinite multiplicity, or for each $\varepsilon > 0$, there exists a $\mu \in \sigma(T)\backslash\{\lambda\}$ such that $|\lambda - \mu| < \varepsilon$.

3.3.2 Remark. It can be shown that $\rho(T)$ is open (cf. [4, Theorem 1.5.12]), and consequently, $\sigma(T)$ is closed, so $\sigma_{\text{ess}}(T) \subseteq \sigma(T)$.

Another useful fact about the essential spectrum is given by the following proposition, which can be found as Theorem 11.6.6 in [4]:

3.3.3 Proposition. *Let $\mathcal{H}$ be a Hilbert space, let T be a closed, hermitian operator on $\mathcal{H}$ and suppose that $\widetilde{T}_1$ and $\widetilde{T}_2$ are two self-adjoint extensions of T such that*

$$\dim(\mathcal{D}(\widetilde{T}_1)/\mathcal{D}(T)) < \infty, \quad \dim(\mathcal{D}(\widetilde{T}_2)/\mathcal{D}(T)) < \infty.$$

Then $\sigma_{\text{ess}}(\widetilde{T}_1) = \sigma_{\text{ess}}(\widetilde{T}_2)$.

Now we come to the reason for introducing the essential spectrum. The following theorem can be found as Theorem 11.3.13 in [4]:

3.3.4 Theorem. *Let $\mathcal{H}$ be a separable, infinite dimensional Hilbert space, and let T be a self-adjoint operator on $\mathcal{H}$. Then the following are equivalent:*

- $\sigma_{\text{ess}}(T) = \emptyset$;
- $\mathcal{H}$ *has an orthonormal basis* $(\phi_j)_{j=1}^{\infty}$ *of eigenvectors of* T *such that the corresponding set of eigenvalues* $(\lambda_j)_{j=1}^{\infty}$ *satisfies* $|\lambda_j| \to \infty$ *as* $j \to \infty$.

We are now ready to prove the following result:

3.3.5 Theorem. *Let* $H = D^2$, $\mathcal{D}(H) = C_0^{\infty}(]a,b[)$ *be the test Hamiltonian of the free particle on the interval* $]a,b[$, *and let* $\widetilde{H}$ *be a self-adjoint realisation of* H. *Then there exists an orthonormal basis* $(\phi_j)_{j=1}^{\infty}$ *of eigenvectors of* $\widetilde{H}$ *such that the corresponding set of eigenvalues* $(E_j)_{j=1}^{\infty}$ *satisfies* $|E_j| \to \infty$ *as* $j \to \infty$.

Proof. Let H_{periodic} be the realisation of H corresponding to periodic boundary conditions. Its eigenfunctions are the exponentials

$$(b-a)^{-\frac{1}{2}} e^{2\pi i k x/(b-a)}, \quad k \in \mathbb{Z}.$$

It is known from elementary Fourier analysis that these functions constitute an orthonormal basis of $L^2(]a,b[)$, and it is readily verified that the corresponding eigenvalues diverge, so by Theorem 3.3.4, it follows that $\sigma_{\text{ess}}(H_{\text{periodic}}) = \emptyset$.

From Theorem 2.3.7, we know that for each self-adjoint realisation $\widetilde{H}'$ of H, we have

$$\dim(\mathcal{D}(\widetilde{H}')/\mathcal{D}(H_{\min})) = \dim(\mathcal{G}(\widetilde{H}')/\mathcal{G}(H_{\min})) = 2,$$

so Proposition 3.3.3 yields $\sigma(\widetilde{H}) = \sigma_{\text{ess}}(H_{\text{periodic}}) = \emptyset$. Applying Theorem 3.3.4 with $T = \widetilde{H}$ gives the result. ■

Using the above theorem, one can show that the unitary evolution group $(U(t))_{t\in\mathbb{R}}$ associated to $\widetilde{H}$ satisfies

$$U(t) \sum_{j=1}^{\infty} \mu_j \phi_j = \sum_{j=1}^{\infty} e^{-itE_j} \mu_j \phi_j,$$

for each $t \in \mathbb{R}$ and each square summable sequence $(\mu_j)_{j=1}^{\infty}$ of complex numbers. If we incorporate the factor $\hbar$, then the equation becomes

$$U(t) \sum_{j=1}^{\infty} \mu_j \phi_j = \sum_{j=1}^{\infty} e^{-itE_{\hbar,j}/\hbar} \mu_j \phi_j,$$

where the subscript $\hbar$ in $E_{\hbar,j}$ indicates that the eigenvalues of the Hamiltonian $\widetilde{H} = \frac{\hbar^2}{2m} D^2$ depend on $\hbar$.
Before we use this result to compute the time evolution of $\phi^{\hbar}_{(p,q)}(x)$, we need one more fact:

3.3.6 Proposition. *Let* H *be the test Hamiltonian of the free particle on the interval* $[a,b]$. *If* $\phi \in \mathcal{D}(H_{\max})$ *is an eigenfunction of* $H_{\max}$, *then* $\phi \in C^{\infty}([a,b])$.

Proof. We prove by induction on n that $\phi \in H^{2n}(]a,b[)$. To prove the statement for $n = 0$, we simply remark that $\phi \in L^2(]a,b[)$. Now suppose that the statement is true for some $n \in \mathbb{N}_0$. Let E be the eigenvalue of ϕ, so that $D^2\phi = E\phi$. Applying the operator D^{2n} on both sides of the equation (in the distributional sense), we obtain

$$H_{\max}(D^{2n}\phi) = D^{2n}(H_{\max}\phi) = ED^n\phi \in L^2]a,b[.$$

We have seen in Example 2.3.12 that $\mathcal{D}(H_{\max}) = H^2(]a,b[)$, so $D^{2n}\phi \in H^2(]a,b[)$, which is equivalent to $D^{2n+2}\phi, D^{2n+1}\phi, D^{2n}\phi \in L^2]a,b[$. Since $\phi \in H^{2n}(]a,b[)$, it follows that $\phi \in H^{2(n+1)}(]a,b[)$. This completes the induction. Applying part (2) of Theorem 1.2.4 yields $\phi \in C^{2n-1}([a,b])$ for each $n \in \mathbb{N}_0$, hence $\phi \in C^\infty([a,b])$, as desired. ■

3.3.7 Remark. The above proposition can be generalised to domains in higher dimensional spaces and second-order elliptic operators of which the coefficients satisfy certain smoothness conditions. However, eigenfunctions of these operators are in general no longer smooth at the boundary. For details, we refer to [5], sections 6.3 and 6.5.

The previous proposition tells us that we may interpret the derivative in the eigenvalue equation

$$-\frac{d^2}{dx^2}\phi = E\phi,$$

or, including the constant $\hbar$ and the mass of the particle m,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi = E\phi,$$

as ordinary derivatives. But from the local existence and uniqueness of solutions to ordinary differential equations, it follows that for fixed E, the above equation has the general solution

$$\phi(x) = \begin{cases} Ae^{i\sqrt{2mE}x/\hbar} + Be^{i\sqrt{2mE}x/\hbar} & E > 0, \\ Ax + B & E = 0, \\ Ae^{\sqrt{-2mE}x/\hbar} + Be^{\sqrt{-2mE}x/\hbar} & E < 0, \end{cases}$$

where A and B are constants determined by the values of ϕ and ϕ' at a chosen point.

According to Theorem 3.3.5, the absolute values of the eigenvalues $(E_j)_{j=1}^\infty$ of the Hamiltonian diverge as $j \to \infty$. We can even say a bit more. First recall the following fact:

3.3.8 Lemma. *Let $I \subseteq \mathbb{R}$ be an interval, let $m, n \in \mathbb{N}$ with $m \leq n$, and let $f\colon I \to \mathbb{R}$ be a function that is m times differentiable. If f has n zeroes, then $f^{(m)}$ has at least $n-m$ zeroes.*

Proof. We prove the claim by induction on m. For $m = 0$, there is nothing to prove. Now suppose the claim is true for some $m \geq 0$, and suppose that $m+1 \leq n$ and that f is $m+1$ times differentiable. If $n = m+1$, then the induction step is trivial. Hence suppose that $m+1 < n$. Let $x_1 < x_2 < \ldots < x_{n-m}$ be zeroes of $f^{(m)}$. It follows from Rolle's theorem that for $j = 1, 2, \ldots, n-m-1$, there exists a $y_j \in]x_j, x_{j+1}[$ such that $f^{(m+1)}(y_j) = 0$. Thus $f^{(m+1)}$ has at least $n-m-1$ zeroes. This completes the induction. ■

3.3.9 Lemma. *Let $H = D^2$, $\mathcal{D}(H) = C_0^\infty(]-1,1[)$ be the test Hamiltonian of the free particle on the interval $]-1,1[$, and let $\widetilde{H}$ be a self-adjoint realisation of H with eigenvalues $(E_j)_{j=1}^\infty$, where the number of times that a single eigenvalue occurs in this sequence is equal to its multiplicity. Then:*

(1) The set $\{j \in \mathbb{N} \colon E_j < 0\}$ is finite.

(2) There exists a number $M \in \mathbb{N}$ such that for each $n \geq 0$, we have

$$|\{j \in \mathbb{N} \colon E_j > 0, \sqrt{E_j} \in [n\pi, (n+1)\pi[\}| \leq M.$$

Proof.
(1) Suppose that $\widetilde{H}$ is the self-adjoint realisation associated to the matrix

$$e^{i\theta} \begin{pmatrix} c & -\overline{d} \\ d & \overline{c} \end{pmatrix} \in \mathrm{U}(2),$$

as described in Example 2.3.8. Now let $k > 0$ be a real number, and suppose that $\phi = Ae^{ikx} + Be^{-ikx} \in \mathcal{D}(\widetilde{H})$. Then the following identity must hold:

$$\begin{pmatrix} \phi(-1) + i\phi'(-1) \\ \phi(1) - i\phi'(1) \end{pmatrix} = e^{i\theta} \begin{pmatrix} c & -\overline{d} \\ d & \overline{c} \end{pmatrix} \begin{pmatrix} \phi(-1) - i\phi'(-1) \\ \phi(1) + i\phi'(1) \end{pmatrix},$$

which means that

$$\begin{pmatrix} e^{-ik}(1-k) & e^{ik}(1+k) \\ e^{ik}(1+k) & e^{-ik}(1-k) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = e^{i\theta} \begin{pmatrix} c & -\overline{d} \\ d & \overline{c} \end{pmatrix} \begin{pmatrix} e^{-ik}(1+k) & e^{ik}(1-k) \\ e^{ik}(1-k) & e^{-ik}(1+k) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix},$$

and this equivalent to the statement that the following expression vanishes:

$$\begin{pmatrix} e^{i\theta}ce^{-ik}(1+k) + (-e^{i\theta}\overline{d}e^{ik} - e^{-ik})(1-k) & (-e^{i\theta}\overline{d}e^{-ik} - e^{ik})(1+k) + e^{i\theta}ce^{ik}(1-k) \\ (e^{i\theta}de^{-ik} - e^{ik})(1+k) + e^{i\theta}\overline{c}e^{ik}(1-k) & e^{i\theta}\overline{c}e^{-ik}(1+k) + (e^{i\theta}de^{ik} - e^{-ik})(1-k) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}.$$

There exist nontrivial $A, B \in \mathbb{R}$ such that the above expression vanishes if and only if the determinant of the matrix in this expression vanishes. It can be shown that this determinant is given by

$$\begin{aligned} &-2ie^{i\theta}(\sin(2k-\theta)(1+k)^2 + \sin(2k+\theta)(1-k)^2 - 4\mathrm{Im}(d)k - 2\mathrm{Re}(c)\sin(2k)(1-k^2)) \\ &= -4ie^{i\theta} \left(\begin{array}{l} (\cos(\theta) + \mathrm{Re}(c))k^2\sin(2k) - 2\sin(\theta)k\cos(2k) \\ +(\cos(\theta) - \mathrm{Re}(c))\sin(2k) - 2\mathrm{Im}(d)k \end{array} \right). \end{aligned}$$

Thus k^2 is an eigenvalue of $\widetilde{H}$ if and only if

$$\begin{aligned} 0 =& (\cos(\theta) + \mathrm{Re}(c))k^2\sin(2k) - 2\sin(\theta)k\cos(2k) \\ &+ (\cos(\theta) - \mathrm{Re}(c))\sin(2k) - 2\mathrm{Im}(d)k. \end{aligned} \tag{3.1}$$

Substituting $k \to -ik$, we see that $-k^2$ is an eigenvalue of $\widetilde{H}$ if and only if

$$\begin{aligned} 0 =& -(\cos(\theta) + \mathrm{Re}(c))k^2\sinh(2k) - 2\sin(\theta)k\cosh(2k) \\ &+ (\cos(\theta) - \mathrm{Re}(c))\sinh(2k) - 2\mathrm{Im}(d)k. \end{aligned} \tag{3.2}$$

Now suppose for the sake of contradiction that $\widetilde{H}$ has infinitely many negative eigenvalues $E_1', E_2', \ldots$, and let $k_j := \sqrt{-E_j'}$ for each $j \geq 1$. Then $k_j \to \infty$ as $j \to \infty$ by 3.3.5. On the other hand, consider the function $f \colon]0, \infty[\to \mathbb{R}$ given by

$$\begin{aligned} f(x) :=& -(\cos(\theta) + \mathrm{Re}(c))x^2\sinh(2x) - 2\sin(\theta)x\cosh(2x) \\ &+ (\cos(\theta) - \mathrm{Re}(c))\sinh(2x) - 2\mathrm{Im}(d)x. \end{aligned}$$

The right-hand side of the above equation is a sum of four terms. For large values of x, the absolute value of the first term will be larger than that of the sum of the remaining terms unless $\cos(\theta) + \mathrm{Re}(c) = 0$. Since $k_j \to \infty$ as $j \to \infty$, this implies $\cos(\theta) = -\mathrm{Re}(c)$. Now we observe that the second term in the equation above dominates the expression for large values of x unless $\sin(\theta) = 0$, so $\sin(\theta) = 0$, and consequently, $\cos(\theta) = \pm 1$. Finally, since the first and second terms vanish, the third term will dominate the expression for large values of x unless $\cos(\theta) = \mathrm{Re}(c)$. But then we have $\cos(\theta) = \mathrm{Re}(c) = -\cos(\theta)$, so $\cos(\theta) = 0$, which contradicts our earlier observation that $\cos(\theta) = \pm 1$. Thus $\widetilde{H}$ has only finitely many negative eigenvalues.

(2) For each $n \in \mathbb{N}$, let

$$M_n := |\{j \in \mathbb{N} \colon E_j > 0, \sqrt{E_j} \in [n\pi, (n+1)\pi[\}|.$$

Looking at equation (3.1), we identify the following cases:

- $\mathrm{Re}(c) = \cos(\theta) = 0$: In this case, the equation reduces to

  $$0 = k(\sin(\theta)\cos(2k) + \mathrm{Im}(d)).$$

  The assumption $\cos(\theta) = 0$ implies that $\sin(\theta) = \pm 1$, so $k = 0$ or $\cos(2k) = \frac{\mathrm{Im}(d)}{\sin(\theta)} = \pm\mathrm{Im}(d)$. Thus for a fixed $n \geq 1$, the equation has at most 2 solutions in $[n\pi, (n+1)\pi[$. Taking multiplicities into account, we infer that $M_n \leq 4$.
- $\mathrm{Re}(c) = -\cos(\theta) \neq 0$: Now the equation becomes

  $$0 = \mathrm{Re}(c)\sin(2k) + \sin(\theta)k\cos(2k) + \mathrm{Im}(d)k.$$

  Consider the function

  $$x \mapsto \mathrm{Re}(c)\sin(2x) + \sin(\theta)x\cos(2x) + \mathrm{Im}(d)x$$

  on the real line. Its second derivative is given by

  $$x \mapsto -4((\mathrm{Re}(c) + \sin(\theta))\sin(2x) + \sin(\theta)x\cos(2x)). \tag{3.3}$$

  - If $\sin(\theta) = 0$, then the second derivative has exactly two zeroes in $[n\pi, (n+1)\pi[$ for each $n \in \mathbb{N}$, so by Lemma 3.3.8, the function can have at most four zeroes in $[n\pi, (n+1)\pi[$, and therefore $M_n \leq 8$.
  - If $\sin(\theta) = -\mathrm{Re}(c)$, then $\sin(\theta) \neq 0$, and we can use the same argument has in the case $\sin(\theta) = 0$ to show that $M_n \leq 8$.
  - If $0 \neq \sin(\theta) \neq -\mathrm{Re}(c)$, then function in equation (3.3) vanishes if and only if

    $$0 = \tan(2x) + \frac{x\sin(\theta)}{\mathrm{Re}(c) + \sin(\theta)}. \tag{3.4}$$

    Consider the functions $f_m \colon](m - 1/2)\pi/2, (m + 1/2)\pi/2[\to \mathbb{R}$, $m \in \mathbb{Z}$, given by

    $$f_m(x) := \tan(2x) + \frac{x\sin(\theta)}{\mathrm{Re}(c) + \sin(\theta)},$$

with derivative

$$f_m'(x) := 2(\tan^2(2x) + 1) + \frac{\sin(\theta)}{\mathrm{Re}(c) + \sin(\theta)},$$

These derivatives each have at most two zeroes, so f_m has at most three zeroes. Hence equation (3.4) has at most three solutions on $](m-1/2)\pi/2, (m+1/2)\pi/2[$ for each $m \in \mathbb{Z}$, and therefore, equation (3.1) has at most five solutions in $](m-1/2)\pi/2, (m+1/2)\pi/2[$. Since

$$\begin{aligned}[n\pi, (n+1)\pi[&= \{(2n+1/2)\pi/2, (2n+3/2)\pi/2\} \\ &\qquad \cup \bigcup_{j=0}^{2}](2n+j-1/2)\pi/2, (2n+j+1/2)\pi/2[,\end{aligned}$$

it follows that equation (3.1) has at most seventeen solutions in $[n\pi, (n+1)\pi[$, so $M_n \leq 34$.

- Now suppose $\mathrm{Re}(c) \neq -\cos(\theta)$. Equation (3.1) can be written as

$$0 = ((\cos(\theta) + \mathrm{Re}(c))k^2 + \cos(\theta) - \mathrm{Re}(c))\sin(2k) - 2\sin(\theta)k\cos(2k) - 2\mathrm{Im}(d)k.$$

Consider the function

$$x \mapsto ((\cos(\theta) + \mathrm{Re}(c))x^2 + \cos(\theta) - \mathrm{Re}(c))\sin(2x) - 2\sin(\theta)x\cos(2x) - 2\mathrm{Im}(d)x.$$

Differentiating this function twice yields

$$\begin{aligned} x \mapsto & (-4(\cos(\theta) + \mathrm{Re}(c))x^2 + 2(\cos(\theta) + \mathrm{Re}(c)) - 4(\cos(\theta) - \mathrm{Re}(c)) + 8\sin(\theta))\sin(2x) \\ & \quad + (8\sin(\theta) + 8(\cos(\theta) + \mathrm{Re}(c)))x\cos(2x) \\ & = (-4(\cos(\theta) + \mathrm{Re}(c))x^2 - 2\cos(\theta) + 6\mathrm{Re}(c) - 8\sin(\theta))\sin(2x) \\ & \quad + 8(\sin(\theta) + \cos(\theta) + \mathrm{Re}(c))x\cos(2x). \end{aligned}$$

Define the polynomials $p(x) := 8(\sin(\theta) + \cos(\theta) + \mathrm{Re}(c))x$ and $q(x) := 4(\cos(\theta) + \mathrm{Re}(c))x^2 + 2\cos(\theta) - 6\mathrm{Re}(c) + 8\sin(\theta)$. Then for each $m \in \mathbb{Z}$, the second derivative of the function defined above is equal to 0 at $x \in](m-1/2)\pi/2, (m+1/2)\pi/2[$ with $q(x) \neq 0$ if and only if $\tan(2x) = p(x)/q(x)$. Observe that p is a polynomial of degree 1, whereas q is a polynomial of degree 2, so there exists an $r > 0$ such that

$$\left| \frac{d}{dx}(p(x)/q(x)) \right| < 2 \leq 2(\tan^2(2x) + 1) = \frac{d}{dx}(\tan(2x)),$$

for each $x \in \mathbb{R}$ with $|x| \geq r$ and $4x + \pi \notin 2\pi\mathbb{Z}$. It follows from the mean value theorem that for each $m \in \mathbb{N}$, with $(m-1/2)\pi/2 \geq r$, there exists at most one point in $](m-1/2)\pi/2, (m+1/2)\pi/2[$ where the second derivative of the function above vanishes, so the function itself vanishes at at most three points on this interval. Consequently, equation 3.1 has at most eleven solutions on the interval $[n\pi, (n+1)\pi[$ for each $n \in \mathbb{N}$ with $n \geq r$, and therefore $M_n \leq 22$. The number of eigenvalues E_j with square root smaller than r is finite by Theorem 3.3.5, so the assertion follows.

$\blacksquare$

To compute the time evolution of a modified state $\phi^{\hbar}_{(p,q)}$, we must decompose it with respect to the orthonormal basis $(\phi_j)_{j=1}^{\infty}$ of eigenvectors associated to some self-adjoint extension $\widetilde{H}$ of $H = \frac{\hbar^2}{2m}D^2$. To do this, it is convenient to compute the time evolution of the restriction of the corresponding coherent state $\Psi^{\hbar}_{(p,q)}$ to the given interval $]a,b[$. Because $U(t)$ is a unitary operator for each $t \in \mathbb{R}$, we have

$$\|U(t)(\phi^{\hbar}_{(p,q)} - \Psi^{\hbar}_{(p,q)})\|_{L^2(]a,b[)} = \|\phi^{\hbar}_{(p,q)} - \Psi^{\hbar}_{(p,q)}\|_{L^2(]a,b[)}.$$

In other words, with respect to the L^2-norm on $]a,b[$, the error is constant. The decomposition of $\Psi^{\hbar}_{(p,q)}$ with respect to the basis $(\phi^{\hbar}_j)_{j=1}^{\infty}$ is given by

$$\Psi^{\hbar}_{(p,q)} = \sum_{j=1}^{\infty} \langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(]a,b[)} \phi^{\hbar}_j.$$

But if $(p,q) \in \mathbb{R}\times]a,b[$, then for small values of $\hbar$, the function $\Psi^{\hbar}_{(p,q)}$ will be exponentially localised on $]a,b[$, so that we may approximate the coefficient $\langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(]a,b[)}$ by the inner product $\langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})}$. Now note that $\phi^{\hbar}_j$ is a linear combination of two exponential functions. In particular, if the eigenvalue E_j of ϕ_j is positive, then the latter inner product is a linear combination of Fourier transforms of $\Psi^{\hbar}_{(p,q)}$, and since the state $\Psi^{\hbar}_{(p,q)}$ is a Gaussian, it is (relatively) easy to compute these Fourier transforms explicitely. Let us make these statements precise:

3.3.10 Theorem. *Let $m \in]0,\infty[$, let $\widetilde{D^2}$ be a self-adjoint realisation of D^2 on the interval $I := [a,b]$, and let $(k_j)_{j=1}^{\infty}$ be a sequence of elements of $[0,\infty[\cup i[0,\infty[$ such that $(-k_j^2)_{j=1}^{\infty}$ is the monotone increasing sequence of eigenvalues of D^2 in which the number of times that a certain eigenvalue appears in the sequence is equal to its multiplicity. Furthermore, let $(\phi_j)_{j=1}^{\infty}$ be a corresponding orthonormal basis of eigenvectors of $\widetilde{D^2}$, with*

$$\phi_j(x) = \begin{cases} A_{j,+}e^{k_j x} + A_{j,-}e^{-k_j x} & k_j \neq 0, \\ A_{j,+} + A_{j,-}x & k_j = 0, \end{cases}$$

and define the functions $(\overline{\phi}_j)_{j=1}^{\infty}$ by

$$\overline{\phi}_j(x) := \begin{cases} \overline{A_{j,+}}e^{\overline{k_j} x} + \overline{A_{j,-}}e^{-\overline{k_j} x} & k_j \neq 0, \\ \overline{A_{j,+}} + \overline{A_{j,-}}x & k_j = 0, \end{cases}$$

For each $\hbar \in [0,\infty[$, define the self-adjoint operator

$$\widetilde{H}_{\hbar} := \frac{\hbar^2}{2m}\widetilde{D^2},$$

and let $(U_{\hbar}(t))_{t\in\mathbb{R}}$ be its associated unitary evolution group. Then for each $t \in \mathbb{R}$ and each $(p,q) \in \mathbb{R}\times]a,b[$, we have

$$\lim_{\hbar\to 0}\left(U_{\hbar}(t)\phi^{\hbar}_{(p,q)} - (4\pi\hbar)^{1/4}e^{ip(q+ip)/2\hbar}\sum_{j=1}^{\infty}\overline{\phi}_j(q+ip)e^{\hbar k_j^2(1+it/m)/2}\phi_j \right) = 0.$$

Proof. First note that by part (1) of Lemma 3.3.9, $\widetilde{D^2}$ has a smallest eigenvalue, which justifies the existence of the sequence $(k_j)_{j=1}^{\infty}$ with the properties listed in the theorem. We prove the theorem in five steps:

(I) Let $c \in \mathbb{C}$, let $c_1 := \mathrm{Re}(c)$ and let $c_2 := \mathrm{Im}(c)$. Then for each $\hbar \in]0, \infty[$, we have

$$\begin{aligned}
\int_{\mathbb{R}} \Psi^{\hbar}_{(p,q)}(x) e^{cx}\, dx &= (\pi\hbar)^{-1/4} e^{-ipq/2\hbar} \int_{\mathbb{R}} e^{-(x-q)^2/2\hbar + c_1 x} e^{i(p/\hbar + c_2)x}\, dx \\
&= (\pi\hbar)^{-1/4} e^{ipq/2\hbar} e^{q(c_1+ic_2)} \int_{\mathbb{R}} e^{-x^2/2\hbar + c_1 x} e^{i(p/\hbar+c_2)x}\, dx \\
&= (\pi\hbar)^{-1/4} e^{ipq/2\hbar} e^{q(c_1+ic_2)} e^{c_1^2\hbar/2} \int_{\mathbb{R}} e^{-(x-c_1\hbar)^2/2\hbar} e^{i(p/\hbar+c_2)x}\, dx \\
&= (\pi\hbar)^{-1/4} e^{ipq/2\hbar} e^{q(c_1+ic_2)} e^{c_1^2\hbar/2} e^{ic_1(p-c_2\hbar)} \int_{\mathbb{R}} e^{-x^2/2\hbar} e^{i(p/\hbar+c_2)x}\, dx \\
&= \hbar^{1/2} (\pi\hbar)^{-1/4} e^{ipq/2\hbar} e^{q(c_1+ic_2)} e^{c_1^2\hbar/2} e^{ic_1(p-c_2\hbar)} \int_{\mathbb{R}} e^{-x^2/2} e^{i(p/\hbar+c_2)\hbar^{1/2}x}\, dx \\
&= (2\pi\hbar)^{1/2} (\pi\hbar)^{-1/4} e^{ipq/2\hbar} e^{q(c_1+ic_2)} e^{c_1^2\hbar/2} e^{ic_1(p-c_2\hbar)} e^{-\hbar(p/\hbar+c_2)^2/2} \\
&= (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} e^{c(q+ip)} e^{\hbar c^2/2}.
\end{aligned}$$

Since $\overline{e^{cx}} = e^{\bar{c}x}$ for each $x \in \mathbb{R}$, we have

$$\langle e^{cx}, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} = (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} e^{\bar{c}(q+ip)} e^{\hbar \bar{c}^2/2},$$

so

$$\langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} = (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} e^{\hbar \overline{k_j}^2/2} \overline{\phi_j}(q+ip), \tag{3.5}$$

whenever $k_j \neq 0$ or if $k_j = 0$ and ϕ_j is constant. Moreover, using partial integration, we see that

$$\begin{aligned}
\int_{\mathbb{R}} \Psi^{\hbar}_{(p,q)}(x)(x-q)\, dx &= (\pi\hbar)^{-1/4} e^{-ipq/2\hbar} \int_{\mathbb{R}} (x-q) e^{-(x-q)^2/2\hbar} e^{ip/\hbar x}\, dx \\
&= (\pi\hbar)^{-1/4} e^{ipq/2\hbar} \int_{\mathbb{R}} x e^{-x^2/2\hbar} e^{ip/\hbar x}\, dx \\
&= \hbar^{1/2} (\pi\hbar)^{-1/4} e^{ipq/2\hbar} \int_{\mathbb{R}} x e^{-x^2/2} e^{ip\hbar^{-1/2}x}\, dx \\
&= ip\hbar (\pi\hbar)^{-1/4} e^{ipq/2\hbar} \int_{\mathbb{R}} e^{-x^2/2} e^{ip\hbar^{-1/2}x}\, dx \\
&= ip(2\pi\hbar)(\pi\hbar)^{-1/4} e^{ipq/2\hbar} e^{-p^2/2\hbar} \\
&= ip(4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar},
\end{aligned}$$

so

$$\langle x, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} = (q+ip)(4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar},$$

which means that equation (3.5) holds for each $j \in \mathbb{N}$. Since k_j is either real or purely imaginary, we have $\overline{k_j}^2 = k_j^2$, so equation (3.5) may be written as

$$\langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} = (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} e^{\hbar k_j^2/2} \overline{\phi_j}(q+ip).$$

(II) Next, we examine the difference $\langle \phi_j, \Psi^\hbar_{(p,q)} \rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^\hbar_{(p,q)} \rangle_{L^2(I)}$. First, suppose that $c \in \mathbb{R}$. Then, we have

$$\begin{aligned}
\left| \int_{\mathbb{R}\setminus[a,b]} e^{-(x-q)^2/2\hbar + cx} e^{ipx/\hbar}\, dx \right| &\leq \int_{\mathbb{R}\setminus[a,b]} e^{-(x-q)^2/2\hbar + cx}\, dx \\
&= e^{cq} \int_{\mathbb{R}\setminus[a-q,b-q]} e^{-x^2/2\hbar + cx}\, dx \\
&= e^{cq + c^2\hbar/2} \int_{\mathbb{R}\setminus[a-q,b-q]} e^{-(x-c\hbar)^2/2\hbar}\, dx \\
&= e^{cq + c^2\hbar/2} \int_{\mathbb{R}\setminus[a-q-c\hbar,b-q-c\hbar]} e^{-x^2/2\hbar}\, dx \\
&= \hbar^{1/2} e^{cq + c^2\hbar/2} \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q-c\hbar),\hbar^{-1/2}(b-q-c\hbar)]} e^{-x^2/2}\, dx,
\end{aligned}$$

for each $\hbar > 0$. Let

$$C_{1,\hbar,q,c} := \hbar^{1/2} \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q-c\hbar),\hbar^{-1/2}(b-q-c\hbar)]} e^{-x^2/2}\, dx,$$

and note that it converges to 0 as $\hbar \to 0$. Furthermore, using the estimates above, we see that

$$\begin{aligned}
\left| \int_{\mathbb{R}\setminus[a,b]} e^{-(x-q)^2/2\hbar} e^{i(p/\hbar + c)x}\, dx \right| &\leq \int_{\mathbb{R}\setminus[a,b]} e^{-(x-q)^2/2\hbar}\, dx \\
&= \hbar^{1/2} \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q),\hbar^{-1/2}(b-q)]} e^{-x^2/2}\, dx,
\end{aligned}$$

for each $\hbar > 0$. Let us denote the expression in the last line by $C_{2,\hbar,q}$, and observe that $\lim_{\hbar\to 0} \hbar^{-1/2} C_{2,\hbar,q} = 0$. Suppose now in addition that $|c| > |p/\hbar|$ for some fixed value of $\hbar \in (0,\infty)$. Using partial integration, we obtain the following estimate:

$$\begin{aligned}
&\left| \int_{\mathbb{R}\setminus[a,b]} e^{-(x-q)^2/2\hbar} e^{i(p/\hbar + c)x}\, dx \right| \\
&\quad = \hbar^{1/2} \left| \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q),\hbar^{-1/2}(b-q)]} e^{-x^2/2} e^{i(p/\hbar + c)\hbar^{1/2}x}\, dx \right| \\
&\quad = |c + p/\hbar|^{-1} \left| \left[e^{-x^2/2} e^{i(p/\hbar+c)\hbar^{1/2}x} \right]_{\hbar^{-1/2}(a-q)}^{\hbar^{-1/2}(b-q)} + \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q),\hbar^{-1/2}(b-q)]} x e^{-x^2/2} e^{i(p/\hbar+c)\hbar^{1/2}x}\, dx \right| \\
&\quad \leq (|c| - |p/\hbar|)^{-1} \left(e^{-(b-q)^2/\hbar} + e^{-(a-q)^2/\hbar} + \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q),\hbar^{-1/2}(b-q)]} |x| e^{-x^2/2}\, dx \right).
\end{aligned}$$

Let us call the expression between parentheses in the last line $C_{3,\hbar,q}$, and observe that $\lim_{\hbar\to 0} C_{3,\hbar,q} = 0$. Finally, let us remark that

$$\left| \int_{\mathbb{R}\setminus[a,b]} x e^{-(x-q)^2/2\hbar} e^{ipx/\hbar}\, dx \right| \leq \hbar^{1/2} \int_{\mathbb{R}\setminus[\hbar^{-1/2}(a-q),\hbar^{-1/2}(b-q)]} (\hbar^{1/2}|x| + |q|) e^{-x^2/2}\, dx.$$

Call the expression on the right-hand side $C_{4,\hbar,q}$, and note that it converges to 0 as $\hbar \to 0$.

(III) In addition to the previous estimates, we require an estimate for the coefficients $A_{j,+}$ and $A_{j,-}$. If $-ik_j > 0$, let $c := -ik_j$. Then, we have

$$
\begin{aligned}
1 = \|\phi_j\|_{L^2(I)}^2 &= \int_a^b |A_{j,+}e^{icx} + A_{j,-}e^{-icx}|^2 \, dx \\
&= \int_a^b |A_{j,+}|^2 + |A_{j,-}|^2 + A_{j,+}\overline{A_{j,-}}e^{2icx} + \overline{A_{j,+}}A_{j,-}e^{-2icx} \, dx \\
&= (b-a)(|A_{j,+}|^2 + |A_{j,-}|^2) + (2ic)^{-1}\left[A_{j,+}\overline{A_{j,-}}e^{2icx} - \overline{A_{j,+}}A_{j,-}e^{-2icx}\right] \\
&= (b-a)(|A_{j,+}|^2 + |A_{j,-}|^2) + (2ic)^{-1}\left(A_{j,+}\overline{A_{j,-}}(e^{2icb} - e^{2ica}) + \overline{A_{j,+}}A_{j,-}(e^{-2ica} - e^{-2icb})\right) \\
&= (b-a)(|A_{j,+}|^2 + |A_{j,-}|^2) + c^{-1}\sin((b-a)c)(A_{j,+}\overline{A_{j,-}}e^{ic(a+b)} + \overline{A_{j,+}}A_{j,-}e^{-ic(a+b)}) \\
&\geq (b-a)(|A_{j,+}|^2 + |A_{j,-}|^2) - 2c^{-1}|A_{j,+}||A_{j,-}| \\
&\geq (b-a-c^{-1})(|A_{j,+}|^2 + |A_{j,-}|^2).
\end{aligned}
$$

Thus, if $c > (b-a)^{-1}$, then $|A_{j,+}|^2 + |A_{j,-}|^2 \leq (b-a-c^{-1})^{-1}$, and hence $(|A_{j,+}| + |A_{j,-}|)^2 \leq 2(b-a-c^{-1})^{-1}$.

(IV) Now we examine the series

$$
\sum_{j=1}^{\infty} |\langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(I)}|^2. \tag{3.6}
$$

From our estimates in part (II) and the triangle inequality, it follows that

$$
\begin{aligned}
&\sum_{j\in\mathbb{N};k_j>0} |\langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(I)}|^2 \\
&\qquad \leq e^{k_1|q| + k_1^2\hbar/2} \sum_{j\in\mathbb{N};k_j>0} (|A_{j,+}|C_{1,\hbar,q,k_j} + |A_{j,-}|C_{1,\hbar,q,-k_j})^2.
\end{aligned}
$$

Note that the sums in this inequality are finite by part (1) of Lemma 3.3.9, and that the right-hand side converges to 0 as $\hbar \to 0$ because $C_{1,\hbar,q}$ does so. Next, we note that

$$
\sum_{j\in\mathbb{N};k_j=0} |\langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(I)}|^2 \leq \sum_{j\in\mathbb{N};k_j=0} (|A_{j,+}|C_{2,\hbar,q} + |A_{j,-}|C_{4,\hbar,q})^2.
$$

The sums in this inequality consist of at most two terms, and the right-hand side converges to 0 as $\hbar \to 0$. Now assume that $p \neq 0$. Fix $\hbar_0 > 0$ such that $|p|/\hbar_0 \geq 2(b-a)^{-1}$. It follows from part (2) of Lemma 3.3.9 that there exist constants $M > 0$ such that

$$
|\{j \in \mathbb{N} \colon k_j \in i[|p|/\hbar_0 + 1, |p|/\hbar + 1[\}| \leq M/\hbar,
$$

for each $\hbar > 0$, so using our estimates obtained in (II) and (III), we deduce that

$$
\begin{aligned}
&\sum_{j\in\mathbb{N};k_j\in i]0,|p|/\hbar+1[} |\langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^{\hbar}_{(p,q)}\rangle_{L^2(I)}|^2 \\
&\quad \leq \sum_{j\in\mathbb{N};k_j\in i]0,|p|/\hbar+1[} (|A_{j,+}| + |A_{j,-}|)^2 C_{2,\hbar,q}^2 \\
&\quad \leq C_{2,\hbar,q}^2 \left(4M((b-a)\hbar)^{-1} + \sum_{j\in\mathbb{N};k_j\in i]0,|p|/\hbar_0+1[} (|A_{j,+}| + |A_{j,-}|)^2\right),
\end{aligned}
$$

and the last line converges to 0 as $\hbar \to 0$. Finally, we estimate the part of the series corresponding to the part of the sequence $(k_j)_{j=1}^{\infty}$ with $|k_j| \geq |p|/\hbar + 1$ and $\hbar < \hbar_0$. Again by part (2) of Lemma 3.3.9, there exists a natural number N such that for each $n \in \mathbb{N}$, we have

$$|\{j \in \mathbb{N} \colon k_j \in i[|p|/\hbar + n, |p|/\hbar + n + 1[\}| \leq N,$$

so

$$\sum_{j \in \mathbb{N}; k_j \in i[|p|/\hbar + 1, \infty[} |\langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(I)}|^2 \leq 4(b-a)^{-1} N C^2_{3,\hbar,q} \sum_{j=1}^{\infty} j^{-2},$$

which converges to 0 as $\hbar \to 0$. Thus for $p \neq 0$, the series in equation 3.6 is well-defined and converges to 0 as $\hbar \to 0$. The same statement holds true for the slightly easier case $p = 0$, and is left to the reader.

(V) We are now in the position to prove the theorem. It follows from (IV) that

$$\sum_{j=1}^{\infty} (\langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} - \langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(I)}) \phi_j \in L^2(I),$$

and that this element converges to 0 as $\hbar \to 0$. Hence the expression

$$\sum_{j=1}^{\infty} \langle \phi_j, \Psi^{\hbar}_{(p,q)} \rangle_{L^2(\mathbb{R})} \phi_j = (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} \sum_{j=1}^{\infty} e^{\hbar k_j^2/2} \overline{\phi_j}(q + ip) \phi_j,$$

defines an element of $L^2(I)$, and

$$\lim_{\hbar \to 0} \left(\Psi^{\hbar}_{(p,q)} - (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} \sum_{j=1}^{\infty} e^{\hbar k_j^2/2} \overline{\phi_j}(q + ip) \phi_j \right) = 0.$$

From part (3) of Lemma 3.1.1, we know that $\lim_{\hbar \to 0} \phi^{\hbar}_{(p,q)} - \Psi^{\hbar}_{(p,q)} = 0$, so

$$\lim_{\hbar \to 0} \left(\phi^{\hbar}_{(p,q)} - (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} \sum_{j=1}^{\infty} e^{\hbar k_j^2/2} \overline{\phi_j}(q + ip) \phi_j \right) = 0.$$

If we apply the unitary evolution group $(U_\hbar(t))_{t \in \mathbb{R}}$ to the expression within the limit, then the L^2-norm of that expression remains unchanged, and therefore

$$\lim_{\hbar \to 0} \left(U_\hbar(t) \phi^{\hbar}_{(p,q)} - (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} \sum_{j=1}^{\infty} e^{\hbar k_j^2 (1+it/m)/2} \overline{\phi_j}(q + ip) \phi_j \right) = 0,$$

as desired. ■

The formula

$$(4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} \sum_{j=1}^{\infty} e^{\hbar k_j^2 (1+it/m)/2} \overline{\phi_j}(q + ip) \phi_j, \tag{3.7}$$

prompts the question whether it can be simplified for realisations of the Hamiltonian for which there exists a 'nice' orthonomal basis of eigenfunctions $(\phi_j)_{j=1}^{\infty}$. The most logical choice here would be to consider H_{periodic}, the realisation corresponding to periodic boundary conditions. However, even in this case, where formula (3.7) becomes a Fourier series, it still cannot be simplified. Thus we must resort to other means to study the time evolution of our states $\phi_{(p,q)}^{\hbar}$, which we shall do in the next subsection.

It is worth noting that formula (3.7) does tell us which energies determine the time evolution of $\phi_{(p,q)}^{\hbar}$. For each $j \in \mathbb{N}$, let

$$c_j := (4\pi\hbar)^{1/4} e^{ip(q+ip)/2\hbar} e^{\hbar k_j^2/2} \overline{\phi_j}(q+ip).$$

Now consider a $j \in \mathbb{N}$ such that k_j is purely imaginary, and let $\kappa := -ik_j$. Then

$$|c_j| = (4\pi\hbar)^{1/4} e^{-p^2/2\hbar} e^{-\hbar\kappa^2/2} |\overline{A_{j,+}} e^{\kappa(-iq+p)} - \overline{A_{j,-}} e^{\kappa(iq-p)}|.$$

If $p \geq 0$, then this is roughly equal to

$$(4\pi\hbar)^{1/4} |A_{j,+}| e^{-p^2/2\hbar} e^{-\hbar\kappa^2/2} e^{\kappa p} = (4\pi\hbar)^{1/4} |A_{j,+}| e^{-\hbar(p/\hbar-\kappa)^2/2}.$$

If we now assume that all values of $A_{j,+}$ and $A_{j,-}$ with $j \geq 1$ are of the same order of magnitude, then it follows that for fixed $\hbar > 0$, the time evolution of the system is mainly determined by the energies corresponding to the values of j with $|k_j| \approx |p|/\hbar$, or more accurately, the difference between $|k_j|$ and $|p|/\hbar$ is of order $\hbar^{-1/2}$. This statement remains true if we assume that $p \leq 0$, and is useful if one wishes to numerically approximate the time evolution of the system.

3.4 MATLAB simulations

Having attempted to study the time evolution of the states $\phi_{(p,q)}^{\hbar}$ analytically without success, we shall now employ numerical methods to examine the behaviour of these states. More specifically, we have written and executed MATLAB-programs to study the action of various unitary evolution groups on $\phi_{(p,q)}^{\hbar}$. The unitary evolution groups that we have analysed correspond to the self-adjoint realisations of the Hamiltonian $\frac{\hbar^2}{2m}D^2$ on the interval $[0,1]$ associated to the following boundary conditions:

- $\phi(0) = \phi(1) = 0$ (Dirichlet boundary conditions), with orthonormal basis and corresponding energies
$$\phi_j(x) := \sqrt{2}\sin(j\pi x), \quad E_j = \frac{(j\pi\hbar)^2}{2m}, \quad (j \geq 1).$$
- $\phi'(0) = \phi'(1) = 0$ (Neumann boundary conditions), with orthonormal basis and corresponding energies
$$\phi_j(x) := \sqrt{2}\cos(j\pi x), \quad E_j = \frac{(j\pi\hbar)^2}{2m}, \quad (j \geq 0).$$
- $\phi(0) = i\phi'(1)$, $\phi(1) = i\phi'(0)$. This set of boundary conditions is somewhat peculiar, since it has an eigenfunction that has negative energy:
$$\phi_0(x) := (e^2-1)^{-1/2}(e^x + ie^{1-x}), \quad E_0 = -\frac{\hbar^2}{2m}.$$

All of the other eigenfunctions orthogonal to the one above have positive energy. Together with ϕ_0, the following functions constitute an orthonormal basis of eigenvectors:

$$\phi_j(x) := (2((j\pi)^2+1))^{-1/2}(((-1)^j j\pi-1)e^{j\pi x}+((-1)^j j\pi+1)e^{-j\pi x}), \quad E_j = \frac{(j\pi\hbar)^2}{2m},$$

where $j \geq 1$.

- $\phi(1) = e^{i\theta}\phi(0)$, $\phi'(1) = e^{i\theta}\phi'(0)$ ('generalised' periodic boundary conditions), where $\theta \in [0, 2\pi[$, with orthonormal basis and corresponding energies

$$\phi_j(x) := e^{i(2\pi j+\theta)x}, \quad E_j = \frac{(\hbar(2\pi j+\theta))^2}{2m}, \quad (j \in \mathbb{Z}).$$

Basically, each of the programs performs the following four actions:

(1) First, all variables such as the initial position q_0, initial momentum p_0, time stepsize dt, etcetera are initialised. Furthermore, the program divides the interval $[0, 1]$ into a number of smaller intervals of equal length, computes the value of $\phi^{\hbar}_{(p_0,q_0)}$ at the end points of those intervals and stores these values in a vector.

(2) In the second step, the function $\phi^{\hbar}_{(p_0,q_0)}$ is decomposed with respect to part of one of the orthonormal bases listed above by approximating the L^2-inner product on $[0, 1]$ with a finite sum. The approximations of these coefficients are again stored in a vector.

(3) With these approximations, the program determines the time evolution of our state $\phi^{\hbar}_{(p_0,q_0)}$ between prespecified values of the start time T_{start} and the end time T_{start}. More specifically, the program approximates $U_{\hbar}(T_{\text{start}} + n \cdot dt)\phi^{\hbar}_{(p_0,q_0)}$, where $n \in \mathbb{N}_0$ satisfies $0 \leq n \cdot dt \leq T_{\text{end}} - T_{\text{start}}$. The result of these computations is stored in a matrix.

(4) In the final step, the program produces a number of frames, each associated to a point in time $T_{\text{start}} + n \cdot dt$ and containing two diagrams, as seen in figure 1. The diagram on the left displays the square of the absolute value of $U_{\hbar}(T_{\text{start}} + n \cdot dt)\phi^{\hbar}_{(p_0,q_0)}$, with on the horizontal axis the position q ranging from 0 to 1. The diagram on the right is a density plot of the Husimi function of $U_{\hbar}(T_{\text{start}}+n\cdot dt)\phi^{\hbar}_{(p_0,q_0)}$ in a part of phase space. The horizontal axis corresponds to position, whereas the vertical axis corresponds to momentum. The colour blue indicates that the value of the Husimi function is relatively small, whereas red means that this value is large at a certain point. In the figure, the Husimi function is concentrated at the point $(0.4, 0.5)$, which is consistent with the values of q_0 and p_0.

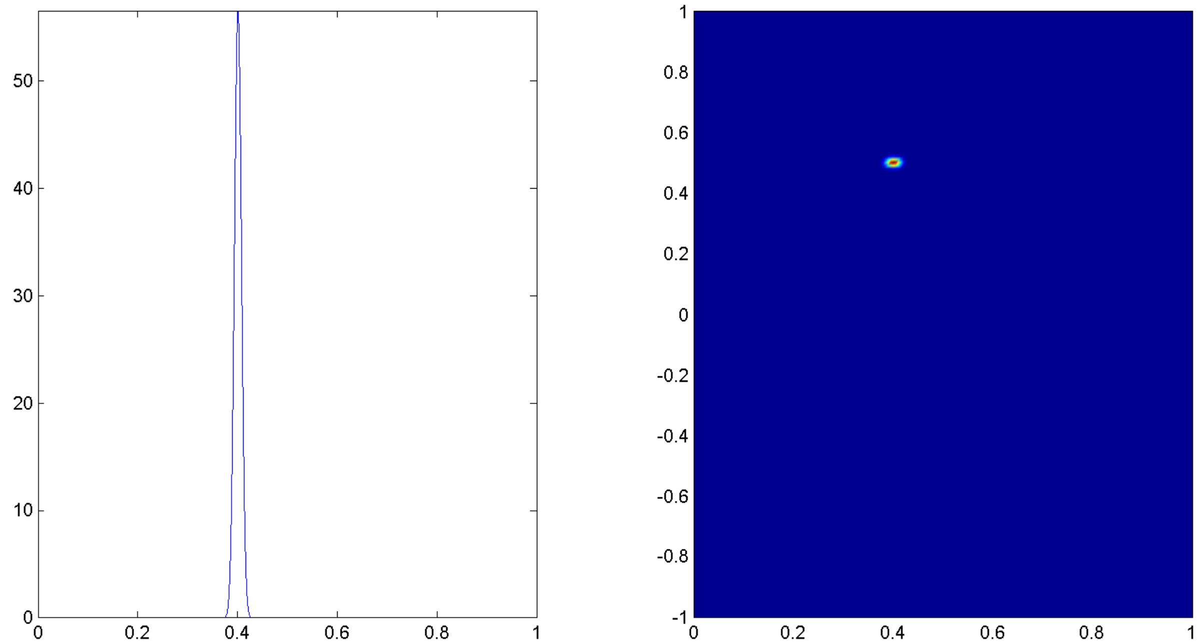


Figure 1: One of the frames produced by the program. In the diagram on the left, the square of the absolute value of the function $\phi^{\hbar}_{(p_0,q_0)}$ is displayed, with $\hbar = 10^{-4}$, $p_0 = 0.5$ and $q_0 = 0.4$. To the right, the Husimi function of $\phi^{\hbar}_{(p_0,q_0)}$ has been plotted.

We have chosen to study the aforementioned realisations because their eigenfunctions can be computed by hand, and therefore do not need to be approximated numerically. The upshot of this is that the difference in L^2-norm between the actual function $\phi^{\hbar}_{(p_0,q_0)}$, and the approximation that is obtained by decomposing the function with respect to (part of) one of the above orthonormal bases, is time independent. More precisely, let $(c_j)_j$ be the sequence of coefficients obtained by the program in step (2). Strictly speaking, this is a finite sequence. Let us extend this sequence by zero. Then the program effectively works with the function $\sum_j c_j \phi_j$, as opposed to the actual function $\phi^{\hbar}_{(p_0,q_0)} = \sum_j \langle \phi_j, \phi^{\hbar}_{(p_0,q_0)} \rangle_{L^2(]0,1[)} \phi_j$. The difference in L^2-norm is given by $(\sum_j |c_j - \langle \phi_j, \phi^{\hbar}_{(p_0,q_0)} \rangle_{L^2(]0,1[)}|^2)^{1/2}$, and this difference remains constant when we apply the unitary operator $U_{\hbar}(t)$ to both the actual function and its approximation.

Let us discuss the results of the simulations. We examined the behaviour of the function $\phi^{\hbar}_{(p_0,q_0)}$ with $q_0 = 0.4$, $p_0 = 0.5$ and $\hbar = 10^{-4}$ for $(T_{\text{start}}, T_{\text{end}}, dt) = (0, 4, 0.025)$ and $(T_{\text{start}}, T_{\text{end}}, dt) = (0, 2000, 8)$, both with mass $m = 1$. The simulations showed that the behaviour of both the square of the absolute value of the wave function, and the Husimi function, are nearly identical for the first three realisations! Each of these three sets of boundary conditions corresponds to a type of motion where particles collide elastically with the boundary. The generalised periodic boundary conditions correspond to periodic motion across the interval. In all cases, the plots of the Husimi function showed that the variation in the position coordinate fluctuates, whereas the variation in the momentum remains constant, apart from the fact that the momentum can undergo a change of sign for the realisations corresponding to the elastic collision-like behaviour. **Most importantly,**

incomplete motion or anything that hints at this phenomenon as one takes the limit $\hbar \to 0$ was not observed in any of the simulations.

It was to be expected that the generalised periodic boundary conditions generate periodic motion. The fact that Dirichlet boundary conditions can be used to simulate a particle colliding elastically with the boundary is no surprise either; these boundary conditions are typically imposed when studying quantum billiards, and can be derived heuristically by demanding that the potential $V(x)$ be infinite outside the billiard. The orthonormal basis associated to the set of mixed boundary condtions closely resembles the basis associated to Dirichlet boundary conditions, so this explains the similarity between their corresponding time evolutions. However, it is remarkable that the Neumann boundary condition appears to be physically equivalent in the limit $\hbar \to 0$ to the Dirichlet boundary conditions as well. The rest of thesis will be devoted to an idea that might shed some light on these findings.

4 Preliminaries from differential geometry

In order to outline our idea mentioned at the end of the previous section, we need to develop a good understanding of some concepts from differential geometry, such as manifolds with boundary and Riemannian manifolds, and in particular, geodesics.

4.1 Manifolds with boundary

The theory discussed here can be found in various textbooks on differential geometry, see for example [11].

4.1.1 Smooth maps and differentiable structures

First, let us recall the following notions:

4.1.1 Definition. *(smoothness of functions on open subsets of $\mathbb{R}^n$)* Let $m, n \in \mathbb{N}_0$, let $U \subseteq \mathbb{R}^n$ be an open subset, let $f\colon U \to \mathbb{R}^m$ be a function and let $x_0 \in U$.

- We say that f is C^0 *at* x_0 iff f is continuous at x_0.
- Let $k \in \mathbb{N}_0$. We say that f is $k+1$ *times continuously differentiable* or C^{k+1} *at* x_0 iff for each multi-index $\alpha \in \mathbb{N}_0^n$ of length k, there exists an open neighbourhood $U_\alpha \subseteq U$ of x_0 such that for each $x \in U_\alpha$, $f|_{U_\alpha}$ is C^k at x, and the function $\partial^\alpha f\colon U_\alpha \to \mathbb{R}^m$ given by $x \mapsto \partial^\alpha f(x)$ is continuously differentiable at x_0.
- We say that f is *smooth* or *infinitely differentiable* or C^∞ *at* $x_0 \in U$ iff there exists an open neighbourhood $U_\infty \subseteq U$ of x_0 such that for each $x \in U_\infty$ and each $k \in \mathbb{N}$, the function f is C^k at x.
- We say that f is *smooth* or *infinitely differentiable* or C^∞ iff f is smooth at every point of U.

4.1.2 Definition. *(smoothness of functions on arbitrary subsets of $\mathbb{R}^n$)* Let $k, m, n \in \mathbb{N}_0$, let $X \subseteq \mathbb{R}^n$, let $f\colon X \to \mathbb{R}^m$ be a map.

- Let $x_0 \in X$. The function f is said to be C^k *(smooth) at* x_0 iff there exists an open neighbourhood $U \subseteq \mathbb{R}^n$ of x_0 and a function $g\colon U \to \mathbb{R}^m$ that is C^k (smooth) at x_0 (in the sense of Definition 4.1.1) and extends $f|_{U\cap X}$.
- The function f is said to be C^k *(smooth)* iff f is C^k (smooth) at x for each $x \in X$.

4.1.3 Remark. Note that the previous definition of a smooth map $f\colon X \to \mathbb{R}^m$ makes sense even if the domain X is empty; in that case, f is always smooth. Moreover, if the domain X of f is open in $\mathbb{R}^n$, then the above definitions of smoothness (at a point) coincide. Finally, if $f\colon \mathbb{R}^n \supseteq X \to \mathbb{R}^m$ is smooth at a point $x \in X$, and $g\colon \mathbb{R}^m \supseteq Y \to \mathbb{R}^l$ is a function such that $f(X) \subseteq Y$ and that is smooth in $f(x)$, then $g \circ f$ is smooth at x.

4.1.4 Definition. Let $n \in \mathbb{N}$. The set $H_n := \{(x_1, \ldots, x_n) \in \mathbb{R}^n\colon x_n \geq 0\}$ with the subspace topology induced by $\mathbb{R}^n$ (with its euclidean topology) is called the *closed n-dimensional upper half-space*. The interior of H_n with respect to $\mathbb{R}^n$, denoted by H_n°, is called the *open n-dimensional upper half-space*.

Now we introduce the notion of a manifold with boundary. It is almost exactly the same as the notion of an 'ordinary' differentiable manifold, and we use the same basic terminology to describe familiar concepts such as charts, atlases, et cetera:

4.1.5 Definition. Let M be a topological space such that:

(1) M is Hausdorff;

(2) M is second countable;

(3) There exists a family of pairs $(U_\alpha, \phi_\alpha)_{\alpha \in J}$, where each pair consists of an open subset $U_\alpha \subseteq M$ and a map $\phi_\alpha \colon U_\alpha \to \mathbb{R}^n$ such that:

 (i) $\{(U_\alpha)\}_{\alpha \in J}$ is an open cover of J;

 (ii) ϕ_α is a homeomorphism onto an open subset of H_n;

 (iii) For each $\alpha, \beta \in J$, the maps $\phi_\beta \circ \phi_\alpha^{-1}|_{\phi_\alpha(U_\alpha \cap U_\beta)}$ and $\phi_\alpha \circ \phi_\beta^{-1}|_{\phi_\beta(U_\alpha \cap U_\beta)}$ are smooth.

- A pair (U_α, ϕ_α) is called a *chart*.
- The family of sets $(U_\alpha, \phi_\alpha)_{\alpha \in J}$ is called an *atlas of M*.
- Two atlases of M are said to be *compatible* if their union is an atlas of M.
- Compatibility is an equivalence relation on the collection of atlases of M; an equivalence class $\mathcal{D}_M$ with respect to this relation is called a *differentiable structure on M*; a maximal element of an equivalence class is called a *maximal atlas of M*.
- A pair $(M, \mathcal{D}_M)$ is called a *smooth n-dimensional manifold with boundary.*

Before proving that compatibility is indeed an equivalence relation whose equivalence classes contain a maximum, we introduce some terminology that is unique to manifolds with boundary:

4.1.6 Definition. Let $(M, \mathcal{D}_M)$ be a smooth n-dimensional manifold with boundary. A point $p \in M$ is called an *interior point of M* iff there exists a chart (U, ϕ) in the maximal atlas such that $p \in U$ and $\phi(p) \in H_n^\circ$. If such a chart does not exist, then p is said to be a *boundary point of M*. The collection of all interior points of M is denoted by $\mathrm{Int}(M)$, and the collection of boundary points by $\mathrm{Bound}(M)$.

4.1.7 Proposition. *Let $(M, \mathcal{D}_M)$ be an smooth n-dimensional manifold with boundary.*

(1) Compatibility of atlases of M is an equivalence relation (hence $\mathcal{D}_M$ is well defined).

(2) Every differentiable structure has a unique maximal element with respect to inclusion; thus we may identify $\mathcal{D}_M$ with that maximal element.

(3) If $p \in \mathrm{Int}(M)$, then for each chart $(U, \phi) \in D_M$ with $p \in U$, we have $\phi(p) \in H_n^\circ$.

(4) $\mathrm{Bound}(M)$ *can be endowed with the structure of an $n-1$-dimensional smooth manifold.*

(5) $\mathrm{Int}(M)$ *is open in M. Consequently,* $\mathrm{Bound}(M)$ *is closed in M.*

Proof.
(1) Clearly, compatiblity of atlases of M is a reflexive and symmetric relation. We shall prove that it is transitive. Suppose $\mathcal{A}_i = (U_{i,\alpha}, \phi_{i,\alpha})_{\alpha \in J_i}$ is an atlas of M for $i = 1, 2, 3$, suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are compatible and that $\mathcal{A}_2$ and $\mathcal{A}_3$ are compatible. Let $(U_{1,\alpha}, \phi_{1,\alpha}) \in \mathcal{A}_1$, let $(U_{3,\gamma}, \phi_{3,\gamma}) \in \mathcal{A}_3$. If $U_{1,\alpha} \cap U_{3,\gamma} = \emptyset$, then the transition functions $\phi_{3,\gamma} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{3,\gamma})}$ and $\phi_{1,\alpha} \circ \phi_{3,\gamma}^{-1}|_{\phi_{3,\gamma}(U_{1,\alpha} \cap U_{3,\gamma})}$ are clearly smooth.

Now suppose $U_{1,\alpha} \cap U_{3,\gamma} \neq \emptyset$. Let $x \in \phi_{1,\alpha}(U_{1,\alpha} \cap U_{3,\gamma})$. Then there exists a unique $p \in U_{1,\alpha} \cap U_{3,\gamma}$ such that $\phi_{1,\alpha}(p) = x$. $\mathcal{A}_2$ is an atlas of M, so there exists a chart $(U_{2,\beta}, \phi_{2,\beta}) \in \mathcal{A}_2$ such that $p \in U_{2,\beta}$. The atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are compatible, so $\phi_{2,\beta} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{2,\beta})}$ is smooth. In particular, the map

$$\phi_{2,\beta} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{2,\beta} \cap U_{3,\gamma})}$$

is smooth at x. Since $\mathcal{A}_2$ and $\mathcal{A}_3$ are compatible atlases, we can argue in the same way that the map

$$\phi_{3,\gamma} \circ \phi_{2,\beta}^{-1}|_{\phi_{2,\beta}(U_{1,\alpha} \cap U_{2,\beta} \cap U_{3,\gamma})},$$

is smooth at $\phi_{2,\beta}(p)$, so the map

$$\begin{aligned}&\phi_{3,\gamma} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{2,\beta} \cap U_{3,\gamma})} \\ &\quad = (\phi_{3,\gamma} \circ \phi_{2,\beta}^{-1}|_{\phi_{2,\beta}(U_{1,\alpha} \cap U_{2,\beta} \cap U_{3,\gamma})}) \circ (\phi_{2,\beta} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{2,\beta} \cap U_{3,\gamma})}),\end{aligned}$$

is smooth at x, hence $\phi_{3,\gamma} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{3,\gamma})}$ is smooth at x. We conclude that $\phi_{3,\gamma} \circ \phi_{1,\alpha}^{-1}|_{\phi_{1,\alpha}(U_{1,\alpha} \cap U_{3,\gamma})}$ is a smooth map. Of course, we can prove in the same way that the inverse is smooth, so $\mathcal{A}_1$ and $\mathcal{A}_3$ are compatible atlases, which proves that compatibility of atlases is a transitive relation.

(2) Let $\mathcal{A}_{\max}$ be the union of all elements of $\mathcal{D}_M$. It is clear that $\mathcal{A} \subseteq \mathcal{A}_{\max}$ for each $\mathcal{A} \in \mathcal{D}_M$. We show that the transition function corresponding to two charts (U_1, ϕ_1) and (U_2, ϕ_2) in $\mathcal{A}_{\max}$ is smooth. Indeed, by definition of $\mathcal{A}_{\max}$, there exists $\mathcal{A}_1, \mathcal{A}_2 \in \mathcal{D}_M$ such that $(U_j, \phi_j) \in \mathcal{A}_j$ for $j = 1, 2$. But then $\mathcal{A}_1 \cup \mathcal{A}_2 \in \mathcal{D}_M$ by definition of $\mathcal{D}_M$, so $\phi_2 \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2)}$ is smooth. The differentiable structure $\mathcal{D}_M$ is nonempty, so $\mathcal{A}_{\max}$ contains an atlas, hence $\mathcal{A}_{\max}$ is an atlas itself.

(3) Let $p \in \operatorname{Int}(M)$. Then there exists a chart $(U_1, \phi_1) \in D_M$ such that $p \in U_1$ and $\phi_1(p) \in H_n^\circ$. Now suppose $(U_2, \phi_2) \in \mathcal{D}_M$ is another chart such that $p \in U_2$. The map $\phi_1 \circ \phi_2^{-1}|_{\phi_2(U_1 \cap U_2)}$ is smooth, so in particular, it is smooth at $\phi_2(p)$, which implies that there exists an open neighbourhood $V \subseteq \mathbb{R}^n$ of $\phi_2(p)$ and a function $f\colon V \to \mathbb{R}^n$ that is smooth at $\phi_2(p)$ and that extends $\phi_1 \circ \phi_2^{-1}|_{\phi_2(U_1 \cap U_2) \cap V}$. Next, consider the set

$$U := (\phi_2 \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2) \cap H_n^\circ})^{-1}(V).$$

We claim that U is an open subset of $\mathbb{R}^n$. First note that the restriction of a continuous map to a subset of its domain is continuous with respect to the induced topology on that subset. Therefore, $\phi_2 \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2) \cap H_n^\circ}$ is continuous, and U is open with respect to $\phi_1(U_1 \cap U_2) \cap H_n^\circ$. Furthermore, ϕ_1 is a homeomorphism onto its open image in H_n, so $\phi_1(U_1 \cap U_2)$ is open in H_n. It is clear that H_n° is open in H_n, so $\phi_1(U_1 \cap U_2) \cap H_n^\circ$ is open in H_n, hence open in H_n°, hence $\phi_1(U_1 \cap U_2) \cap H_n^\circ$ is open in $\mathbb{R}^n$ and it follows that U is an open subset of $\mathbb{R}^n$, which proves the claim.

Furthermore, note that U contains $\phi_1(p)$, and that the composition

$$f \circ \phi_2 \circ \phi_1^{-1}|_U \colon U \to \mathbb{R}^n,$$

is the identity map onto its image. Thus the derivative of the composition in $\phi_1(p)$ is the identity map, so by the chain rule, the derivative of $\phi_2 \circ \phi_1^{-1}|_{\phi_1(U_1 \cap U_2)}$ at $\phi_1(p)$ is invertible. The inverse function theorem now implies that there exists an open neighbourhood $W \subseteq \mathbb{R}^n$ of $\phi_1(p)$ that is mapped homeomorphically onto its open image $\phi_2 \circ \phi_1^{-1}(W) \subseteq \mathbb{R}^n$. On the other hand, we know that $\phi_2(p) \in \phi_2 \circ \phi_1^{-1}(W) \subseteq H_n$, so $\phi_2(p) \in H_n^\circ$, as desired.

(4) M is Hausdorff and second countable, so $\mathrm{Bound}(M)$ with the induced topology is also Hausdorff and second countable.

Construct an atlas for $\mathrm{Bound}(M)$ as follows. Let $(U, \phi) \in D_M$ be a chart such that U has nonempty intersection with $\mathrm{Bound}(M)$. It follows from the previous part of the proposition that $\phi(U \cap \mathrm{Bound}(M)) = \phi(U) \cap (\mathbb{R}^{n-1} \times \{0\})$. ϕ is a homeomorphism onto its open image in H_n, so $\phi|_{U \cap \mathrm{Bound}(M)}$ is a homeomorphism onto an open subset of $\mathbb{R}^{n-1} \times \{0\}$. The subspace $\mathbb{R}^{n-1} \times \{0\}$ is of course homeomorphic to $\mathbb{R}^{n-1}$, so we can use the above restriction to define a homeomorphism $\psi \colon V \to \mathbb{R}^{n-1}$, where $V := U \cap \mathrm{Bound}(M)$.

We can perform this construction for each chart $(U, \phi) \in D_M$ having non empty intersection with the boundary of M to obtain a family $\{(V_\alpha, \psi_\alpha)\}_{\alpha \in I}$ of charts on $\mathrm{Bound}(M)$. One readily checks that this family has all properties of an atlas. Thus $\mathrm{Bound}(M)$ with the differentiable structure $\mathrm{Bound}(M)$ associated to this atlas is an $n-1$-dimensional smooth manifold.

(5) Let $p \in \mathrm{Int}(M)$. Then there exists a chart $(U, \phi) \in \mathcal{D}_M$ such that $\phi(p) \in H_n^\circ$. But then $\phi^{-1}(H_n^\circ)$ is an open subset of U and a subset of $\mathrm{Int}(M)$ that contains p. The set U is open in M, so $\phi^{-1}(H_n^\circ)$ is an open neighbourhood of p in M that is contained in $\mathrm{Int}(M)$. We conclude that $\mathrm{Int}(M)$ is open in M. ■

4.1.8 Example.

(1) Every smooth manifold M (with a differentiable structure) is automatically a smooth manifold with boundary, and $\mathrm{Bound}(M) = \emptyset$. Conversely, if M is a smooth manifold with boundary such that $\mathrm{Bound}(M) = \emptyset$, then by part (3) of Proposition 4.1.7, M is a smooth manifold.

(2) For each $n \in \mathbb{N}$, the half-space H_n itself can naturally be endowed with the structure of an n-dimensional smooth manifold with boundary with $\mathrm{Int}(H_n) = H_n^\circ$ and $\mathrm{Bound}(H_n) = \partial H_n$.

(3) For each $n \in \mathbb{N}$, the closed unit ball in $\mathbb{R}^n$ can be given the structure of an n-dimensional smooth manifold with boundary such that the interior of the manifold is the open unit ball in $\mathbb{R}^n$ and the boundary is S^{n-1}. In particular, the closed interval $[-1, 1]$ has boundary $\{-1, 1\}$.

(4) Every open subset of a smooth manifold with boundary is again a smooth manifold with boundary of the same dimension.

4.1.2 The tangent space and smooth maps

Now that we have defined the notion of a smooth manifold with boundary, we would like to define some of the concepts which are already defined in the setting of ordinary smooth manifolds, such as the tangent space and the tangent map of a smooth map.

First, we need to extend the notion of a derivative at a point of a function defined on an open subset of $\mathbb{R}^n$ to functions defined on open subsets of H_n. It is not obvious that the concept of a derivative makes sense for an arbitrary smooth map defined on an open subset of a half-space; it certainly does not make sense for functions defined on an arbitrary subset of $\mathbb{R}^n$. As a counterexample, one can think of a constant function $f\colon \{0\} \to \mathbb{R}$, which can be extended smoothly to $\mathbb{R}$ in infinitely many ways, and few of these extensions will have the same derivative at 0. However, in the case of open subsets of half-spaces, we have the following lemma:

4.1.9 Lemma. *Let $k, m, n \in \mathbb{N}$, let $U_0 \subseteq H_n$ be an open subset (relative to H_n), let $f\colon U_0 \to \mathbb{R}^m$ be a map and let $x_0 \in U_0$. Suppose that U_1 and U_2 are two open neighbourhoods of x_0 in $\mathbb{R}^n$, that $g_1\colon U_1 \to \mathbb{R}^m$ and $g_2\colon U_2 \to \mathbb{R}^m$ are two maps that are both C^k at x_0, and that $g_j|_{U_0 \cap U_i} = f|_{U_0 \cap U_i}$ for $j = 1, 2$. Then for each multi-index $\alpha \in \mathbb{N}_0^n$ of length k, we have $\partial^\alpha g_1(x_0) = \partial^\alpha g_2(x_0)$.*

Proof. We prove the lemma by induction on the length k of the multi-index α. For $\alpha = 0$, there is nothing to prove. Now suppose that the assertion is true for all multi-indices β of length $k \in \mathbb{N}$, and that the functions g_1 and g_2 are C^{k+1} at x_0. Fix a multi-index $\alpha \in \mathbb{N}_0^n$ of length $k+1$. Then there exists a multi-index $\beta \in \mathbb{N}_0^n$ of length k and an i with $1 \leq i \leq n$ such that $\alpha - \beta = e_i$, where e_i is the i-th standard basis vector of $\mathbb{R}^n$. Let $U_{\beta,1} \subseteq U_1$ and $U_{\beta,2} \subseteq U_2$ be open subsets such that the functions $g_i|_{U_{\beta,i}}$ are C^k at x for each $x \in U_{\beta,i}$ for $i = 1, 2$, and let $U := U_0 \cap U_{\beta,1} \cap U_{\beta,2}$. In view of $g_1|_U = f|_U = g_2|_U$ and the induction hypothesis, we have

$$\partial^\beta g_1(x) = \partial^\beta g_2(x) \quad \text{for each } x \in U.$$

Now note that U is an open subset of H_n, and that g_1 and g_2 are C^{k+1} at x_0. Hence we have:

$$\partial^\alpha g_j(x_0) = \partial_i \partial^\beta g_j(x_0) = \lim_{t>0, x+te_i \in U} \frac{\partial^\beta g_j(x + te_i) - \partial^\beta g_j(x_0)}{t},$$

for $j = 1, 2$. Comparing this equation with the previous one, we obtain

$$\partial^\alpha g_1(x_0) = \partial^\alpha g_2(x_0).$$

This completes the induction. ■

4.1.10 Definition. Let $k, m, n \in \mathbb{N}$, let $U \subseteq H_n$ be an open subset of H_n, let $f\colon U \to \mathbb{R}^m$ and suppose that f is C^k (smooth) at x_0. Then there exists an open neighbourhood $V \subseteq \mathbb{R}^n$ of x_0 and a C^k (smooth) function $g\colon V \to \mathbb{R}^m$ such that $f|_{U \cap V} = g|_{U \cap V}$. For each $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \leq k$, let $\partial^\alpha f(x_0) := \partial^\alpha g(x_0)$. Moreover, we define the *total derivative $f'(x_0)$ of f at x_0* by $f'(x_0) := g'(x_0)\colon \mathbb{R}^n \to \mathbb{R}^m$.

The previous lemma says that $\partial^\alpha g(x_0)$ in the above definition is independent of the choice of g, hence $\partial^\alpha f(x_0)$ is well defined. Furthermore, since the derivative of a C^1 function at a point on an open subset of $\mathbb{R}^n$ can be expressed entirely in terms of its first order partial derivatives, the linear map $f'(x_0)$ is well defined. One can recover the chain rule for functions between open subsets of half-spaces from the chain rule for functions defined on open subsets of $\mathbb{R}^n$.

We are now ready to construct the tangent space. We will follow the method outlined in [9, 1.27]. Suppose $(M, \mathcal{D}_M)$ is a n-dimensional smooth manifold with boundary, and let $p \in M$. Again, identify $\mathcal{D}_M$ with its maximal atlas, let

$$A_p := \{((U,\phi),u) \in \mathcal{D}_M \times \mathbb{R}^n \colon p \in U\}.$$

Define the relation $\sim$ on A_p as follows:

$$((U,\phi),u) \sim ((V,\psi),v) \Leftrightarrow (\psi \circ \phi^{-1}|_{\phi(U\cap V)})'(\phi(p))(u) = v.$$

4.1.11 Lemma.

(1) The relation $\sim$ is an equivalence relation on A_p.

(2) If $((U,\phi),u_1), ((U,\phi),u_2) \in A_p$ are elements of the same equivalence class, then $u_1 = u_2$.

(3) Let $((U,\phi),u) \in A_p$. For each chart $(V,\psi) \in \mathcal{D}_M$, there exists a $v \in \mathbb{R}^n$ such that $((U,\phi),u) \sim ((V,\psi),v)$.

Proof. Let $((U,\phi),u) \in A_p$. The identity map on an open subset of H_n can be extended smoothly to the identity map on $\mathbb{R}^n$, so $(\phi \circ \phi^{-1}|_{\phi(U)})'(\phi(p))(u) = \mathrm{Id}|'_{\phi(U)}(\phi(p))(u)$, which proves that $\sim$ is reflexive.

Now suppose that $((U,\phi),u), ((V,\psi),v), ((W,\chi),w) \in A_p$, that $((U,\phi),u) \sim ((V,\psi),v)$ and that $((V,\psi),v) \sim ((W,\psi),w)$. By the chain rule, we have

$$\begin{aligned}
(\chi \circ \phi^{-1})|'_{\phi(U\cap W)}(\phi(p))(u) &= (\chi \circ \psi^{-1} \circ \psi \circ \phi^{-1}|_{\phi(U\cap V\cap W)})'(\phi(p))(u) \\
&= (\chi \circ \psi^{-1}|_{\psi(V\cap W)})'(\psi(p)) \circ (\psi \circ \phi^{-1}|_{\phi(U\cap V)})'(\phi(p))(u) \\
&= (\chi \circ \psi^{-1}|_{\psi(V\cap W)})'(\psi(p))(v) \\
&= w,
\end{aligned}$$

so $\sim$ is transitive.

For the symmetry of $\sim$, note that

$$\begin{aligned}
(\phi \circ \psi^{-1}|_{\psi(U\cap V)})'(\psi(p))(v) &= (\phi \circ \psi^{-1}|_{\psi(U\cap V)})'(\psi(p)) \circ D(\psi \circ \phi^{-1}|_{\phi(U\cap V)})(\phi(p))(u) \\
&= (\mathrm{Id}|_{\phi(U\cap V)})'(\phi(p))(u) \\
&= u.
\end{aligned}$$

We conclude that $\sim$ is an equivalence relation, which proves (1).

To prove (2), simply note that $u_1 = \mathrm{Id}|'_{\phi(U)}(\phi(p))(u_1) = u_2$.

For (3), we remark that $v := (\psi \circ \phi^{-1}|_{\phi(U\cap V)})'(\phi(p))(u)$ satisfies $((U,\phi),u) \sim ((V,\psi),v)$. ∎

Now suppose $(U,\phi) \in \mathcal{D}_M$ is a chart such that $p \in U$. Then the map $\theta_{(U,\phi)} \colon \mathbb{R}^n \to A_p/\sim$ given by $u \mapsto ((U,\phi),u)$ is injective by part (2) of the previous lemma, and surjective by part (3). Thus $\theta_{(U,\phi)}$ is a bijection, and we may use this bijection to transfer the vector space structure of $\mathbb{R}^n$ to $A_p/\sim$.

If (V,ψ) is another chart such that $p \in V$, then we can define $\theta_{(V,\psi)}$ in the same way we defined $\theta_{(U,\phi)}$, and it is easy to see that

$$\theta_{(U,\phi)} = \theta_{(V,\psi)} \circ (\psi \circ \phi^{-1}|_{\phi(U\cap V)})'(\phi(p)).$$

The map $(\psi \circ \phi^{-1}|_{\phi(U\cap V)})'(\phi(p))$ is a linear automorphism of $\mathbb{R}^n$, so the vector space structure on $A_p/\sim$ is independent of the choice of (U,ϕ).

4.1.12 Definition. The set $A_p/\sim$ with the above vector space structure is called the *tangent space of* $(M,\mathcal{D}_M)$ *at* p and denoted by by T_pM.

4.1.13 Definition. Let $(N,\mathcal{D}_N)$ be an n-dimensional smooth manifold with boundary, let $(M,\mathcal{D}_M)$ be an m-dimensional smooth manifold with boundary, let $f \colon N \to M$ be a map, and let $p \in N$.

- We say that f *is* C^k *(smooth) at* p iff there exist charts $(U,\phi) \in \mathcal{D}_N$ and $(V,\psi) \in \mathcal{D}_M$ such that $p \in U$, $f(U) \subseteq V$, and $\psi \circ f \circ \phi^{-1}|_{\phi(U)}$ is C^k (smooth) at $\phi(p)$ in the sense of Definition 4.1.2;
- If f is C^1 at p, then we define the *tangent map* $T_pf \colon T_pN \to T_{f(p)}M$ *of* f *at* p by

$$T_pf := \theta_{(V,\psi)} \circ (\psi \circ f \circ \phi^{-1})'(\phi(p)) \circ \theta^{-1}_{(U,\phi)};$$

- We say that f *is* C^k *(smooth)* iff f is C^k (smooth) at each point $p \in N$;
- If f is a C^k (smooth) bijection with a C^k (smooth) inverse f, then we call f a C^k *(smooth) diffeomorphism.*

4.1.14 Lemma. *The above definition of the tangent map* T_pf *does not depend on the choices of the charts* (U,ϕ) *and* (V,ψ).

Proof. Suppose (U_1,ϕ_1) and (U_2,ϕ_2) are charts on N whose domains contain p and suppose (V_1,ψ_1) and (V_2,ψ_2) are two charts on M whose domains contain $f(p)$, and suppose that $f(U_i) \subseteq V_i$ for $i = 1,2$. Let T_pf_1 and T_pf_2 be the associated tangent maps. We are done if we can show that the following cube is commutative:

$$
\begin{tikzcd}
 & T_pN \arrow[rr, "T_pf_2"] & & T_{f(p)}M \\
T_pN \arrow[ur, equal] \arrow[rr, "T_pf_1"] & & T_{f(p)}M \arrow[ur, equal] & \\
 & \mathbb{R}^n \arrow[uu, dotted, "\theta_{(U_2,\phi_2)}"'] \arrow[rr, dotted, "(\psi_2 \circ f \circ \phi_2^{-1})'(\phi_2(p))"'] & & \mathbb{R}^m \arrow[uu, "\theta_{(V_2,\psi_2)}"'] \\
\mathbb{R}^n \arrow[uu, "\theta_{(U_1,\phi_1)}"] \arrow[ur, dotted, "(\phi_2 \circ \phi_1^{-1})'(\phi_1(p))"'] \arrow[rr, "(\psi_1 \circ f \circ \phi_1^{-1})'(\phi_1(p))"'] & & \mathbb{R}^m \arrow[uu, "\theta_{(V_1,\psi_1)}"'] \arrow[ur, "(\psi_2 \circ \psi_1^{-1})'(\psi_1 \circ f(p))"'] &
\end{tikzcd}
$$

Indeed, the front and the back sides of the above cube are commutative by definition of T_pf_1 and T_pf_2, respectively. Also, from our discussion on the vector space structure on the tangent spaces, we infer that the left-hand and right-hand sides of the cube are commutative. Finally, it follows from the chain rule that the bottom of the cube is commutative. Since all arrows in this diagram, except perhaps the ones pointing to the right, are isomorphisms of vector spaces, we conclude that the entire cube is commutative. In particular, the top square commutes, which implies that $T_pf_1 = T_pf_2$, as desired. ■

4.1.3 Products and fibre bundles

Now that we have constructed the tangent space, we want to talk about vector bundles over a manifold with boundary, in particular about the cotangent bundle, since we are interested in the phase space of a manifold with boundary. In order to construct vector bundles, it is convenient to have the notion of a product of two manifolds at hand. There is one problem: although it is rather straightforward to define the notion of a product in the category of smooth manifolds, it is not so clear what a product of two manifolds with boundary is. For example, the product of $H_1 = [0, \infty)$ with itself has a 'corner' at $(0, 0)$. Though it is possible to 'straighten out' the boundary of $H_1 \times H_1$ (viewed as a subset of $\mathbb{R}^2$), one loses the nice differentiable structure inherited from $\mathbb{R}^2$ in doing so.

There are two ways to deal with this issue: first, one can consider the larger category of *manifolds with corners*, where there is a well-defined notion of a product; see [13] for a discussion on these objects. The second method, which we will discuss here, is to only consider the product of a manifold with empty boundary, an 'ordinary' manifold, with a manifold with boundary. This will turn out to be sufficient for our purposes.

4.1.15 Proposition. *Let $(F, \mathcal{D}_F)$ be a smooth p-dimensional manifold and let $(M, \mathcal{D}_M)$ be an m-dimensional manifold with boundary. Then there exists a unique topology and a unique differentiable structure $\mathcal{D}$ on $F \times M$ such that $(F \times M, D)$ is a smooth $p+m$-dimensional manifold with boundary* $\mathrm{Bound}(F \times M) = F \times \mathrm{Bound}(M)$*, the projection maps $P_F \colon F \times M \to F$ and $P_M \colon F \times M \to M$ are smooth, and $(F \times M, \mathcal{D})$ satisfies the following universal property:*
Let N be a smooth manifold with boundary, and let $f \colon N \to F$ and $g \colon N \to M$ be

smooth maps. Then there exists a unique smooth map h that makes the following diagram commutative:

$$\begin{tikzcd}
N \arrow[dr, dashed, "h"] \arrow[drr, bend left, "f"] \arrow[ddr, bend right, "g"'] & & \\
& F \times M \arrow[r, "P_F"] \arrow[d, "P_M"'] & F \\
& M &
\end{tikzcd}$$

Proof. For the topology on $F \times M$, take the product topology. Then $F \times M$ is Hausdorff and second countable because F and M are. For each chart $(U, \phi) \in \mathcal{D}_F$ and each chart $(V, \psi) \in \mathcal{D}_M$, define a map $\phi \times \psi \colon U \times V \to \mathbb{R}^p \times H_m = H_{p+m}$ by $(x, y) \mapsto (\phi(x), \psi(y))$. $U \times V$ is an open subset of $F \times M$ and $\phi \times \psi$ is a homeomorphism onto its image, that is open in H_{p+m}, because ϕ and ψ are homeomorphisms onto their images, which are open in $\mathbb{R}^p$ and H_m respectively. One readily verifies that

$$\mathcal{A} := \{(U \times V, \phi \times \psi) \colon (U, \phi) \in \mathcal{D}_F,\ (V, \psi) \in \mathcal{D}_M\},$$

is an atlas of $F \times M$, and that P_F and P_M are smooth maps with respect to the differentiable structure $\mathcal{D}$ on $F \times M$ corresponding to $\mathcal{A}$.

Now suppose $(x_0, y_0) \in F \times M$ and suppose that $(U \times V, \phi \times \psi) \in \mathcal{A}$ is a chart such that $(x_0, y_0) \in U \times V$. Then by part (3) of Proposition 4.1.7 the following are equivalent:

- $(x_0, y_0) \in \operatorname{Int}(F \times M)$;
- $(\phi(x_0), \psi(y_0)) \in H^\circ_{p+m}$;
- $\psi(y_0) \in H^\circ_m$;
- $y_0 \in \operatorname{Int}(M)$.

Thus $\operatorname{Int}(F \times M) = F \times \operatorname{Int}(M)$ and $\operatorname{Bound}(F \times M) = F \times \operatorname{Bound}(M)$.

Next, let N be an n-dimensional manifold with boundary and let $f \colon N \to F$ and $g \colon N \to M$ be smooth maps. Then the map $h := f \oplus g \colon N \to F \times M$ given by $p \mapsto (f(p), g(p))$ is the unique map that makes the diagram above commutative. It is continuous by the universal property for products of topological spaces. To see that it is smooth, we prove that h is smooth at each point $p \in N$. By the smoothness of f and g, there exists a chart (W, χ) of N such that $p \in N$ and there exist $(U, \phi) \in \mathcal{D}_F$ and $(V, \psi) \in \mathcal{D}_M$ such that $f(W) \subseteq U$ and $g(W) \subseteq V$, and $\phi \circ f \circ \chi^{-1}$ and $\psi \circ g \circ \chi^{-1}$ are smooth. But then $h(W) \subseteq U \times V$, and

$$(\phi \times \psi) \circ h \circ \chi^{-1} = (\phi \circ f \circ \chi^{-1}) \oplus (\psi \circ g \circ \chi^{-1}),$$

is smooth. Thus h is smooth, and $(F \times M, \mathcal{D})$ satisfies the universal property. Uniqueness of the topology and the differentiable structure can now be deduced from the universal property by taking $N = F \times M$, $f = P_F$ and $g = P_M$. ■

Now we can introduce the notion of a fibre bundles over a manifold with boundary, where the fibre is a manifold with empty boundary:

4.1.16 Definition. Let B and E be manifolds with boundary, let F be a manifold and, let $\pi\colon E \to B$ be a smooth surjection. Then (E, B, F, π) is called a *smooth fibre bundle over B with fibre F* iff for each $p \in B$, there exists an open neighbourhood $U \subseteq B$ of p and a diffeomorphism $\Phi\colon \pi^{-1}(U) \to F \times U$ such that $\pi|_{\pi^{-1}(U)} = P_U \circ \Phi$, where $P_U\colon F \times U \to U$ is the canonical projection on U.
A pair (U, Φ) is called a *local trivialisation of E*.

The following proposition provides us with a way of constructing a fibre bundle.

4.1.17 Proposition. *Let $(B, \mathcal{D}_B)$ be an m-dimensional manifold with boundary, let $(F, \mathcal{D}_F)$ be a p-dimensional manifold, and let $(F_p)_{p\in B}$ be a family of sets. Let $E := \coprod_{p\in B} F_p$ be the disjoint union of these sets, and let $\pi\colon E \to B$ be the canonical projection onto B.*
In addition, let $(U_\alpha, \Phi_\alpha)_{\alpha\in I}$ be a family of pairs, each of which consists of an open subset $U_\alpha \subseteq B$ and a bijection $\Phi_\alpha\colon \pi^{-1}(U_\alpha) \to F \times U_\alpha$ with the following properties:

(i) *$(U_\alpha)_{\alpha\in I}$ is an open cover of B.*

(ii) *For each $\alpha \in I$, let $P_{U_\alpha}\colon F \times U_\alpha \to U_\alpha$ be the canonical projection on U_α. Then $\pi|_{\pi^{-1}(U_\alpha)} = P_{U_\alpha} \circ \Phi_\alpha$.*

(iii) *For each $\alpha, \beta \in I$, the map $\Phi_\beta \circ \Phi_\alpha^{-1}|_{F\times(U_\alpha\cap U_\beta)}$ is smooth.*

Then there exists a topology τ and a differentiable structure $\mathcal{D}$ on E that turns E into a $p+m$-dimensional manifold with boundary $\pi^{-1}(\mathrm{Bound}(B))$ and that turns (E, B, F, π) into a fibre bundle. Furthermore, τ and $\mathcal{D}$ are the only topology and differentiable structure for which $\pi^{-1}(U_\alpha)$ is open in E and Φ_α is a diffeomorphism for each $\alpha \in I$.

Proof. Define the topology τ as follows:
Let τ be the collection of all subsets $X \subseteq E$ such that for each $x \in X$, there exists an $\alpha \in I$ with $\pi(x) \in U_\alpha$ and $\Phi_\alpha(X \cap \pi^{-1}(U_\alpha))$ is open in $F \times U_\alpha$.
From this definition, it is easily seen that $\pi^{-1}(U_\alpha) \in \tau$ for each $\alpha \in I$.

Before we prove that τ is a topology, we prove the following claim:

$X \in \tau$ if and only if for each $\alpha \in I$, $\Phi_\alpha(X \cap \pi^{-1}(U_\alpha))$ is open in $F \times U_\alpha$.

Proof of the claim: the 'if'-part is a direct consequence of the definition and the fact that $(U_\alpha)_{\alpha\in I}$ covers B. To prove the converse, let $X \in \tau$, let $\alpha \in I$ and let $y \in \Phi_\alpha(X \cap \pi^{-1}(U_\alpha))$. Then there exists a $\beta \in I$ such that $\pi \circ \Phi_\alpha^{-1}(y) \in U_\beta$ and such that $\Phi_\beta(X \cap \pi^{-1}(U_\beta))$ is open in $F \times U_\beta$. Thus the set $\Phi_\beta(X \cap \pi^{-1}(U_\beta))$ is an open subset of $F \times U_\beta$ containing $\Phi_\beta \circ \Phi_\alpha^{-1}(y)$. It follows from this and from property (ii) of the family of pairs $(U_\alpha, \Phi_\alpha)_{\alpha\in I}$ that the set $W := \Phi_\beta(X \cap \pi^{-1}(U_\beta)) \cap (F \times U_\alpha)$ is an open subset of $F \times (U_\alpha \cap U_\beta)$ containing $\Phi_\beta \circ \Phi_\alpha^{-1}(y)$. By property (iii), $\Phi_\alpha \circ \Phi_\beta^{-1}(W)$ is an open subset of $F \times (U_\alpha \cap U_\beta)$ containing y, and hence it is an open subset of $F \times U_\alpha$ containing y. Moreover, $\Phi_\alpha \circ \Phi_\beta^{-1}(W)$ is contained in $\Phi_\alpha(X \cap \pi^{-1}(U_\alpha))$, again by property (ii). We conclude that $\Phi_\alpha(X \cap \pi^{-1}(U_\alpha))$ is open in $F \times U_\alpha$, which proves the claim.

One readily verifies from the definition of τ that $\emptyset, E \in \tau$. Now let $(X_\beta)_{\beta\in J}$ be a family of subsets of E. Then for each $\alpha \in I$, we have

$$\begin{aligned}\Phi_\alpha\left(\left(\bigcup_{\beta\in J} X_\beta\right)\cap \pi^{-1}(U_\alpha)\right) &= \Phi_\alpha\left(\bigcup_{\beta\in J}(X_\beta\cap\pi^{-1}(U_\alpha))\right)\\ &= \bigcup_{\beta\in J}\Phi_\alpha(X_\beta\cap\pi^{-1}(U_\alpha)),\end{aligned}$$

and a similar identity holds when we replace the unions with intersections, since Φ_α is a bijection. But then the claim implies that τ is closed under arbitrary unions and finite intersections, so τ is indeed a topology on E.

Next, let $\alpha \in I$. We show that Φ_α is a homeomorphism. Let $V \subseteq F \times U_\alpha$ be open. Let $x \in \Phi_\alpha^{-1}(V)$. Then $\pi(x) \in U_\alpha$, and $\Phi_\alpha(\Phi_\alpha^{-1}(V)\cap\pi^{-1}(U_\alpha)) = V$ is open in $F\times U_\alpha$, so $\Phi_\alpha^{-1}(V) \in \tau$ by definition of τ. Thus Φ_α is continuous.

Now suppose U is an open subset of $\pi^{-1}(U_\alpha)$. Then U is open in E, because $\pi^{-1}(U_\alpha)$ is open in E. It follows from the claim we proved earlier that $\Phi_\alpha(U\cap\pi^{-1}(U_\alpha)) = \Phi_\alpha(U)$ is open in $F\times U_\alpha$. We conclude that the bijection Φ_α is open, and therefore a homeomorphism.

As in the case of product manifolds, (E,τ) is Hausdorff and second countable because both B and F are Hausdorff and second countable.

To construct a differentiable structure on E, first equip all of the sets $F\times U_\alpha$ with the differentiable structure $\mathcal{D}_{F\times U_\alpha}$ of a product manifold as described in Proposition 4.1.15, then define a collection $\mathcal{A}$ of pairs by

$$\mathcal{A} = \{(\Phi_\alpha^{-1}(W), \chi\circ\Phi_\alpha|_{\Phi_\alpha^{-1}(W)})\colon (W,\chi)\in\mathcal{D}_{F\times U_\alpha},\ \alpha\in I\}.$$

Then $\mathcal{A}$ is an atlas of E. In particular, property (iii) implies that all of the transition functions are smooth, and that the functions $(\Phi_\alpha)_{\alpha\in I}$ are diffeomorphisms. Let $\mathcal{D}$ be the associated differentiable structure on E. Furthermore, the following are equivalent:

- $x \in \mathrm{Int}(E)$;
- There exist $\alpha\in I$ and $(W,\chi)\in\mathcal{D}_{F\times U_\alpha}$ such that $x\in\Phi_\alpha^{-1}(W)$ and $\chi\circ\Phi_\alpha(x)\in H^\circ_{p+m}$;
- There exists $\alpha\in I$ such that $\pi(x)\in U_\alpha$ and $\Phi_\alpha(x)\in\mathrm{Int}(F\times U_\alpha) = F\times\mathrm{Int}(U_\alpha)$;
- There exists $\alpha\in I$ such that $x\in\pi^{-1}(\mathrm{Int}(U_\alpha))$;
- $x\in\pi^{-1}(\mathrm{Int}(B))$.

Thus $\mathrm{Int}(E) = \pi^{-1}(\mathrm{Int}(B))$, and $\mathrm{Bound}(E) = \pi^{-1}(\mathrm{Bound}(B))$. It is clear that (E,B,F,π) is a fibre bundle, with local trivialisations $(U_\alpha,\Phi_\alpha)_{\alpha\in I}$.

To prove the statements about uniqueness, we note that the following are equivalent:

- X is open in E;
- For each $\alpha\in I$, $X\cap\pi^{-1}(U_\alpha)$ is open in $\pi^{-1}(U_\alpha)$;
- For each $\alpha\in I$, $\Phi_\alpha(X\cap\pi^{-1}(U_\alpha))$ is open in $F\times U_\alpha$.

We conclude that the topology τ is the unique topology such that for each $\alpha\in I$, $\pi^{-1}(U_\alpha)$ is open in E and Φ_α is a homeomorphism. Now suppose $\mathcal{D}'$ is a differentiable structure on E such that Φ_α is a diffeomorphism for each $\alpha\in I$. Then the atlas $\mathcal{A}$ is compatible with all atlases in $\mathcal{D}'$, hence $\mathcal{D}' = \mathcal{D}$ and the differentiable structure is unique. ■

4.1.18 Definition. Let (E, B, F, π) be a fibre bundle over B, and let $(U_\alpha, \Phi_\alpha)_{\alpha \in I}$ be a collection of local trivialisations such that $(U_\alpha)_{\alpha \in i}$ is an open cover of B. Suppose that F is a vector space, that for each $p \in B$, the set $\pi^{-1}(\{p\})$ is a vector space, and that for each $\alpha \in I$ such that $p \in U_\alpha$, the map $\pi^{-1}(\{p\}) \to F$, given by $x \mapsto P_F \circ \Phi_\alpha(x)$, is linear (and hence an isomorphism of vector spaces), where $P_F \colon F \times B \to F$ is the canonical projection. Then (E, B, F, π) is called a *vector bundle over B*.

4.1.19 Example.

(1) Any product $E = F \times B$ of a manifold F with a manifold with boundary B can be viewed as a fibre bundle with fibre F, and the map π is the canonical projection P_B onto B. A bundle of this form is called *trivial*.

(2) The tangent bundle $TM := \coprod_{p \in M} T_pM$ of an n-dimensional manifold with boundary M can be given the structure of a vector bundle over M. Let $\pi \colon TM \to M$ be the canonical projection. For each chart (U, ϕ) of M, consider the bijection $\Psi \colon \mathbb{R}^n \times U \to \pi^{-1}(U)$ given by $(u, p) \mapsto [(U, \phi), u] \in T_pM$. The inverses of these maps can be used to endow TM with the structure of a vector bundle using Proposition 4.1.17.

(3) Similar to the previous example, one can consider the cotangent bundle $T^*M := \coprod_{p \in M} T_p^*M$ of an n-dimensional manifold with boundary M, where for each $p \in M$, T_p^*M is the dual space of T_pM. For each chart (U, ϕ) of M, let $\Psi \colon \mathbb{R}^n \times U \to \pi^{-1}(U)$ be the map that maps for fixed $p \in U$ the standard basis $(e_1, \dots, e_n)$ of $\mathbb{R}^n$ to the dual basis of the basis $([(U, \phi), e_1], \dots, [(U, \phi), e_n])$ of T_pM, which we denote by $([(U, \phi), e_1]^*, \dots, [(U, \phi), e_n]^*)$.

(4) Finally, we define the bundle of mixed tensors of type (k, l) $(k, l \in \mathbb{N}_0)$, with

$$T_l^k M := \coprod_{p \in M} (\otimes^k T_p^* M) \otimes (\otimes^l T_p M),$$

and given a chart (U, ϕ) of M, the local frame $([(U, \phi), e_1], \dots, [(U, \phi), e_n])$ induces a local frame of $T_l^k M$, which can be used to define a local trivialisation

$$\Phi \colon \pi^{-1}(U) \to \mathbb{R}^{n^{(k+l)}} \times U,$$

where $\pi \colon T_l^k M \to M$ is the canonical projection. Composing Φ with the map $\mathrm{Id}_{\mathbb{R}^{n^{(k+l)}}} \times \phi$ yields a chart of $T_l^k M$. In what follows, we shall refer to this chart as the chart associated to (U, ϕ).

4.2 Symplectic geometry and Hamilton's equations

The purpose of this subsection is to recall some of the abstract framework behind Hamilton's equations, and to introduce Hamiltonian vector fields.

4.2.1 Definition. Let $(Q, \mathcal{D}_Q)$ be a smooth manifold with boundary. A *Riemannian metric g* on Q is a smooth section of the bundle $T_0^2 Q$ with the additional property that for each $q \in Q$, the element g_q, viewed as a bilinear form on T_qQ, is an inner product. Such a triple $(Q, \mathcal{D}_Q, g)$ is called a *Riemannian manifold with boundary*

Moreover, let us introduce the notion of a vector field:

4.2.2 Definition. Let $(Q, \mathcal{D}_Q)$ be a smooth manifold with boundary. A smooth section of the canonical projection $TQ \to Q$ is called a *vector field on* Q. The set of all vector fields on Q is denoted by $\mathcal{X}(Q)$.

4.2.3 Remark. The set $\mathcal{X}(Q)$, together with pointwise addition and scalar multiplication, is a vector space. Moreover, the map $C^\infty(Q) \times \mathcal{X}(Q) \to \mathcal{X}(Q)$, given by $(f, X) \mapsto f \cdot X$, where $(f \cdot X)(q) := f(q)X(q)$ for each $q \in Q$, defines a $C^\infty(Q)$-module structure on $\mathcal{X}(Q)$.

4.2.4 Definition. Let V be a real vector space. A *symplectic form* ω on V is a nondegenerate bilinear form on V that is skew-symmetric, i.e.

$$\omega(u, v) = -\omega(v, u) \quad \text{for each } u, v \in V.$$

The pair (V, ω) is called a *symplectic vector space.*

4.2.5 Proposition. *Let $(Q, \mathcal{D}_Q)$ be a smooth, n-dimensional manifold with boundary. Define the 1-form $\alpha \in T^*(T^*Q)$ as follows: Let (U, ϕ) be a chart of Q introducing local coordinates $(q_1, \ldots, q_n)$, and let $(\pi^{-1}(U), \Phi)$ be the induced chart on T^*Q with corresponding local coordinates $(p_1, \ldots, p_n, q_1, \ldots, q_n)$, where $\pi \colon T^*Q \to Q$ is the canonical projection. Now define α on $\pi^{-1}(U)$ by*

$$\alpha := \sum_{j=1}^{n} p_j dq_j.$$

Then:

(1) The definition of α is independent of the choice of the chart (U, ϕ);

(2) The exterior derivative $\omega := d\alpha$ is given in local coordinates by

$$\omega = \sum_{j=1}^{n} dp_j \wedge dq_j.$$

This definition is again independent of the choice of the chart (U, ϕ). Furthermore, for each $(p_0, q_0) \in T^*Q$, $\omega_{(p_0,q_0)}$ is a symplectic form on $T_{(p_0,q_0)}(T^*Q)$, and $d\omega = 0$.

Proof. See [3, Section 2.2]. ■

4.2.6 Definition. Let $(Q, \mathcal{D}_Q)$ be a smooth, n-dimensional manifold with boundary. The 2-form ω obtained in the previous proposition is called the *canonical symplectic form on* T^*Q.

Next, we define the classical Hamiltonian. Suppose $(Q, \mathcal{D}_Q, g)$ is a Riemannian manifold with boundary. Then for each $q \in Q$, the inner product g_q defines an isomorphism $T_qQ \to T_q^*Q$, by mapping a vector v to the linear functional $w \mapsto g_q(v, w)$. This functional is commonly written as $v^\flat$, and the inverse of the above isomorphism is denoted by $\theta \mapsto \theta^\sharp$.

Now suppose that $(Q, \mathcal{D}_Q, g)$ is the configuration space of a particle with mass m. If the particle is free, then its *classical Hamiltonian* is by definition the function $H\colon T^*Q \to \mathbb{R}$, given by $(p, q) \mapsto g_q(p^\sharp, p^\sharp)/2m$. If there is a (nonzero) potential $V\colon Q \to \mathbb{R}$ on the configuration space, then the Hamiltonian is given by

$$(p, q) \mapsto g_q(p^\sharp, p^\sharp)/2m + V(q).$$

Let us introduce some notation for later use. Let $q \in Q$, and suppose $(e_1, \ldots, e_n)$ is a local frame induced by a chart of Q defined on some open neighbourhood of q, so that $e_1(q), \ldots, e_n(q)$ is a basis of T_qQ. Then we define $g_{jk}(q) := g_q(e_j(q), e_k(q))$ for $j, k = 1, 2, \ldots, n$. Note that the map

$$T_q^*Q \times T_q^*Q \to \mathbb{R}, \quad (p_1, p_2) \mapsto g_q(p_1^\sharp, p_2^\sharp),$$

defines an inner product on T_q^*Q. If $(e_1^*, \ldots, e_n^*)$ is the local frame dual to $(e_1, \ldots, e_n)$, then we define $g^{jk}(q) := g_q(e_j^*(q), e_k^*(q))$ for $j, k = 1, 2, \ldots, n$. Of course, the functions $(g_{jk})_{1 \leq j,k \leq n}$ and $(g^{jk})_{1 \leq j,k \leq n}$ depend on the chosen chart. Furthermore, when regarded as matrices, they are inverses of each other.

4.2.7 Proposition. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth n-dimensional manifold with boundary, let ω be the canonical symplectic form on T^*Q, and let $f\colon T^*Q \to \mathbb{R}$ be a smooth function. Then there exists a unique vector field $v_f \in \mathcal{X}(T^*Q)$ such that for each $(p, q) \in T^*Q$, we have $\omega_{(p,q)}(v_f(p, q), \cdot) = -df_{(p,q)}$. Finally, if I is an open interval, then $I \ni t \mapsto (p(t), q(t)) \in T^*Q$ is an integral curve of v_f if and only if in local coordinates, it is a solution to the following system of differential equations:*

$$\frac{dp_j}{dt}(t) = \frac{\partial f}{\partial q_j}(p(t), q(t)), \quad \frac{dq_j}{dt}(t) = -\frac{\partial f}{\partial p_j}(p(t), q(t)), \quad j = 1, 2, \ldots, n.$$

Thus, taking $f = H$, the integral curves of the corresponding vector field v_H are solutions to Hamilton's equations, which means that the motion of a particle is given by the flow of v_H.

Proof. Uniqueness follows from the fact that $\omega_{(p,q)}$ is nondegenerate for each $(p, q) \in T^*Q$. As for existence, one can check that the vector field

$$v_f = \sum_{j=1}^{n} \frac{\partial f}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial}{\partial p_j},$$

has the required properties. ■

4.2.8 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth n-dimensional manifold with boundary, let ω be the canonical symplectic form on T^*Q

- A smooth function $f\colon T^*Q \to \mathbb{R}$ is called an *observable on* (T^*Q, ω).
- Let f be an observable on (T^*Q, ω). Then the vector field v_f defined in the previous proposition is called the *Hamiltonian vector field of f*.

4.2.9 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth n-dimensional manifold with boundary, let ω be the canonical symplectic form on T^*Q, and let $f, g\colon T^*Q \to \mathbb{R}$ be smooth functions with Hamiltonian vector fields v_f and v_g. Then the *Poisson bracket* $\{f, g\}$ *of* f *and* g is given by

$$\{f, g\} := \omega(v_f, v_g) = \sum_{j=1}^{n} \frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q_j} - \frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j}.$$

4.2.10 Remark. According to classical mechanics, the time evolution of a function f on the phase space T^*Q of a smooth n-dimensional Riemannian manifold $(Q, \mathcal{D}_Q, g)$ with empty boundary is determined by the equation

$$\frac{df}{dt} = \{H, f\},$$

where H is the Hamiltonian of the system. In particular, if one takes $f = q_j$ and $f = p_j$ for $j = 1, 2, \ldots, n$, then one obtains Hamilton's equations.

4.3 Geodesics

In the absence of a potential, the trajectory of a particle on $\mathbb{R}^n$ is a straight line. The notion of a geodesic generalises this idea to arbitrary Riemannian manifolds.

4.3.1 Geodesics on manifolds with empty boundary

To a Riemannian metric on a smooth manifold with empty boundary, we can associate a canonical bilinear map on the space of vector fields.

4.3.1 Proposition. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with empty boundary. Then there exists a unique bilinear map $\nabla\colon \mathcal{X}(Q) \times \mathcal{X}(Q) \to \mathcal{X}(Q)$, $(X, Y) \mapsto \nabla_X Y$ with the following properties:*
Let $X, Y, Z \in \mathcal{X}(Q)$, and let $f \in C^\infty(Q)$. Then, we have

(i) $\nabla_{fX} Y = f \nabla_X Y$;

(ii) $\nabla_X (fY) = f \nabla_X Y + L_X(f) Y$;

(iii) $\nabla_X Y - \nabla_Y X - [X, Y] = 0$;

(iv) $L_X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)$.

Moreover, if $q \in Q$, and (U, ϕ) is a chart such that $q \in U$, then

$$\nabla_X Y(q) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} \left(X^k \frac{dY^j}{dx_k} + \sum_{l=1}^{n} \Gamma^j_{kl} X^k Y^l \right) \right) e_j(q),$$

where X^j and Y^j are the components of the vector fields X and Y at the point q with respect to the local frame $(e_1, \ldots, e_n)$ on $T_q Q$ induced by the chart (U, ϕ), and Γ^j_{kl} are the so-called Christoffel symbols, given by

$$\Gamma^j_{kl} := \frac{1}{2} \sum_{m=1}^{n} g^{jm} \left(\frac{\partial g_{lm}}{\partial x_k} + \frac{\partial g_{mk}}{\partial x_l} - \frac{\partial g_{kl}}{\partial x_m} \right).$$

.

For a proof of this proposition, see [9, Theorem 2.51].

4.3.2 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with empty boundary. The map ∇ defined in the previous proposition is called the *Levi-Civita connection* or the *canonical connection of the metric g*.

First, we note that the Levi-Civita connection is local in the following sense:

4.3.3 Lemma. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth, n-dimensional Riemannian manifold with empty boundary, and let $q_0 \in Q$.*

(1) Let $X_1, X_2, Y \in \mathcal{X}(Q)$, and suppose that $X_1(q_0) = X_2(q_0)$. Then $\nabla_{X_1} Y(q_0) = \nabla_{X_2} Y(q_0)$.

(2) Let $X, Y_1, Y_2 \in \mathcal{X}(Q)$, and suppose that there exists an open neighbourhood $U \subseteq Q$ of q_0 such that $Y_1|_U = Y_2|_U$. Then $\nabla_X Y_1(q_0) = \nabla_X Y_2(q_0)$.

Proof.
(1) We first prove that if $X_1|_U = X_2|_U$ for some open neighbourhood $U \subseteq Q$ of q, then $\nabla_{X_1} Y(q_0) = \nabla_{X_2} Y(q_0)$. Without loss of generality, we can assume that U is the domain of a chart (U, ϕ) of Q. Let $\tilde{\alpha} \colon \mathbb{R}^n \to [0,1]$ be a bump function such that $\operatorname{supp}(\tilde{\alpha}) \subset \phi(U)$ and $\tilde{\alpha}$ is equal to 1 on some neighbourhood of $\phi(q_0)$. Then $\tilde{\alpha} \circ \phi$ defines a bump function on U that can be extended to a function α in $C_0^\infty(Q)$ by extension by zero. Since $X_1|_U = X_2|_U$, we have $\alpha X_1 = \alpha X_2$, hence

$$\nabla_{X_j} Y(q_0) = \alpha(q_0) \nabla_{X_j} Y(q_0) = \nabla_{\alpha X_j} Y(q_0),$$

so $\nabla_{X_1} Y(q_0) = \nabla_{X_2} Y(q_0)$. This proves our claim. Remark that to a vector field $X \in \mathcal{X}(U)$ we can always associate an element $\tilde{X} \in \mathcal{X}(Q)$ such that X and $\tilde{X}$ coincide on some neighbourhood of q_0 by means of a bump function. It follows from the claim that the vector $\nabla_X Y(q_0) := \nabla_{\tilde{X}} Y(q_0) \in T_{q_0} Q$ is well defined, and that the vector field $\nabla_X Y \in \mathcal{X}(U)$ is indeed smooth.

Now suppose that $X_1(q_0) = X_2(q_0)$, let (U, ϕ) be a chart with $q_0 \in U$, and let $(e_1, \ldots, e_n)$ be the local frame on U induced by the chart (U, ϕ). Then there exist functions $f_{j,k} \in C^\infty(U)$ for $j = 1, 2$ and $k = 1, \ldots, n$ such that

$$X_j(q) = \sum_{k=1}^{n} f_{j,k}(q) e_k(q),$$

for $j = 1, 2$. The assumption $X_1(q_0) = X_2(q_0)$ implies $f_{1,k}(q_0) = f_{2,k}(q_0)$ for $k = 1, \ldots, n$. On the other hand, we have

$$\nabla_{X_j} Y(q_0) = \nabla_{\sum_{k=1}^{n} \alpha f_{j,k} e_k} Y(q_0) = \sum_{k=1}^{n} \alpha(q_0) f_{j,k}(q_0) \nabla_{e_k} Y(q_0) = \sum_{k=1}^{n} f_{j,k}(q_0) \nabla_{e_k} Y(q_0),$$

for $j = 1, 2$, so $\nabla_{X_1} Y(q_0) = \nabla_{X_2} Y(q_0)$, which proves the assertion.

(2) This proof is very similar to the proof of (1); suppose that U is the domain of some chart (U, ϕ), and fix a bump function $\alpha \in C_0^\infty(Q)$ as before. Then we have $\alpha Y_1 = \alpha Y_2$, and

$$\nabla_X(\alpha Y_j)(q_0) = \alpha(q_0)\nabla_X Y_j(q_0) + L_X(\alpha)(q_0)Y_j(q_0) = \nabla_X Y_j(q_0),$$

for $j = 1, 2$, since α is equal to 1 in a neighbourhood of q_0. Thus $\nabla_X Y_1(q_0) = \nabla_X Y_2(q_0)$, as desired. ■

The Levi-Civita connection maps a pair of vector fields on Q to another vector field on Q. However, the vector fields that we are interested in are in general not defined on all of Q, but rather on a curve in Q.

4.3.4 Definition. Let $(Q, \mathcal{D}_Q)$ be a smooth manifold with boundary, let $I \subseteq \mathbb{R}$ be an interval, and let $\gamma\colon I \to Q$ be a curve on Q. Then a *vector field along* γ is a lift $X\colon I \to TQ$ of γ to TQ.

4.3.5 Example. If $\gamma\colon I \to Q$ is a curve on a smooth manifold with boundary Q, then the map $\gamma'\colon I \to TQ$ given by $t \mapsto T_t\gamma(1)$ is a vector field along γ. It is the vector field that to each $t \in I$ assigns the tangent vector of the curve γ at t.

Let $(Q, \mathcal{D}_Q, g)$ be a smooth, n-dimensional Riemannian manifold with boundary, let $\gamma\colon I \to Q$ be a curve on Q, let X be a vector field along γ, and let $t_0 \in I$. We want to define $\nabla_{\gamma'}X(t_0)$, the derivative of X along the curve γ at t_0 in the direction tangent to the curve. Of course, we could do this using the expression in 4.3.1 for the covariant derivative in terms of local coordinates. We must be careful, however: γ' and X are not vector fields on Q, and in general they cannot be extended to vector fields on Q, since γ may not be injective.

It is nevertheless still possible to define $\nabla_{\gamma'}X(t_0)$. Remark that by part (2) of Lemma 4.3.3, for each $Y \in \mathcal{X}(Q)$, the vector $\nabla_{\gamma'(t_0)}Y(t_0) \in T_{\gamma(t_0)}Q$ is well defined, since the vector $\gamma'(t_0)$ can always be extended to a vector field on Q.

- If $\gamma'(t_0) = 0$, then $\nabla_{\gamma'(t_0)}Y(t_0) = 0$, so we set $\nabla_{\gamma'}X(t_0) := 0$.
- If $\gamma'(t_0) \neq 0$, then γ is an immersion at t_0, and hence an immersion on an open neighbourhood $J \subseteq I$ of t_0. By the rank theorem, there exists a chart (U, ϕ) on Q such that $\gamma(J) \subseteq U$ (restrict J to some smaller open neighbourhood of t_0 if necessary), and $\phi \circ \gamma(t) = (t, 0, 0, \ldots, 0)$. Now let $P_1\colon \mathbb{R}^n \to \mathbb{R}$ be the projection onto the first coordinate. Then the map

  $$Y = X \circ P_1 \circ \phi|_{(P_1\circ\phi)^{-1}(J)},$$

  is a vector field on an open neighbourhood of $\gamma(t_0)$ with the property that $Y(\gamma(t)) = X(t)$ for each $t \in J$. By Lemma 4.3.3, the vector $\nabla_{\gamma'(t_0)}Y(t_0)$ is well defined, and we set

  $$\nabla_{\gamma'}X(t_0) := \nabla_{\gamma'(t_0)}Y(t_0).$$

4.3.6 Lemma. *Let $(Q, \mathcal{D}_Q, g)$, γ, X and t_0 be as above. Then $\nabla_{\gamma'}X(t_0)$ is well defined.*

Proof. Clearly $\nabla_{\gamma'}X(t_0)$ is well defined if $\gamma'(t_0) = 0$, so suppose $\gamma'(t_0) \neq 0$, and construct two vector fields $Y_1, Y_2 \in \mathcal{X}(Q)$ on Q with the property that there exists an open neighbourhood $J \subseteq I$ of t_0 such that $Y_j(\gamma(t)) = X(t)$ for each $t \in J$ and $j = 1, 2$. Suppose that γ' has components $(\gamma'_1, \ldots, \gamma'_n)$ with respect to the local frame induced by some chart (U, ϕ) of Q with $t_0 \in U$. Let $(Y_j^{(1)}, \ldots, Y_j^{(n)})$ be the components of Y_j with respect to the same local frame. Since $Y_1(\gamma(t)) = X(t) = Y_2(\gamma(t))$ for each $t \in J$, we have

$$Y_1^{(l)}(t_0) = Y_2^{(l)}(t_0),$$

and

$$\begin{aligned}\sum_{k=1}^{n} \gamma'_k(t_0) \frac{\partial Y_1^{(l)}}{\partial x_k}(\phi \circ \gamma(t_0)) &= \frac{d}{dt}(Y_1(\gamma(t)))|_{t=t_0} = \frac{d}{dt}(Y_2(\gamma(t)))|_{t=t_0} \\ &= \sum_{k=1}^{n} \gamma'_k(t_0) \frac{\partial Y_2^{(l)}}{\partial x_k}(\phi \circ \gamma(t_0)),\end{aligned}$$

for $l = 1, \ldots, n$. Applying the explicit formula for ∇ stated in Proposition 4.3.1 yields $\nabla_{\gamma'(t_0)} Y_1(t_0) = \nabla_{\gamma'(t_0)} Y_2(t_0)$, which proves the assertion. ■

4.3.7 Remark. In the literature, the vector field $\nabla_{\gamma'} X$ along γ is sometimes denoted by $\frac{D}{dt} X$.

4.3.8 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with boundary.

- A curve $\gamma \colon I \to Q$ with the property that $\nabla_{\gamma'} \gamma'(t) = 0$ for each $t \in I$ is called a *geodesic*.
- If γ is a geodesic on Q with the property that there exists no geodesic $\tilde{\gamma}$ properly extending γ, then γ is called a *maximal geodesic*.
- The Riemannian manifold Q is said to be *geodesically complete* iff each maximal geodesic has domain $I = \mathbb{R}$.

4.3.9 Remark. For obvious reasons, the equation $\nabla_{\gamma'} \gamma' = 0$ is called the *geodesic equation*. Expressing γ in local coordinates, we can write it as

$$\gamma''_l + \sum_{j=1}^{n} \sum_{k=1}^{n} \Gamma^l_{jk} \gamma'_j \gamma'_k = 0.$$

4.3.10 Proposition. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with empty boundary, let $q_0 \in Q$. Then there exists a neighbourhood $U \subseteq Q$ of q_0 and $\varepsilon > 0$ such that for each element (q, v) such that $q \in U$ and $g_q(v, v)^{1/2} < \varepsilon$, there exists a unique geodesic $\gamma \colon I \to Q$ such that $[-1, 1] \subset I$, $\gamma(0) = q$ and $\gamma'(0) = v$.*

Proof. It follows from Proposition 4.3.1 that the equation $\nabla_{\gamma'} \gamma' = 0$ locally reduces to a system of second-order ordinary differential equations, and such a system has a unique solution locally. For details, we refer to [9], Theorem 2.84 and Corollary 2.85. ■

4.3.11 Corollary. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with empty boundary. Then there exists an open neighbourhood Ω of the image of the zero section $Q \to TQ$ with the property that for each $(q, v) \in \Omega$, there exists a unique geodesic $\gamma \colon I \to Q$ such that $[-1, 1] \subset I$, $\gamma(0) = q$ and $\gamma'(0) = v$.*

Now we come to the first important construction related to geodesics:

4.3.12 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with empty boundary, and let Ω be the open neighbourhood from the previous corollary. We define the *exponential map* $\exp\colon \Omega \to Q$ as the map that assigns to each element $(q, v) \in \Omega$ the point $\gamma(1)$, where $\gamma\colon I \to Q$ is the unique geodesic such that $[-1, 1] \subset I$, $\gamma(0) = q$ and $\gamma'(0) = v$.

For a proof of the following proposition, we refer to [9, Proposition 2.88(i)]

4.3.13 Proposition. *The exponential map* $\exp\colon \Omega \to Q$ *is smooth.*

The exponential map can be used to define the family of maps $(\Phi_t)_{t\in[-1,1]}$ from Ω to TQ, given by

$$\Phi_t(q, v) := \left(\exp(q, tv), \frac{d}{ds}\exp(q, sv)|_{s=t}\right).$$

The metric g induces an isomorphism $T_qQ \to (T_qQ)^*$ of vector spaces for each $q \in Q$, and hence an isomorphism $TQ \to T^*Q$ of vector bundles over Q. By abuse of notation, we shall denote this isomorphism by g as well. We can use this map to define a family of maps $(\Psi_t)_{t\in[-1,1]}$ from $g(\Omega)$ to T^*Q, where $\Psi_t := g \circ \Phi_t \circ g^{-1}$.

Now fix $(q, p) \in g(\Omega)$. Then we can consider the curve $[-1, 1] \to T^*Q$, given by $t \mapsto \Phi_t(q, p)$. Let $\Xi(q, p)$ be the derivative of this curve at 0. In this way, we obtain a vector field on $g(\Omega)$. The significance of this vector field is given by the following theorem, which can be found as Theorem 2.124 in [9]:

4.3.14 Theorem. Ξ *is the restriction of the Hamiltonian vector field* v_H *on* T^*Q *with its canonical symplectic form to the open subset* $g(\Omega) \subseteq T^*Q$.

In other words, free particles travel along the geodesics of the configuration space Q. Therefore, the motion of a free particle is complete if and only if the maximal geodesic along which it moves, has domain $\mathbb{R}$. Consequently, classical motion on Q is complete if and only if Q is geodesically complete. In the next section, we will look at a criterion for geodesic completeness. First, however, let us mention the following weak version the tubular neighbourhood theorem, which will play an important role later on:

4.3.15 Lemma. *Let g be a Riemannian metric on an open neighbourhood $U \subseteq \mathbb{R}^n$ of $x_0 \in \mathbb{R}^{n-1} \times \{0\}$ with its canonical differentiable structure, let $\tilde{U} := U \cap (\mathbb{R}^{n-1} \times \{0\})$, and let $\mathbf{n}\colon \tilde{U} \to TU$ be the unique smooth vector field on $\tilde{U}$ that assigns to each $x \in \tilde{U}$ the vector $n(x) \in T_xU$ such that:*

- $g_x(\mathbf{n}(x), \mathbf{n}(x)) = 1$*;*
- $g_x(\mathbf{n}(x), v) = 0$ *for each* $v \in T_x(\tilde{U}) \subseteq T_x\mathbb{R}^n$*;*
- $\mathbf{n}(x) \in H_n^\circ \subseteq \mathbb{R}^n$*, where we identify* $\mathbb{R}^n$ *with* T_xU *in the canonical way.*

Then there exists an open neighbourhood $V \subseteq \mathbb{R}^{n-1}$ of the projection of x_0 onto $\mathbb{R}^{n-1}$, and an $\varepsilon > 0$ such that the map $f\colon V \times (-\varepsilon, \varepsilon) \to \mathbb{R}^n$, given by $(x, t) \mapsto \exp((x, 0), t\mathbf{n}(x, 0))$, is a diffeomorphism onto its image, which is an open neighbourhood of x_0.

Proof. First, we remark that $\mathbf{n}(x)$ can be constructed by applying the Gram-Schmidt algorithm to the standard basis of $\mathbb{R}^n \cong T_xU$, and that this algorithm only involves smooth operations, so $\mathbf{n}$ is indeed a smooth map. It follows that the map

$$f_1\colon \{x \in \mathbb{R}^{n-1}\colon (x, 0) \in \tilde{U}\} \times \mathbb{R} \to TU,$$

given by $(x,t) \mapsto ((x,0), t\mathbf{n}(x,0))$, is a smooth map. The exponential map $\exp\colon \Omega \to \mathbb{R}^n$ is also smooth, hence the composition $\exp \circ f_1|_{f_1^{-1}(\Omega)}$ is smooth. Observe that

$$\exp \circ f_1|_{f_1^{-1}(\Omega)}(x_0) = x_0,$$

that

$$\frac{\partial}{\partial x_j}(\exp \circ f_1|_{f_1^{-1}(\Omega)})(x_0) = e_j,$$

for $j = 1, 2, \ldots, n-1$, where e_j is the j-th standard basis vector, and that

$$\frac{\partial}{\partial x_n}(\exp \circ f_1|_{f_1^{-1}(\Omega)})(x_0) = n(x_0) \in H_n^\circ.$$

The vectors $(e_1, e_2, \ldots, e_{n-1}, \mathbf{n}(x_0))$ form a basis of $\mathbb{R}^n \simeq T_{x_0}U$, hence $(\exp \circ f_1|_{f_1^{-1}(\Omega)})'(x_0)$ is invertible, so by the inverse function theorem, there exists an open neighbourhood $W \subseteq f_1^{-1}(\Omega)$ of x_0 such that $\exp \circ f_1|_W$ is a diffeomorphism onto an open subset of U. Now find $V \subseteq \mathbb{R}^{n-1}$ and $\varepsilon > 0$ such that $V \times (-\varepsilon, \varepsilon) \subseteq W$, and let f be the restriction of $\exp \circ f_1$ to that set. ■

4.3.2 The Riemannian distance

4.3.16 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth Riemannian manifold with empty boundary, and let $\gamma\colon I \to Q$ be a piecewise curve on Q. Then we define $L(\gamma)$, the *length of the curve* γ, by

$$L(\gamma) := \int_I g_{\gamma(t)}(\gamma'(t), \gamma'(t))^{1/2}\, dt.$$

4.3.17 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth, connected Riemannian manifold with empty boundary, let $q_1, q_2 \in Q$, and let X_{q_1,q_2} be the set of all piecewise C^1 curves on Q from q_1 to q_2. Then the *Riemannian distance* $d(q_1, q_2)$ from q_1 to q_2 is given by

$$d(q_1, q_2) := \inf\{L(\gamma)\colon \gamma \in X_{q_1,q_2}\}.$$

We call the function d the *Riemannian distance function on* $(Q, \mathcal{D}_Q, g)$.

Thus we obtain a function $d\colon Q \times Q \to [0, \infty)$. It is easy to check that $d(q,q) = 0$ for each $q \in Q$, and that d is symmetric and satisfies the triangle inequality. We can even say a bit more:

4.3.18 Proposition. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth, connected Riemannian manifold with empty boundary, let τ be the topology on Q, and let d be the Riemannian distance function on Q. Then d is a distance function on Q, and τ is the metric topology on Q induced by d.*

Proof. See [9, Proposition 2.91]. ■

Now we come to the criterion for geodesic completeness. For a proof of this theorem, we refer to [9, Corollary 2.105]

4.3.19 Theorem. (Hopf-Rinow, 1931) *Let $(Q, \mathcal{D}_Q, g)$ be a smooth, connected Riemannian manifold with empty boundary, and let d be the Riemannian distance function on Q. Then the following are equivalent:*

(1) $(Q, \mathcal{D}_Q, g)$ *is a geodesically complete manifold;*

(2) (Q, d) *is a complete metric space.*

4.3.20 Remark.

- Many of the previous statements remain true to some extent if the metric g is not smooth, but merely continuous or C^k. Proposition 4.3.18 is still true if g is continuous. Furthermore, the exponential map is C^k if the Riemannian metric is C^k, and the Hopf-Rinow theorem is still true if g is only C^1.
- Observe that throughout our discussion of geodesics, we assumed that the Riemannian manifolds had empty boundary. This is a necessary assumption; for instance, the Hopf-Rinow theorem is trivially false for $Q = [0, 1]$ with the standard metric inherited from $\mathbb{R}$.

5 Modifying phase space

Let $\Omega \subset \mathbb{R}^n$ be a bounded open subset with smooth boundary, and consider a particle with postive mass moving on $\overline{\Omega}$ with a certain (nonzero) velocity. Its motion can be described by a curve $t \mapsto (p(t), q(t)) \in \mathbb{R}^n \times \Omega$, where as before, $p(t)$ and $q(t)$ represent the momentum and the position of the particle respectively. In the absence of a potential, the particle will reach the boundary of Ω in finite time, T, say. When it does so, there are a number of possibilities; the particle may leave the domain $\overline{\Omega}$ altogether, but our simulations suggest it might also undergo an elastic collision with the boundary, or reappear at another point of the boundary and continue its journey across $\overline{\Omega}$ with possibly different momentum. In this section, we will be concerned with the last two types of motion, both of which are complete.

For an example of the first of these two kinds of motion, let $\Omega =]0,1[$, and imagine that $q(0) \in]0,1[$ and $p(0) > 0$, so that the particle moves to the right and that $q(T) = 1$. After the collision, the particle will move to the left with momentum $-p(0)$ until it reaches position 0, where it will be reflected, and its momentum after the second collision is equal to the initial momentum $p(0)$. Of course, the particle will continue to move back and forth across $[0,1]$, its momentum alternating between $p(0)$ and $-p(0)$ whenever it is defined.

For the second kind of motion, consider again the particle on $[0,1]$ with the same initial conditions, but with a periodic boundary, meaning that if a particle reaches 1 at time T, a copy of itself will appear at position 0 with the same momentum $p(T)$ the original particle had. At time $t > T$, the original particle vanishes and we dub the copy the new 'original' particle. We will do the same thing for particles moving to the left reaching position 0, where the copy appears at position 1. Our particle will move to the right, and appears at position 0 at time T, whereupon the particle, now located to the left of the interval, will move to the right with the same momentum $p(0)$ it had before it reached the right-hand side of the interval, and it will continue to move to the right until it reaches 1 again, and the whole process repeats itself indefinitely.

Note that in the first situation, the position $q(t)$ of the particle was defined for all t, but the momentum $p(t)$ was not. In the second situation, the momentum was defined everywhere, but the position was not. Thus in both cases, the curve $t \mapsto (p(t), q(t))$ was ill-defined. Alternatively, if we assign a definite momentum or position to the particle at the boundary, then the corresponding curve is discontinuous. To solve these problems, in this section, we shall consider an associated smooth curve on another manifold, and consider the curve in $\mathbb{R}^n \times \overline{\Omega}$ associated to the motion of the particle as the projection of the smooth curve on the original manifold $\mathbb{R}^n \times \overline{\Omega}$ with boundary $\mathbb{R}^n \times \partial\Omega$.

5.1 The double of a manifold with boundary

5.1.1 Construction

At last we get to the heart of the matter. Given a configuration space Q that is a manifold with boundary, we want to look at extensions of Q by attaching copies of Q to itself such that the resulting manifold no longer has a boundary. At the same time, we are interested in what happens at the level of the corresponding phase space $M = T^*Q$ of Q. One interesting construction, which can be found in [20] and [14], is the *double of* *a manifold with boundary*, in our case the configuration space Q, which is very easy to define: simply take the (disjoint) union of two copies $Q_i := \{(i, q) : q \in Q\}$ $(i = 1, 2)$ of Q, and subsequently take the quotient with respect to the following equivalence relation:

$$(1, q) \sim (2, q') \Leftrightarrow q = q' \in \mathrm{Bound}(Q).$$

For the topology on this space, take the one that is naturally associated to the operations of taking the disjoint union and the quotient.

5.1.1 Definition. We call the space $Q_1 \sqcup Q_2 / \sim$ with the above topology the *double of the topological space* Q, and denote it by $\mathrm{Double}(Q)$.

Later on, we shall see how this construction can be generalised in a useful way. This generalisation shares many properties with $\mathrm{Double}(Q)$, and since the double of a manifold is such an intuitive construction, we shall simply continue to study $\mathrm{Double}(Q)$ for the moment.

5.1.2 Proposition. *Let Q be a manifold with boundary and let* $\mathrm{Double}(Q)$ *be its double.*

(1) The maps $\iota_i \colon Q \to \mathrm{Double}(Q)$, $i = 1, 2$, given by $q \mapsto [i, q]$, are homeomorphisms onto their images.

(2) The canonical projection $P \colon \mathrm{Double}(Q) \to Q$, given by $[i, q] \mapsto q$, is continuous and open.

(3) $\mathrm{Double}(Q)$ *is Hausdorff.*

(4) $\mathrm{Double}(Q)$ *is second countable.*

Proof.
(1) It is clear from the definition of $\mathrm{Double}(Q)$ and the definitions of the disjoint union and quotient topologies that the maps ι_i, $i = 1, 2$, are continuous bijections onto their images. Now suppose $U \subseteq Q$ is open and let $i \in \{1, 2\}$. Then the set $V := \iota_1(U) \cup \iota_2(U)$ is open in $\mathrm{Double}(Q)$, hence $\iota_i(U) = \iota_i(U) \cap V$ is open in $\iota_i(U)$. Thus ι_i is a homeomorphism onto its image.

(2) Let $U \subseteq Q$ be open. Then $P^{-1}(U) = \iota_1(U) \cup \iota_2(U)$ is open in $\mathrm{Double}(Q)$, hence P is continuous. Now suppose $V \subseteq \mathrm{Double}(Q)$ is open in $\mathrm{Double}(Q)$. Let $q \in P(V)$. Then there exists an $i \in \{1, 2\}$ such that $[i, q] = \iota_i(q) \in V$, and by the previous part of the proposition, $\iota_i^{-1}(V) \subseteq P(V)$ is an open subset of Q containing q. Thus $P(V)$ is open, and it follows that P is open.

(3) Let $[i, q], [j, r] \in \mathrm{Double}(Q)$, and suppose $[i, q] \neq [j, r]$. Then there are two possibilities:

- $q \neq r$: Since Q is Hausdorff, there exist open sets $U, V \subset Q$ such that $q \in U$, $r \in V$ and $U \cap V = \emptyset$. But then $\iota_1(U) \cup \iota_2(U)$ and $\iota_1(V) \cup \iota_2(V)$ are two disjoint open subsets of $\mathrm{Double}(Q)$ containing $[i, q]$ and $[j, r]$ respectively.
- $q = r$ and $i \neq j$: Then $q \in \mathrm{Int}(Q)$, and $\iota_1(\mathrm{Int}(Q))$ and $\iota_2(\mathrm{Int}(Q))$ are two disjoint open subsets of $\mathrm{Double}(Q)$ containing $[1, q]$ and $[2, q]$ respectively.

We conclude that Double(Q) is Hausdorff.

(4) Pick a countable base $(U_m)_{m\in\mathbb{N}}$ of the topology on Q. Let $U \subseteq \mathrm{Double}(Q)$ be open and let $[i,q] \in U$. Without loss of generality, we may assume that $i = 1$. Then $\iota_j^{-1}(U)$ is open in Q for $j = 1, 2$, and $\iota_1^{-1}(U)$ contains q. Again, there are two possibilities:

- $q \notin \iota_2^{-1}(U)$: Then in particular, $q \notin \mathrm{Bound}(Q)$, so we can pick an $m \in \mathbb{N}$ such that $U_m \subseteq \iota_1^{-1}(U) \cap \mathrm{Int}(Q)$ is an open neighbourhood of q. Then $\iota_1(U_m)$ is an open neighbourhood of $[1,q]$ contained in U.
- $q \in \iota_2^{-1}(U)$: Then pick $m \in \mathbb{N}$ such that $U_m \subseteq \iota_1^{-1}(U) \cap \iota_2^{-1}(U)$ is an open neighbourhood of q. Then $\iota_1(U_m) \cup \iota_2(U_m)$ is an open neighbourhood of $[1,q]$ contained in U.

We conclude that

$$\{\iota_i(U_m) \colon m \in \mathbb{N},\ U_m \subseteq \mathrm{Int}(Q),\ i = 1,2\} \cup \{\iota_1(U_m) \cup \iota_2(U_m) \colon m \in \mathbb{N}\}$$

is a countable base of the topology on Double(Q), so Double(Q) is second countable. ■

Thus we have obtained a new topological space, which is both Hausdorff and second countable. We want to endow Double(Q) with the structure of a manifold with empty boundary such that ι_1 and ι_2 are smooth embeddings of Q into Double(Q), i.e., aside from being homeomorphisms onto their images, ι_1 and ι_2 should be smooth maps with the property that the tangent map is injective at every point of Q.

It is clear how to define charts whose domains are contained in $\iota_1(\mathrm{Int}(Q))$ or $\iota_2(\mathrm{Int}(Q))$. There is, however, no canonical differentiable structure on Q, a problem that presents itself when one tries to define charts whose domains have nonempty intersection with $\iota_1(\mathrm{Bound}(Q)) = \iota_2(\mathrm{Bound}(Q))$. Consider, for example, $Q = H_2 = \mathbb{R} \times [0, \infty[$, and let $Q_i := \iota_i(Q)$ for $i = 1, 2$. One is tempted to identify Q_1 with H_2 via the inclusion $Q \hookrightarrow \mathbb{R}^2$ and Q_2 with $-H_2 = \mathbb{R}\times]-\infty, 0]$ via the same inclusion and subsequently taking the mirror image in the x-axis, and thus identify Double(Q) with $\mathbb{R}^2$.

But there are other ways to include Q into $\mathbb{R}^2$ such that Q is mapped onto itself: take, for instance, the map $(x, y) \mapsto (x + \lambda y, y)$, where $\lambda \in \mathbb{R}\backslash\{0\}$. This map is a smooth homeomorphism onto its image, and its inverse is smooth as well; hence it is a chart of Q. If we now use this chart to construct a chart $\mathrm{Double}(Q) \to \mathbb{R}^2$, then we see that the curve $\gamma\colon \mathbb{R} \to \mathrm{Double}(Q)$,

$$\gamma(t) := \begin{cases} [2, |t|] & t < 0 \\ [1, |t|] & t \geq 0 \end{cases},$$

is not smooth with respect to the second differentiable structure, whereas it is smooth with respect to the differentiable structure defined in the previous paragraph. Thus we have obtained an infinite family of different differentiable structures on Double(Q), and the question is which one we should pick. There is a way to define a family of differentiable structures on Double(Q) such that for any two members $\mathcal{D}$ and $\mathcal{D}'$ of this family, the associated manifolds $(\mathrm{Double}(Q), \mathcal{D})$ and $(\mathrm{Double}(Q), \mathcal{D}')$ are diffeomorphic; cf. [14, Theorem 6.3]. Still, we would like to single out a specific differentiable structure. For an arbitrary manifold with boundary, this is impossible.

If the manifold carries a Riemannian metric, however, then we will see that there exists a canonical differential structure on the double of that manifold. Note that it is necessary in both classical and quantum mechanics to have a Riemannian metric on the configuration space in order to be able to define the Hamiltonian of the system. There is one drawback: in general, the inherited Riemannian structure on $\mathrm{Double}(Q)$ is not smooth everywhere, but only on $P^{-1}(\mathrm{Int}(Q))$.

5.1.3 Lemma. *Let $U \subseteq H_n$ be an open subset of H_n. We view U as a smooth manifold with boundary by endowing it with its canonical differentiable structure. Let g be a Riemannian metric on U, and let $x_0 \in U$. Then there exists an open neighbourhood $V \subseteq \mathbb{R}^n$ of x_0 carrying a Riemannian metric $\tilde{g}$ such that $V \cap H_n \subseteq U$ and $\tilde{g}|_{V\cap H_n} = g|_{V\cap H_n}$.*

Proof. Since the Riemannian metric g is defined on the open subset U of H_n, we may regard it as a smooth map A from U to the set of real $n \times n$-matrices $M(n, \mathbb{R})$, given by $x \mapsto A(x) := (g(x)_{ij})_{1\le i,j\le n}$, and $A(x)$ is symmetric and $v^T A(x)v > 0$ for each column vector $v \in \mathbb{R}^n$. The fact that A is smooth implies that there exists a smooth map $\tilde{A}$ on some open neighbourhood $\tilde{V} \subseteq \mathbb{R}^n$ of x_0 such that $\tilde{A}|_{\tilde{V}\cap H_n} = A|_{\tilde{V}\cap H_n}$. But then the map $\hat{A}\colon \tilde{V} \to M(n, \mathbb{R})$, given by

$$x \mapsto \frac{1}{2}(\tilde{A}(x) + (\tilde{A}(x))^T),$$

is another smooth map with the same property as $\tilde{A}$, and in addition, for each $x \in \tilde{V}$, the matrix $\hat{A}(x)$ is symmetric. Thus we may assume without loss of generality that $\tilde{A}(x)$ is symmetric for each $x \in \tilde{V}$.

Now we show that there exists an open neighbourhood $V \subseteq \tilde{V}$ of x_0 such that $v^T \tilde{A}(x)v > 0$ for each column vector $v \in \mathbb{R}^n\backslash\{0\}$ and each $x \in V$. Let $S = S^{n-1}$ be the unit sphere in $\mathbb{R}^n$. The map $f\colon \tilde{V} \times \mathbb{R}^n \to \mathbb{R}$, defined by $(x, v) \mapsto v^T \tilde{A}(x)v$ is smooth, so $f^{-1}(]0, \infty[)$ is open in $\tilde{V} \times \mathbb{R}^n$. Hence, for each $v \in S$, there exist open subsets $V_v \subseteq \tilde{V}$ and $W_v \subseteq \mathbb{R}^n$ such that $(x_0, v) \in V_v \times W_v$. Now choose a family of pairs of open sets $\{(V_v, W_v)\}_{v\in S}$ with this property. Then $\{W_v\}_{v\in S}$ is an open cover of S, so by compactness, there exist $v_1, \ldots, v_m$ such that $S \subseteq \bigcup_{j=1}^m W_{v_j}$. Let $V := \bigcap_{j=1}^m V_{v_j}$. Then $V \subseteq \tilde{V}$ is an open neighbourhood of x_0 with the property that for each $x \in V$ and each $v \in S$, we have $v^T \tilde{A}(x)v > 0$, and hence, for each $x \in V$ and each $v \in \mathbb{R}^n\backslash\{0\}$, we have $v^T \tilde{A}(x)v > 0$. Thus we can use the function $\tilde{A}|_V$ to define a Riemannian metric $\tilde{g}$ on V with the desired property. ■

5.1.4 Theorem. *Let $(Q, \mathcal{D}_Q, g)$ be an n-dimensional Riemannian manifold with boundary. Then there exists a differentiable structure $\mathcal{D}$ on* $\mathrm{Double}(Q)$ *such that ι_1 and ι_2 are smooth embeddings, and there exists a unique continuous section $\tilde{g}\colon \mathrm{Double}(Q) \to T_0^2\mathrm{Double}(Q)$ such that $g = \iota_1^*(\tilde{g}) = \iota_2^*(\tilde{g})$.*

Proof. (I) We shall begin by constructing an atlas $\mathcal{A}$ on $\mathrm{Double}(Q)$. First, suppose that $q_0 \in \mathrm{Int}(Q)$, so that $[j, q_0] \in P^{-1}(\mathrm{Int}(Q))$ for $j = 1, 2$. Then we can find a chart (U, ϕ) of Q such that $q_0 \in U \subseteq \mathrm{Int}(Q)$. Now, we let $(\iota_j(U), \phi \circ P|_{\iota_j(U)})$ be elements of $\mathcal{A}$. It follows from Proposition 5.1.2 that $\iota_j(U)$ is an open subset of $\mathrm{Double}(Q)$ and that $\phi \circ P|_{\iota_j(U)}$ is a homeomorphism onto its image. Furthermore, if (V, ψ) is another chart of Q such that $q_0 \in V \subseteq \mathrm{Int}(Q)$, then

$$(\psi \circ P|_{\iota_j(U\cap V)}) \circ (\phi \circ P|_{\iota_j(U\cap V)})^{-1} = \psi \circ (\phi|_{U\cap V})^{-1},$$

so the transition functions are smooth maps.

Next, we specify how to construct charts for points in $P^{-1}(\text{Bound}(Q))$. Let $q_0 \in \text{Bound}(Q)$, and let (U, ϕ) be a chart of Q such that $q_0 \in U$. Let $(\phi^{-1})^*(g)\colon \phi(U) \to T_0^2\phi(U)$ be the map given by

$$(\phi^{-1})^*(g)_x(v, w) := g_{\phi^{-1}(x)}(T_x\phi^{-1}(v), T_x\phi^{-1}(w)).$$

$(\phi^{-1})^*(g)$ is a Riemannian metric on $\phi(U)$, so by Lemma 5.1.3, there exists a set $V \subseteq \phi(U)$ that is open in $\mathbb{R}^n$ and that contains $x_0 := \phi(q_0)$, and there exists a Riemannian metric $\tilde{g}$ on V such that $\tilde{g}|_{V\cap H_n} = (\phi^{-1})^*(g)|_{V\cap H_n}$. By Lemma 4.3.15, there exists an open neighbourhood $W \subset \mathbb{R}^{n-1}$ of the projection of x_0 onto $\mathbb{R}^{n-1}$, and an $\varepsilon > 0$ such that the map

$$\Psi\colon W \times (-\varepsilon, \varepsilon) \to V, \quad (x, t) \mapsto \exp((x, 0), t\mathbf{n}(x, 0)),$$

is a diffeomorphism onto its image, which is open in V and hence in $\mathbb{R}^n$. Here, $\mathbf{n}$ is the unit normal on $W \times \{0\}$ that satisfies $n(x, 0) \in H_n^\circ$ for each $x \in W$.

We claim that W and $\varepsilon > 0$ can be chosen in such a way that $\Psi^{-1}(H_n^\circ) \subseteq H_n^\circ$, $\Psi^{-1}(H_n^\circ) \subseteq -H_n^\circ$ and $\Psi^{-1}(\mathbb{R}^{n-1} \times \{0\}) \subseteq \mathbb{R}^{n-1} \times \{0\}$. Indeed, let $P_n\colon \mathbb{R}^n \to \mathbb{R}$ be the projection on the n-th coordinate, and consider the number

$$a := P_n\left(\frac{\partial\Psi}{\partial x_n}(x_0)\right) = P_n(\mathbf{n}(x_0)).$$

Then $a > 0$, since $\mathbf{n}(x_0) \in H_n^\circ$. Because Ψ is C^1 at x_0, there exists an open neighbourhood $N \subseteq W\times]-\varepsilon, \varepsilon[$ of x_0 such that

$$\frac{\partial(P_n \circ \Psi)}{\partial x_n}(x) = P_n\left(\frac{\partial\Psi}{\partial x_n}(x)\right) > a/2$$

for each $x \in N$. Without loss of generality, we may assume that N is of the form $W'\times]-\varepsilon', \varepsilon'[$. Now let $y \in W'$. Then it follows from the mean value theorem that for each $b \in]0, \varepsilon'[$, we have

$$P_n \circ \Psi(y, b) = P_n(\Psi(y, b) - \Psi(y, 0)) \geq \frac{ab}{2} > 0,$$

and similarly, $P_n \circ \Psi(y, -b) \leq \frac{-ab}{2} < 0$. This proves the claim.

Now choose W and ε with the aforementioned properties. Let $r\colon \mathbb{R}^n \to \mathbb{R}^n$ be the map

$$(x_1, \ldots, x_{n-1}, x_n) \mapsto (x_1, \ldots, x_{n-1}, -x_n), \tag{5.1}$$

let $\tilde{U} := P^{-1}(\phi^{-1} \circ \Psi(W \times [0, \varepsilon[))$, and define the map $\tilde{\phi}\colon \tilde{U} \to W\times]-\varepsilon, \varepsilon[$ as follows:

$$\tilde{\phi}([i, q]) := \begin{cases} \Psi^{-1} \circ \phi(q) & i = 1 \\ r \circ \Psi^{-1} \circ \phi(q) & i = 2 \end{cases}.$$

The set $\Psi(W\times[0, \varepsilon[)$ is an open subset of H_n containing x_0, so $\phi^{-1}\circ\Psi(W\times[0, \varepsilon[)$ is open in U, hence is also open in Q, and it follows that $\tilde{U}$ is open in Double(Q) and contains $[1, q_0]$.

(II) We show that $\tilde{\phi}$ is a well-defined homeomorphism. If $q \in \mathrm{Bound}(Q)$, then $\phi(q) \in \mathbb{R}^{n-1} \times \{0\}$, so $\Psi^{-1} \circ \phi(q) \in \mathbb{R}^{n-1} \times \{0\}$ (note that $\Psi^{-1} \circ \phi(q) = \phi(q)$) and it follows that $\Psi^{-1} \circ \phi(q) = r \circ \Psi^{-1} \circ \phi(q)$, so $\tilde{\phi}$ is well-defined. To see that $\tilde{\phi}$ is a bijection, we note that an inverse of $\tilde{\phi}$ is the map $W \times]-\varepsilon, \varepsilon[\to \tilde{U}$ given by

$$(y,t) \mapsto \begin{cases} \iota_1 \circ \phi^{-1} \circ \Psi(y,t) & t \geq 0 \\ \iota_2 \circ \phi^{-1} \circ \Psi \circ r(y,t) & t < 0 \end{cases}.$$

In view of the fact that ϕ and Ψ are homeomorhisms, we conclude that $\tilde{\phi}$ is continuous at each point of $\tilde{U} \cap P^{-1}(\mathrm{Int}(Q))$ and that $\tilde{\phi}^{-1}$ is continuous at each point of $W \times (]-\varepsilon, \varepsilon[\setminus \{0\})$.

Now suppose $q \in P(\tilde{U}) \cap \mathrm{Bound}(Q)$, and let $Z \subseteq W \times]-\varepsilon, \varepsilon[$ be an open neighbourhood of $x = \tilde{\phi}([1,q]) = \phi(q)$. Then $\phi^{-1}(H_n \cap \Psi(Z))$ and $\phi^{-1}(H_n \cap \Psi \circ r(Z))$ are open subsets of Q containing q, hence

$$Y := P^{-1}((\phi^{-1}(H_n \cap \Psi(Z))) \cap (\phi^{-1}(H_n \cap \Psi \circ r(Z)))) \subseteq \tilde{U}$$

is an open subset of $\mathrm{Double}(Q)$ containing $[1,q]$, with the property that $\tilde{\phi}(Y) \subseteq Z$. Thus $\tilde{\phi}$ is continuous at $[1,q]$, and it follows that $\tilde{\phi}$ is continuous at each point of its domain.

Suppose $x \in W \times \{0\}$, and let $Y \subseteq \tilde{U}$ be an open neighbourhood of $q := \phi^{-1}(x)$. Then $(\iota_1 \circ \phi^{-1})^{-1}(Y)$ and $(\iota_2 \circ \phi^{-1})^{-1}(Y)$ are open subsets of H_n. For each $x' \in \mathbb{R}^n$, we have $x' \in H_n$ if and only if $\Psi(x') \in H_n$, and Ψ is a homeomorphism onto its image, so $(\iota_1 \circ \phi^{-1} \circ \Psi)^{-1}(Y)$ and $(\iota_2 \circ \phi^{-1} \circ \Psi)^{-1}(Y)$ are open subsets of H_n containing x, and hence

$$Z := (\iota_1 \circ \phi^{-1} \circ \Psi)^{-1}(Y) \cap (\iota_2 \circ \phi^{-1} \circ \Psi \circ r)^{-1}(Y) \subseteq W_1 \times (-\delta_1, \delta_1),$$

is an open subset of $\mathbb{R}^n$ containing x, with the property that $\tilde{\phi}^{-1}(Z) \subseteq Y$. Thus $\tilde{\phi}^{-1}$ is continuous at x, and it follows that $\tilde{\phi}$ is a homeomorphism onto its image.

(III) We have shown that $(\tilde{U}, \tilde{\phi})$ is a chart of $\mathrm{Double}(Q)$. We want to show that charts of this type, together with the ones defined on $P^{-1}(\mathrm{Int}(Q))$, form an atlas of $\mathrm{Double}(Q)$. The only nontrivial thing left to check here is that, given two charts (U_1, ϕ_1), (U_2, ϕ_2) of Q whose domains contain a common element q_0 of the boundary, the transition function corresponding to two associated charts $(\tilde{U}_1, \tilde{\phi}_1)$, $(\tilde{U}_2, \tilde{\phi}_2)$ on $\mathrm{Double}(Q)$ is smooth at q_0. On a side note, we remark that it is possible that $(\tilde{U}_1, \tilde{\phi}_1) \neq (\tilde{U}_2, \tilde{\phi}_2)$ even if $(U_1, \phi_1) = (U_2, \phi_2)$, since we made some choices in the construction of the chart $(\tilde{U}, \tilde{\phi})$ from the chart (U, ϕ).

Let $q \in U_1 \cap U_2 \cap \mathrm{Bound}(Q)$, and define $x_1, x_2 \in \mathbb{R}^{n-1}$ by $(x_j, 0) = \phi_j(q)$ for $j = 1, 2$. Then for sufficiently small $\varepsilon > 0$, the curves $\gamma_j \colon]-\varepsilon, \varepsilon[\to V_j$, given by $t \mapsto \Psi_j(x_j, t)$, are geodesics such that $\gamma_j(0) = (x_j, 0)$ and $\gamma_j'(0) = \mathbf{n}_j(x_j, 0)$ for $j = 1, 2$. But $\phi_2 \circ (\phi_1|_{U_1 \cap U_2})^{-1}$ is an isomorphism between the Riemannian manifolds $(\phi_1(U_1 \cap U_2), (\phi_1^{-1})^*(g)|_{\phi_1(U_1 \cap U_2)})$ and $(\phi_2(U_1 \cap U_2), (\phi_2^{-1})^*(g)|_{\phi_2(U_1 \cap U_2)})$ that maps $(x_1, 0)$ to $(x_2, 0)$ and whose tangent map sends $\mathbf{n}_1(x_1, 0)$ to $\mathbf{n}_2(x_2, 0)$. So by the uniqueness of geodesics we have $\phi_2^{-1} \circ \gamma_2|_{[0,\varepsilon[} = \phi_1^{-1} \circ \gamma_1|_{[0,\varepsilon[}$, and therefore

$$\tilde{\phi}_1^{-1}(x_1, t) = \tilde{\phi}_2^{-1}(x_2, t),$$

for each $t \in]-\varepsilon, \varepsilon[$. But this means that $\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}|_{\tilde{\phi}_1(\tilde{U}_1 \cap \tilde{U}_2)}$ is of the form $\psi \times \mathrm{Id}_{\mathbb{R}}$, where $\psi \colon N \to \mathbb{R}^{n-1}$ is a diffeomorphism onto its image on some open neighbourhood $N \subseteq \mathbb{R}^{n-1}$

of x_1, hence $\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}|_{\tilde{\phi}_1(\tilde{U}_1 \cap \tilde{U}_2)}$ is smooth at $(x_1, 0)$. We conclude that the set $\mathcal{A}$ containing all of the constructed charts $(\tilde{U}, \tilde{\phi})$ is an atlas of Double(Q).

(IV) Next, observe that the pairs $(\iota_k^{-1}(\tilde{U}_j), \tilde{\phi}_j \circ \iota_k)$ are charts of Q for $j, k = 1, 2$, which implies that ι_1 and ι_2 are smooth immersions. It follows from part (1) of Proposition 5.1.2 that ι_1 and ι_2 are smooth embeddings.

Now let $[j, q] \in$ Double(Q), and let (U, ϕ) be a chart of Q such that $q \in U$. Construct a chart $(\tilde{U}, \tilde{\phi})$ on Double(Q) such that $[j, q] \in \tilde{U}$ as above. Define

$$\tilde{g}_{[j,q]} \colon T_{[j,q]}\mathrm{Double}(Q) \times T_{[j,q]}\mathrm{Double}(Q) \to \mathbb{R},$$

by

$$\tilde{g}_{[j,q]}([(\tilde{U}, \tilde{\phi}), v], [(\tilde{U}, \tilde{\phi}), w]) := g_q([(\iota_j^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_j), v], [(\iota_j^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_j), w]).$$

We show that $\tilde{g}_{[j,q]}$ is well defined. Suppose $q \in$ Bound(Q). Then for each $v \in \mathbb{R}^n$, we have

$$[(\iota_2^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_2), v] = [(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1), r(v)],$$

where r is the map from equation (5.1). By construction of $\tilde{\phi}$, the vector $[(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1), e_n]$ is orthogonal to the vectors $[(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1), e_j]$, $j = 1, \ldots, n-1$, so the map $T_qQ \to T_qQ$, given by

$$[(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1), v] \mapsto [(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1), r(v)],$$

is unitary. Thus $\tilde{g}_{[j,q]}$ is well defined. It is now easy to see that $g = \iota_j^*(\tilde{g})$ for $j = 1, 2$, and that $\tilde{g} \colon \mathrm{Double}(Q) \to T_0^2\mathrm{Double}(Q)$ is a continuous section of the canonical projection $T_0^2\mathrm{Double}(Q) \to \mathrm{Double}(Q)$, which is smooth on $P^{-1}(\mathrm{Int}(Q))$. ■

5.1.5 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a Riemannian manifold with boundary, let $\mathcal{D}$ be the differentiable structure on Double(Q) associated to the atlas constructed in Theorem 5.1.4, and let $\tilde{g} \colon \mathrm{Double}(Q) \to T_0^2\mathrm{Double}(Q)$ be the (continuous) Riemannian metric from the same theorem. We call the triple $(\mathrm{Double}(Q), \mathcal{D}, \tilde{g})$ the *Riemannian double of* $(Q, \mathcal{D}_Q, g)$.

5.1.6 Remark.

- In the above proof, we had to make certain choices while constructing of $(\tilde{U}, \tilde{\phi})$ from (U, ϕ), and consequently, there is no canonical map sending charts (U, ϕ) of Q to charts $(\tilde{U}, \tilde{\phi})$. Thus we have implicitly used the axiom of choice.
- The metric $\tilde{g}$ on Double(Q) is in general not smooth or even continuously differentiable on $P^{-1}(\mathrm{Bound}(Q))$. As a counterexample, consider the closed unit disc $D \subseteq \mathbb{R}^2$ with the Riemannian metric inherited from the canonical Riemannian metric on $\mathbb{R}^2$. Polar coordinates $(r, \theta) \in]0, 1] \times] - \pi, \pi[$ can be used to parametrise an open neighbourhood of $(1, 0) \in D$. Since the geodesics of $\mathbb{R}^2$ are precisely the curves traversing straight lines with constant velocity, and since the radial coordinate defines curves perpendicular to the boundary of D, the relation between polar coordinates and a possible set of coordinates (u, v) that one could obtain from Lemma 4.3.15 is given by $(u, v) = (\theta, 1 - r)$. The basis vector e_u induced by these coordinates satisfies

$$g_{(u,v)}(e_u, e_u) = (1 - v)^2,$$

so
$$\frac{\partial}{\partial v}(g_{(u,v)}(e_u, e_u))|_{(u,v)=(0,0)} = -2 \neq 0,$$
so if we use these coordinates to obtain coordinates on $\mathrm{Double}(D)$ to parametrise a neighbourhood of $[1,(1,0)]$, then $g_{(u,v)}(e_u, e_u)$ is not differentiable at the point $(1,0) \in D$.

One can ask to what extent the differentiable structure associated to the atlas constructed in the previous theorem, is unique. It is unclear whether it is unique as a C^∞-structure. Assume for the moment that the transition functions of the elements of an atlas $\mathcal{A}$ are not smooth, but merely C^1 at points of the boundary. Let us call such an atlas 'smooth on the interior and C^1 at the boundary', and let us adopt the same terminology for the corresponding notion of a differentiable structure. Then we have the following theorem:

5.1.7 Theorem. *Let $(Q, \mathcal{D}_Q, g)$ be an n-dimensional Riemannian manifold with boundary, and let $\mathcal{A}$ be the atlas constructed in Theorem 5.1.4. Then there exists a unique differentiable structure $\mathcal{D}$ on* $\mathrm{Double}(Q)$ *containing $\mathcal{A}$ that is smooth on the interior and C^1 at the boundary, such that ι_1 and ι_2 are smooth embeddings.*

Proof. (I) Let $\mathcal{A}'$ be an atlas such that $\mathrm{Double}(Q)$ with the corresponding differentiable structure $\mathcal{D}'$ is a differentiable manifold that is smooth on $P^{-1}(\mathrm{Int}(Q))$ and C^1 on $P^{-1}(\mathrm{Bound}(Q))$ and such that ι_1 and ι_2 are smooth embeddings. We may assume without loss of generality that both $\mathcal{A}$ and $\mathcal{A}'$ are maximal. Our goal is to show that $\mathcal{A} \cup \mathcal{A}'$ is an atlas. To do so, it suffices to show that for each $[j,q] \in \mathrm{Double}(Q)$, there exist charts $(U,\phi) \in \mathcal{A}$ and $(V,\psi) \in \mathcal{A}'$, and an open subset $W \subseteq U \cap V$ such that $[j,q] \in W$ and with the property that $\psi \circ \phi^{-1}|_{\phi(W)}$ is a C^1-diffeomorphism onto an open subset of $\mathbb{R}^n$.

Therefore, let $[j,q] \in \mathrm{Double}(Q)$, and let $(V,\psi) \in \mathcal{A}'$. By assumption, $\iota_j \colon Q \hookrightarrow \mathrm{Double}(Q)$ is a smooth embedding, so there exists a chart (U,ϕ) of Q such that $q \in U$, $\iota_j(U) \subseteq V$, and such that
$$\psi \circ \iota_j \circ \phi^{-1} \colon H_n \supseteq \phi(U) \to \mathbb{R}^n.$$
is a smooth map. If $q \in \mathrm{Bound}(Q)$, then we assume that $\iota_j(U) \subseteq V$ for $j = 1, 2$.

First we discuss the case $[j,q] \in P^{-1}(\mathrm{Int}(Q))$. By restricting U to $U \cap \mathrm{Int}(Q)$ if necessary, we may assume that $U \subseteq \mathrm{Int}(Q)$. Then $(\tilde{U}, \tilde{\phi}) := (\iota_j(U), \phi \circ P)$ is an element of $\mathcal{A}$, and
$$\psi \circ \tilde{\phi}^{-1} = \psi \circ \iota_j \circ \phi^{-1} \colon H_n^\circ \supseteq \phi(U) \to \mathbb{R}^n,$$
is a smooth map. In particular, it is differentiable at $\phi(q)$, and
$$(\psi \circ \tilde{\phi}^{-1})'(\phi(q)) = T_{[j,q]}\psi \circ T_q\iota_j \circ (T_q\phi)^{-1}.$$
The map ι_j is an immersion, so $(\psi \circ \tilde{\phi}^{-1})'(\phi(q))$ is injective, and hence is an isomorphism of vector spaces. By the inverse function theorem, there exists an open neighbourhood $\tilde{W} \subseteq H_n^\circ$ of $\phi(q)$ such that $\psi \circ \tilde{\phi}|_{\tilde{W}}$ is a diffeomorphism onto an open subset of $\mathbb{R}^n$, so $W := \tilde{\phi}^{-1}(\tilde{W})$ is the desired neighbourhood of $[j,q]$.

(II) Before we discuss the case $[j,q] \in P^{-1}(\mathrm{Bound}(Q))$, we prove that $P^{-1}(\mathrm{Bound}(Q))$ is an embedded C^1-submanifold of $(\mathrm{Double}(Q), \mathcal{D}')$. Indeed, suppose (U,ϕ) and (V,ψ)

are charts as above. Then the smooth map $\psi \circ \iota_j \circ \phi^{-1}$ has a smooth extension f to some open subset of $\mathbb{R}^n$ containing $\phi(U)$. Again, since ι_j is an immersion, we can use an argument similar to the one in the previous paragraph to show that there exists an open neighbourhood $W \subseteq \mathbb{R}^n$ of $\phi(q)$ such that $f|_W$ is a C^1-diffeomorphism onto an open subset of $\mathbb{R}^n$. By replacing W with $W \cap f^{-1}(\psi(V))$ if necessary, we may assume that the image of $f|_W$ is contained in the image of ψ, so

$$(\tilde{V}, \tilde{\psi}) := (\psi^{-1}(f(W)), (f|_W)^{-1} \circ \psi|_{\psi^{-1}(f(W))}) \in \mathcal{A}',$$

is a chart such that

$$\tilde{\psi}(\tilde{V} \cap P^{-1}(\mathrm{Bound}(Q))) = \tilde{\psi}(\tilde{V}) \cap (\mathbb{R}^{n-1} \times \{0\}).$$

We conclude that $P^{-1}(\mathrm{Bound}(Q))$ is an embedded C^1-submanifold of $(\mathrm{Double}(Q), \mathcal{D}')$.

(III) Now suppose $[j, q] \in P^{-1}(\mathrm{Bound}(Q))$, and let (U, ϕ) and (V, ψ) are charts, as before. Since $P^{-1}(\mathrm{Bound}(Q))$ is an embedded C^1-submanifold of $(\mathrm{Double}(Q), \mathcal{D}')$, we may assume without loss of generality that

$$\psi(V \cap P^{-1}(\mathrm{Bound}(Q))) = \psi(V) \cap (\mathbb{R}^{n-1} \times \{0\}),$$

and, applying a reflection if necessary,

$$\psi(V \cap \iota_1(Q)) = \psi(V) \cap H_n.$$

Construct a chart $(\tilde{U}, \tilde{\phi}) \in \mathcal{A}$ such that $[j, q] \in \tilde{U}$ from the chart (U, ϕ) following the method outlined in the existence-part of the proof. By construction of $\tilde{\phi}$, we have $\tilde{\phi}^{-1}(x) = \iota_j \circ \phi^{-1}(x)$ for each $x \in \tilde{\phi}(\tilde{U}) \cap (\mathbb{R}^{n-1} \times \{0\})$, so

$$\begin{aligned} w_k &:= \lim_{t \to 0} \frac{\psi \circ \tilde{\phi}^{-1}(\tilde{\phi}([j, q]) + te_k) - \psi([j, q])}{t} \\ &= \lim_{t \to 0} \frac{\psi \circ \iota_j \circ \phi^{-1}(\phi(q) + te_k) - \psi([j, q])}{t} \\ &= \frac{\partial}{\partial x_k} (\psi \circ \iota_j \circ \phi^{-1})(\phi(q)), \end{aligned}$$

for $k = 1, \ldots, n-1$. On the other hand, we have

$$\frac{\partial}{\partial x_k} (\psi \circ \iota_j \circ \phi^{-1})(\phi(q)) = T_{[j,q]}\psi \circ T_q \iota_j \circ (T_q \phi)^{-1}(e_k),$$

and ι_j is an immersion, so $(w_1, \ldots, w_{n-1})$ is a linear independent system of vectors. Thus the first $n-1$ partial derivatives of $\psi \circ \tilde{\phi}^{-1}$ at $\phi(q)$ exist, and are linearly independent. We want to show that the n-th partial derivative of $\psi \circ \tilde{\phi}^{-1}$ at $\phi(q)$ exists. Consider the limits

$$w_n^{\pm} := \lim_{t \to 0^{\pm}} \frac{\psi \circ \tilde{\phi}^{-1}(\tilde{\phi}([j, q]) + te_n) - \psi([j, q])}{t}.$$

The limit w_n^+ exists, because we can write $\psi \circ \tilde{\phi}^{-1}$ as a composition of two smooth functions as follows:

$$\psi \circ \tilde{\phi}^{-1}|_{\tilde{\phi}(\iota_1(Q) \cap \tilde{U})} = (\psi \circ \iota_1 \circ (\phi|_{\iota_1^{-1}(\tilde{U})})^{-1}) \circ (\tilde{\phi} \circ \iota_1 \circ (\phi|_{\iota_1^{-1}(\tilde{U})})^{-1})^{-1}.$$

Likewise w_n^- exists, since

$$\psi \circ \tilde{\phi}^{-1}|_{\tilde{\phi}(\iota_2(Q)\cap\tilde{U})} = (\psi \circ \iota_2 \circ (\phi|_{\iota_2^{-1}(\tilde{U})})^{-1}) \circ (\tilde{\phi} \circ \iota_2 \circ (\phi|_{\iota_2^{-1}(\tilde{U})})^{-1})^{-1},$$

Both w_n^+ and w_n^- are orthogonal to w_j for $j = 1, \ldots, n-1$, and they have norm 1 with respect to $(\psi^{-1})^*(\tilde{g})$. Finally, note that $w_n^+, w_n^- \in H_n$ since

$$\psi(V \cap \iota_1(Q)) = \psi(V) \cap H_n,$$

so $w_n^+ = w_n^-$. We infer that the limit

$$w_n := \lim_{t\to 0} \frac{\psi \circ \tilde{\phi}^{-1}(\tilde{\phi}([j,q]) + te_n) - \psi([j,q])}{t},$$

exists and is equal to w_n^+, so the n-th partial derivative of $\psi \circ \tilde{\phi}^{-1}$ at $\phi(q)$ exists. It is easy to see that $\psi \circ \tilde{\phi}^{-1}$ is in fact C^1 on a neighbourhood of $\phi(q) = \tilde{\phi}([j,q])$. The partial derivatives $w_1, \ldots, w_n$ form a basis of $\mathbb{R}^n$, so by the inverse function theorem, there exists an open neighbourhood $\tilde{W} \subset \mathbb{R}^n$ of $\phi(q)$ such that $\psi \circ \tilde{\phi}^{-1}|_{\tilde{W}}$ is a diffeomorphism onto an open subset of $\mathbb{R}^n$. But then $W := \tilde{\phi}^{-1}(\tilde{W})$ is an open neighbourhood in $\mathrm{Double}(Q)$ of $[j,q]$ such that $W \subseteq \tilde{U} \cap V$ and $\psi \circ \tilde{\phi}^{-1}|_{\tilde{\phi}(W)}$ is a C^1-diffeomorphism. This establishes uniqueness of $\mathcal{D}$. ■

Now that we have defined a differentiable structure on the double of the configuration space, we would like to consider motion of a particle moving on $\mathrm{Double}(Q)$. The idea is that we can now lift motion on Q to motion on $\mathrm{Double}(Q)$. Of course, when one talks about motions of particles, one is naturally lead to ask what happens at the level of cotangent bundles. We will now describe a construction that is almost identical to the construction of the double of a manifold, namely the *double of the phase space* $M = T^*Q$. Indeed, this construction is the primary motivation for our study of fibre bundles over manifolds with boundary.

Let $(Q, \mathcal{D}_Q, g)$ be a Riemannian manifold with boundary, let M be its cotangent bundle, take two copies M_1 and M_2 of M, and form their disjoint union $M_1 \sqcup M_2$, as in our construction of $\mathrm{Double}(Q)$ above. Again, we leave the interiors of both copies untouched and glue the boundaries together, but now in a slightly different way.

Let $(p,q) \in \mathrm{Bound}(M)$, let (U, ϕ) be a chart of Q such that $q \in U$. Construct a chart $(\tilde{U}, \tilde{\phi})$ on $\mathrm{Double}(Q)$ as in the proof of the Theorem 5.1.4. Then $(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1)$ is a chart on Q. This chart induces a chart $((\iota_1 \circ \pi)^{-1}(\tilde{U}), \Phi)$ on M, where $\pi \colon M = T^*Q \to Q$ is the canonical projection. Moreover, let $r \colon \mathbb{R}^n \to \mathbb{R}^n$ be the reflection in the hyperplane orthogonal to the n-th standard basis vector, as defined in equation (5.1).

Now let $(\xi, x) := \Phi(p,q) \in \mathbb{R}^n \times (\mathbb{R}^{n-1} \times \{0\})$. We define an equivalence relation $\sim$ on $M_1 \sqcup M_2$ by

$$(1, (p,q)) = (1, \Phi^{-1}(\xi, x)) \sim (2, \Phi^{-1}(r(\xi), x)).$$

5.1.8 Lemma. *The above relation is independent of* $((\iota_1 \circ \pi)^{-1}(\tilde{U}), \Phi)$*, i.e. the equivalence class* $[(1, (\xi, x))]$ *only contains* $(1, (\xi, x))$ *and* $(2, (r(\xi), x))$*.*

Proof. Let (U_1, ϕ_1) and (U_2, ϕ_2) be two charts such that $q \in U_1 \cap U_2 \cap \mathrm{Bound}(Q)$, construct charts $(\tilde{U}_j, \tilde{\phi}_j)$ on Double(Q) such that $q \in \tilde{U}_1 \cap \tilde{U}_2$, and subsequently define charts $((\iota_1 \circ \pi)^{-1}(\tilde{U}_j), \Phi_j)$ on M. Now let $(\xi_j, x_j) := \Phi_j(p, q)$ for $j = 1, 2$. We must show that

$$\Phi_1^{-1}(r(\xi_1), x_1) = \Phi_2^{-1}(r(\xi_2), x_2).$$

By definition of the tangent space at $q \in Q$, we have

$$[(\iota_1^{-1}(\tilde{U}_2), \tilde{\phi}_2 \circ \iota_1), \tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}|'_{\tilde{\phi}_1(\tilde{U}_1 \cap \tilde{U}_2)}(x_1)(v)] = [(\iota_1^{-1}(\tilde{U}_1), \tilde{\phi}_1 \circ \iota_1), v],$$

for each $v \in \mathbb{R}^n$. Let A be the matrix associated to the linear map $\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}|'_{\tilde{\phi}_1(\tilde{U}_1 \cap \tilde{U}_2)}(x_1)$. From the definition of the charts $(\tilde{U}_j, \tilde{\phi}_j)$, $j = 1, 2$, and part (3) of Propostion 4.1.7, that both $\mathbb{R}^{n-1} \times \{0\}$ and the subspace of $\mathbb{R}^n$ spanned by e_n are invariant subspaces of this linear map, so A is of the form

$$A = \begin{pmatrix} & & & 0 \\ & \tilde{A} & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix},$$

where $\tilde{A}$ is an invertible $(n-1) \times (n-1)$-matrix. But then, using elementary linear algebra, we find that $\xi_2 = (A^{-1})^* \xi_1$, so

$$\begin{aligned} r(\xi_2) &= r((A^{-1})^* \xi_1) \\ &= \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} & & & 0 \\ & (\tilde{A}^{-1})^* & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \xi_1 \\ &= \begin{pmatrix} & & & 0 \\ & (\tilde{A}^{-1})^* & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \xi_1 \\ &= (A^{-1})^* r(\xi_1), \end{aligned}$$

and it follows that $\Phi_1^{-1}(r(\xi_1), x_1) = \Phi_2^{-1}(r(\xi_2), x_2)$, as desired. ∎

5.1.9 Definition. We call the space $M_1 \sqcup M_2/\sim$ with the topology associated to taking the disjoint union and the quotient the *double of the phase space of* Q, and write it as PDouble(Q).

The next step is to introduce a differentiable structure on PDouble(Q). As in the case of Double(Q), we can define continuous inclusions $\kappa_j \colon M \hookrightarrow \mathrm{PDouble}(Q)$, which are homeomorphisms onto their images. We will only show how to define a chart on $\kappa_1(\mathrm{Bound}(M)) = \kappa_2(\mathrm{Bound}(M))$.

Let $\pi \colon M = T^*Q \to Q$ be the canonical projection. Again, let $(p_0, q_0) \in \mathrm{Bound}(M)$, let (U, ϕ) be a chart of Q such that $q_0 \in U$, and construct a chart $(\tilde{U}, \tilde{\phi})$ on Double(Q)

as in Theorem 5.1.4, with $[1, q_0] \in \tilde{U}$. As before, the chart $(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1)$ on Q induces a chart $((\iota_1 \circ \pi)^{-1}(\tilde{U}), \Phi)$ on M, which allows us to define a chart $(\bigcup_{j=1,2} \kappa_j(\iota_1 \circ \pi)^{-1}(\tilde{U}), \tilde{\Phi})$ on $\mathrm{PDouble}(Q)$ as follows:
Let $(p, q) \in (\iota_1 \circ \pi)^{-1}(\tilde{U})$, let $(\xi, x) := \Phi(p, q) \in \mathbb{R}^n \times H_n$. Then we define

$$\tilde{\Phi} \colon (\iota_1 \circ \pi \circ P)^{-1}(\tilde{U}) \to \mathbb{R}^n \times \mathbb{R}^n$$

by

$$\tilde{\Phi}([j, (p, q)]) := \begin{cases} (\xi, x) & j = 1 \\ (r(\xi), r(x)) & j = 2 \end{cases} .$$

The following theorem relates $\mathrm{PDouble}(Q)$ to $T^*\mathrm{Double}(Q)$:

5.1.10 Theorem. *Let $(Q, \mathcal{D}_Q, g)$ be an n-dimensional Riemannian manifold with boundary and let $(\mathrm{Double}(Q), \mathcal{D}_{\mathrm{Double}(Q)})$ be the double of Q with the differentiable structure from Theorem 5.1.4. Then the charts of the form $(\bigcup_{j=1,2} \kappa_j(\iota_1 \circ \pi)^{-1}(\tilde{U}), \tilde{\Phi})$ defined above determine a differentiable structure $\mathcal{D}$ that is smooth on $\bigcup_{j=1,2} \kappa_j(\mathrm{Int}(M))$ and continuous on $\kappa_1(\mathrm{Bound}(M))$. The inclusions $\kappa_i \colon M \hookrightarrow \mathrm{PDouble}(Q)$ are smooth embeddings with respect to this differentiable structure. Moreover, the map $F \colon T^*\mathrm{Double}(Q) \to \mathrm{PDouble}(Q)$, given by*

$$(p, [j, q]) \mapsto [j, (p \circ T_q\iota_j, q)],$$

is an isomorphism of vector bundles over $\mathrm{Double}(Q)$.

Proof. (I) First, let us prove that the charts defined above form an atlas of $\mathrm{PDouble}(Q)$. Let $q_0 \in \mathrm{Bound}(Q)$, and pick two charts (U_j, ϕ_j), $j = 1, 2$, such that $q_0 \in U_1 \cap U_2$. Then construct the following charts:

- $(\tilde{U}_j, \tilde{\phi}_j)$ on $\mathrm{Double}(Q)$;
- $(\iota_1^{-1}(\tilde{U}_j), \tilde{\phi}_j \circ \iota_1)$ on Q;
- $((\iota_1 \circ \pi)^{-1}(\tilde{U}_j), \Phi_j)$ on M;
- $(\bigcup_{i=1,2} \kappa_i((\iota_1 \circ \pi)^{-1}(\tilde{U})), \tilde{\Phi}_j)$ on $\mathrm{PDouble}(Q)$.

We show that $\tilde{\Phi}_2 \circ \tilde{\Phi}_1^{-1}$ is smooth at each point of

$$\tilde{\Phi}_1 \left(\bigcup_{i=1,2} \kappa_i((\iota_1 \circ \pi)^{-1}(\tilde{U}_1 \cap \tilde{U}_2 \cap P^{-1}(\mathrm{Int}(Q)))) \right),$$

and continuous on

$$\tilde{\Phi}_1 \circ \kappa_1((\iota_1 \circ \pi)^{-1}(\tilde{U}_1 \cap \tilde{U}_2 \cap P^{-1}(\mathrm{Bound}(Q)))).$$

Suppose $(\xi, x) \in \tilde{\Phi}_1(\bigcup_{i=1,2} \kappa_i((\iota_1 \circ \pi)^{-1}(\tilde{U}_1 \cap \tilde{U}_2))) \cap (\mathbb{R}^n \times H_n)$. Then, we have

$$\tilde{\Phi}_2 \circ \tilde{\Phi}_1^{-1}(\xi, x) = \Phi_2 \circ \Phi_1(\xi, x) = (((A(x))^{-1})^*\xi, \tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}(x)),$$

where $A(x) := (\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1})'(x)$.

Now suppose $(\xi, x) \in \tilde{\Phi}_1(\bigcup_{j=1,2} \kappa_j((\iota_1 \circ \pi)^{-1}(\tilde{U}_1 \cap \tilde{U}_2))) \cap (\mathbb{R}^n \times -H_n)$. Define $(\xi', x') \in \mathbb{R}^n \times \mathbb{R}^n$ by

$$(r(\xi'), r(x')) := \tilde{\Phi}_2 \circ \tilde{\Phi}_1^{-1}(\xi, x) = \tilde{\Phi}_2([2, \Phi_1^{-1}(r(\xi), r(x))]),$$

so that $[2, \Phi_1^{-1}(r(\xi), r(x))] = [2, \Phi_2^{-1}(r(\xi'), r(x'))]$, and therefore

$$(\xi', x') = \Phi_2 \circ \Phi_1^{-1}(r(\xi), r(x)) = (((A(r(x)))^{-1})^* r(\xi), \tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}(r(x))).$$

If $q \in Q$ and $y \in H_n$ satisfy $\tilde{\phi}_1([1, q]) = y$, then $\tilde{\phi}_1([2, q]) = r(y)$. Thus we obtain $\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}(r(x)) = r(\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}(x))$, so

$$\tilde{\Phi}_2 \circ \tilde{\Phi}_1^{-1}(\xi, x) = (r(\xi'), r(x')) = (r(((A(r(x)))^{-1})^* r(\xi)), \tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}(x)).$$

The smoothness on $\tilde{\Phi}_1(\bigcup_{j=1,2} \kappa_j((\iota_1 \circ \pi)^{-1}(\tilde{U}_1 \cap \tilde{U}_2 \cap P^{-1}(\mathrm{Int}(Q)))))$ now follows from the fact that $\tilde{\phi}_2 \circ \tilde{\phi}_1^{-1}$ and $y \mapsto A(y)$ are smooth maps. The continuity on

$$\tilde{\Phi}_1 \circ \kappa_1((\iota_1 \circ \pi)^{-1}(\tilde{U}_1 \cap \tilde{U}_2 \cap P^{-1}(\mathrm{Bound}(Q))))$$

is a consequence of the fact that

$$r(((A(r(x)))^{-1})^* r(\xi)) = ((A(x))^{-1})^* \xi,$$

whenever $x \in \mathbb{R}^{n-1} \times \{0\}$, which we already showed in the proof of Lemma 5.1.8.

It is clear that the maps κ_j are smooth on $\mathrm{Int}(M)$. To see that they are smooth on the boundary, observe that

$$\tilde{\Phi}_1 \circ \kappa_j \circ \Phi_1^{-1} := \begin{cases} \mathrm{Id}_{\Phi_1((\iota_1 \circ \pi)^{-1}(\tilde{U}_1))} & j = 1 \\ r \times r|_{\Phi_1((\iota_1 \circ \pi)^{-1}(\tilde{U}_1))} & j = 2 \end{cases},$$

where $r \times r$ is the map on $\mathbb{R}^n \times \mathbb{R}^n$ sending (ξ, x) to $(r(\xi), r(x))$.

(II) Next, we will prove that F is an isomorphism of vector bundles. Let us begin by verifying that F is well defined. Let $(p_0, [j_1, q_0]) = (p_0, [j_2, q_0]) \in T^*\mathrm{Double}(Q)$. If $q_0 \in \mathrm{Int}(Q)$, then $j_1 = j_2$, and F is well defined at $(p_0, [j_1, q_0])$.

Now suppose $q_0 \in \mathrm{Bound}(Q)$, and let (U, ϕ) be a chart on Q such that $q_0 \in U$. Construct charts $(\tilde{U}, \tilde{\phi})$, $(\iota_1^{-1}(\tilde{U}), \tilde{\phi} \circ \iota_1)$, $((\iota_1 \circ \pi)^{-1}(\tilde{U}), \Phi)$ and $(\bigcup_{j=1,2} \kappa_j((\iota_1 \circ \pi)^{-1}(\tilde{U})), \tilde{\Phi})$ on $\mathrm{Double}(Q)$, Q, M and $\mathrm{PDouble}(Q)$ respectively. Moreover, let (V, Ψ) be the chart on $T^*\mathrm{Double}(Q)$ induced by $(\tilde{U}, \tilde{\phi})$. Then we have

$$\tilde{\Phi}([1, (p_0 \circ T_{q_0}\iota_1, q_0)]) = \Phi(p_0 \circ T_{q_0}\iota_1, q_0) =: (\xi_0, x_0),$$

where $x_0 = \phi(q_0) \in \mathbb{R}^{n-1} \times \{0\}$. On the other hand, we have

$$\tilde{\Phi}([2, (p_0 \circ T_{q_0}\iota_2, q_0)]) = (r(\xi_0'), r(x_0)) = (r(\xi_0'), x_0),$$

where ξ_0' is given by

$$(\xi_0', x_0) := \Phi(p_0 \circ T_{q_0}\iota_2, q_0) = \Phi(p_0 \circ T_{q_0}\iota_1 \circ ((T_{q_0}\iota_1)^{-1} \circ T_{q_0}\iota_2), q_0) = (r(\xi_0), x_0),$$

since $\tilde{\phi} \circ (\iota_1|_{\iota_1(Q)}^{-1} \circ \iota_2) \circ \tilde{\phi}^{-1} = r|_{\tilde{\phi}(\tilde{U})}$. We infer that

$$\tilde{\Phi}([1, (p_0 \circ T_{q_0}\iota_1, q_0)]) = (\xi_0, x_0) = \tilde{\Phi}([2, (p_0 \circ T_{q_0}\iota_2, q_0)]),$$

so F is indeed a well-defined map.

It is clear that F maps the fibre of $T^*\mathrm{Double}(Q)$ over $[j, q_0]$ to the fibre of $\mathrm{PDouble}(Q)$ over $[j, q_0]$, and that the restriction of F to the first fibre is a linear map. F is injective, because $T_{q_0}\iota_j$ is surjective for $j = 1, 2$, and F is surjective, because $T_{q_0}\iota_j$ is invertible for $j = 1, 2$.

It remains to be shown that F is a diffeomorphism. Since F is bijective, it suffices to show that F is a local diffeomorphism at each point $(p_0, [j, q_0])$ of $T^*\mathrm{Double}(Q)$. This is easy to see if $q_0 \in \mathrm{Int}(Q)$, so we shall only bother with the case $q_0 \in \mathrm{Bound}(Q)$. Note that

$$\tilde{\phi} \circ \iota_1 \circ (\tilde{\phi} \circ \iota_1)^{-1} = \mathrm{Id}|_{\tilde{\phi}(\iota_1^{-1}(\tilde{U}))}, \text{ and } \tilde{\phi} \circ \iota_2 \circ (\tilde{\phi} \circ \iota_1)^{-1} = r|_{\tilde{\phi}(\iota_1^{-1}(\tilde{U}))},$$

so, looking at the corresponding maps between cotangent bundles, we obtain

$$\Psi \circ (\iota_1)_* \circ \Phi^{-1} = \mathrm{Id}|_{\Phi((\iota_1 \circ \pi)^{-1}(\tilde{U}))}, \text{ and } \Psi \circ (\iota_2)_* \circ \Phi^{-1} = (r \times r)|_{\Phi((\iota_1 \circ \pi)^{-1}(\tilde{U}))},$$

where $(\iota_j)_* \colon M \to T^*\mathrm{Double}(Q)$ maps (p, q) to $(p \circ (T_q\iota_j)^{-1}, [j, q])$ for $j = 1, 2$. Let $(\xi, x) \in \Psi(V)$. Then there are two possibilities:

- $x \in H_n$: then $(\xi, x) \in \Phi((\iota_1 \circ \pi)^{-1}(\tilde{U}))$. Let $(p, q) := \Phi^{-1}(\xi, x)$. Then

  $$\Psi^{-1}(\xi, x) = \Psi^{-1} \circ (\Psi \circ (\iota_1)_* \circ \Phi^{-1})(\xi, x) = (p \circ (T_q\iota_1)^{-1}, [1, q]),$$

  so

  $$\tilde{\Phi} \circ F \circ \Psi^{-1}(\xi, x) = \tilde{\Phi} \circ F(p \circ (T_q\iota_1)^{-1}, [1, q]) = \tilde{\Phi}([1, (p, q)]) = \Phi(p, q) = (\xi, x).$$

- $x \in -H_n$: then $(r(\xi), r(x)) \in \Phi((\iota_1 \circ \pi)^{-1}(\tilde{U}))$. Let $(p, q) := \Phi^{-1}(r(\xi), r(x))$. Then

  $$\Psi^{-1}(\xi, x) = \Psi^{-1} \circ (\Psi \circ (\iota_2)_* \circ \Phi^{-1})(r(\xi), r(x)) = (p \circ (T_q\iota_2)^{-1}, [2, q]),$$

  so

  $$\tilde{\Phi} \circ F \circ \Psi^{-1}(\xi, x) = \tilde{\Phi} \circ F(p \circ (T_q\iota_2)^{-1}, [2, q]) = \tilde{\Phi}([2, (p, q)]) = (\xi, x).$$

Thus we obtain

$$\tilde{\Phi} \circ F \circ \Psi^{-1} = \mathrm{Id}_{\Psi(V)},$$

so F is indeed a local diffeomorphism at $(p_0, [j, q_0])$. ■

The above theorem has a nice physical interpretation: given a curve in $T^*\mathrm{Double}(Q)$ passing from $T^*\iota_1(Q)$ to $T^*\iota_2(Q)$, the local momentum coordinate associated to the direction perpendicular to the boundary will undergo a sign change, and the sign of the derivative associated to the local position coordinate will be reversed accordingly. This type of behaviour is characteristic for particles colliding elastically with the boundary, as the sign changes ensure that the curve satisfies the law of reflection.

5.1.2 Completeness

We have constructed the double of a Riemannian manifold with boundary, and we have seen that it is a manifold with empty boundary. Since the boundary was the main obstacle for geodesic completeness, we may now ask whether the Riemannian double is complete.

In general, the answer is no: one obvious counterexample is the open interval $[0, 1)$ with the canonical metric. Its Riemannian double may be identified with $]-1, 0] \cup [0, 1[=]-1, 1[$ with the canonical metric, which is not geodesically complete. This is, of course, a poor counterexample. By the Hopf-Rinow theorem (Theorem 4.3.19), a connected Riemannian manifold is complete as a metric space if and only if it is geodesically complete. Our original manifold with boundary $[0, 1[$ was not complete to begin with, so one cannot expect its double to be complete. Indeed, we have the following fact:

5.1.11 Proposition. *Let $(Q, \mathcal{D}_Q, g)$ be a connected, nonempty Riemannian manifold with boundary, and let $(\mathrm{Double}(Q), \mathcal{D}_{\mathrm{Double}}(Q), g_{\mathrm{Double}(Q)})$ be its Riemannian double.*

(1) The Riemannian double is connected if and only if $\mathrm{Bound}(Q)$ *is nonempty.*

(2) Assume that $\mathrm{Double}(Q)$ *is connected, and let d be its Riemannian distance function. Then the map $\tilde{d}\colon Q \times Q \to [0, \infty)$, given by*

$$\tilde{d}(q_1, q_2) := d([1, q_1], [1, q_2]),$$

is a distance function on Q, and the associated metric topology is the topology on Q.

(3) $(Q, \tilde{d})$ is a complete metric space if and only if $(\mathrm{Double}(Q), d)$ *is a complete metric space.*

Proof.
(1) We have shown in Proposition 5.1.2 that the maps $\iota_j\colon Q \hookrightarrow \mathrm{Double}(Q)$ are topological embeddings for $j = 1, 2$, so $\iota_1(Q)$ and $\iota_2(Q)$ are connected because Q is connected. If $\mathrm{Bound}(Q)$ is nonempty, then $\iota_1(Q)$ and $\iota_2(Q)$ have a nonempty intersection, so $\mathrm{Double}(Q) = \iota_1(Q) \cup \iota_2(Q)$ is connected. On the other hand, if $\mathrm{Bound}(Q)$ is empty, then $\iota_1(Q)$ and $\iota_2(Q)$ are both nonempty clopen subsets of $\mathrm{Double}(Q)$, so $\mathrm{Double}(Q)$ is disconnected.

(2) Note that d is well-defined, because $\mathrm{Double}(Q)$ is connected. From the fact that the map d is a distance function on $\mathrm{Double}(Q)$, it is readily seen that $\tilde{d}$ is a distance function on Q.

Let us call the topologies on Q and $\mathrm{Double}(Q)$ given by the manifold structure the *manifold topologies* of these spaces. It is clear that the map ι_1 is an isometry from $(Q, \tilde{d})$ to $(\iota_1(Q), d_{\iota_1(Q) \times \iota_1(Q)})$, so the map ι_1 is a homeomorphism from Q to $\iota_1(Q)$ with their respective metric topologies. It follows from Proposition 5.1.2 that ι_1 is also a homeomorphism from Q with its manifold topology to $\iota_1(Q)$ with the subspace topology induced by the manifold topology on $\mathrm{Double}(Q)$. Now, the manifold topology on $\mathrm{Double}(Q)$ is precisely the metric topology by Proposition 4.3.18. It is a well-known fact from general topology that the subspace topology induced by the metric topology on some subspace of a metric space is equal to the metric topology induced by the restriction of the distance function to that subspace, so the metric topology on $\iota_1(Q)$ is equal to the manifold topology on $\iota_1(Q)$, and consequently, the metric topology on Q is equal to the manifold topology on Q.

(3) First suppose that $(Q, \tilde{d})$ is complete. The subset $\iota_1(Q)$ is a closed subset of Q, so $(\iota_1(Q), d_{\iota_1(Q)\times\iota_1(Q)})$ is a complete metric space. The map $\iota_1(Q)$ is an isometry from $(Q, \tilde{d})$ to $(\iota_1(Q), d_{\iota_1(Q)\times\iota_1(Q)})$ by definition of $\tilde{d}$, so $(Q, \tilde{d})$ is complete.

To prove the converse statement, suppose $(Q, \tilde{d})$ is complete, and let $([j_k, q_k])_{k\in\mathbb{N}}$ be a Cauchy sequence in $(\mathrm{Double}(Q), d)$. Then there exist infinitely many $k \in \mathbb{N}$ such that $j_k = 1$, or there exist infinitely many $k \in \mathbb{N}$ such that $j_k = 2$. Without loss of generality, we may assume that there exist infinitely many $k \in \mathbb{N}$ such that $j_k = 1$, and in this way, we obtain a subsequence $([1, q_{k_l}])_{l\in\mathbb{N}}$ of $([j_k, q_k])_{k\in\mathbb{N}}$. This subsequence is a Cauchy sequence in $(\mathrm{Double}(Q), d)$ because $([j_k, q_k])_{k\in\mathbb{N}}$ is a Cauchy sequence, so the sequence $(q_{k_l})_{l\in\mathbb{N}}$ is a Cauchy sequence in $(Q, \tilde{d})$. Since $(Q, \tilde{d})$ was assumed to be complete, the sequence $(q_{k_l})_{l\in\mathbb{N}}$ has a limit $q \in Q$. It follows that $([1, q_{k_l}])_{l\in\mathbb{N}}$ converges to $[1, q]$, and since $([j_k, q_k])_{k\in\mathbb{N}}$ is Cauchy, it follows that $\lim_{k\to\infty}[j_k, q_k] = [1, q]$. Thus $(\mathrm{Double}(Q), d)$ is complete. ■

In general, one does not know the metric $\tilde{d}$ on Q explicitly. However, if Q is compact, then we know that $(Q, \tilde{d})$ is complete, because the metric topology is the same as the manifold topology on Q.

Even if the manifold with boundary Q with metric $\tilde{d}$ is complete, this does not guarantee that $\mathrm{Double}(Q)$ is geodesically complete. Indeed, given a pair $([j, q], v)$ with $[j, q] \in \mathrm{Double}(Q)$ and $v \in T_{[j,q]}\mathrm{Double}(Q)$, it is possible that there is no unique curve through $[j, q]$ with tangent vector v satisfying the geodesic equation.

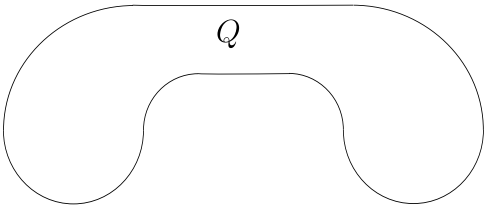
As a counterexample, consider the closed unit disc $D \subset \mathbb{R}^2$. This set is compact, because it is a closed and bounded subset of $\mathbb{R}^2$, so $(D, \tilde{d})$ is complete. The disc is of course a subset of another Riemannian manifold, namely $\mathbb{R}^2$ with the canonical metric. The geodesic corresponding to the pair $((1, 0), (0, 1))$ is the curve $t \mapsto (1, t)$, but the intersection of the image of this curve with D is just the point $(1, 0)$, so there is no way to properly define the geodesic in D corresponding to the pair $((1, 0), (0, 1))$. More generally, if the boundary of the manifold is 'convex', then one does not have local existence of the geodesic.

In contrast, consider a compact set that is locally 'concave', such as the subset of $\mathbb{R}^2$ depicted in figure 2.

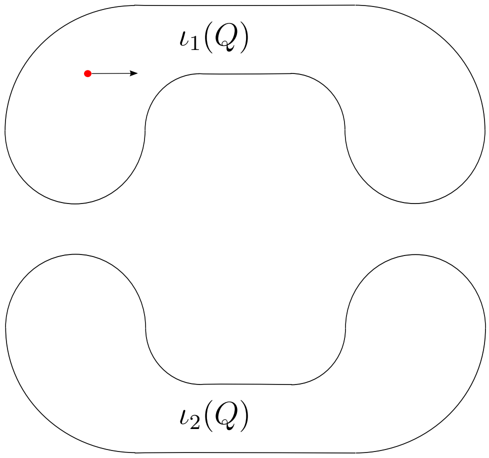
Here, there are multiple straight lines through a given point. In other words, we do not have local uniqueness of the geodesic.

The reason that geodesics exhibit this pathological behaviour is that the Riemannian double $(\mathrm{Double}(Q), \mathcal{D}_{\mathrm{Double}(Q)}, g_{\mathrm{Double}(Q)})$ is in general not a Riemannian manifold; in particular, the Riemannian metric $g_{\mathrm{Double}(Q)}$ is not smooth on the boundary, but merely continuous. Thus the Christoffel symbols (cf. Proposition 4.3.1) are not defined there, and consequently, the geodesic equation makes no sense.

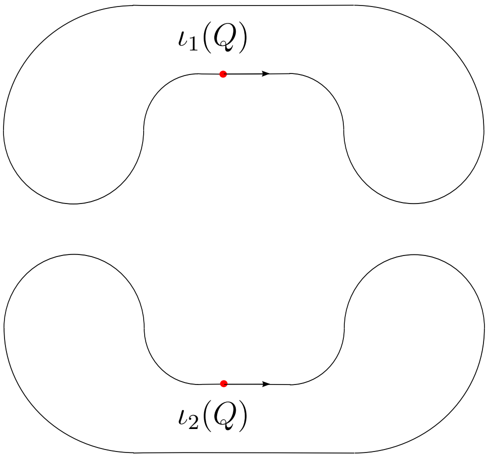
We will work towards formulating a condition on the metric at the boundary that will provide some degree of smoothness of $g_{\mathrm{Double}(Q)}$ on the boundary. Firstly, note that the notion of a geodesic as formulated in Definition 4.3.8 can be extended to Riemannian



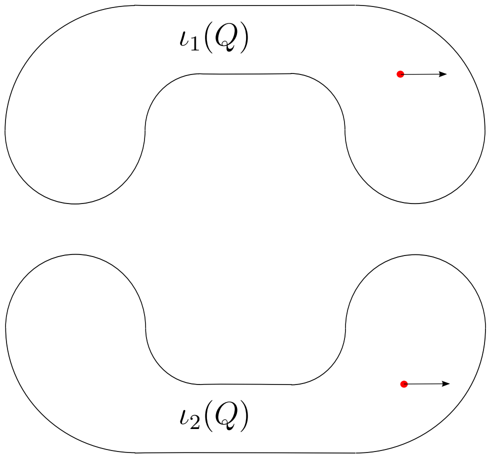
(a) The configuration space Q.



(b) The particle starts on $\iota_1(Q)$.



(c) When the particle is on the embedding of the boundary, it is on both $\iota_1(Q)$ and $\iota_2(Q)$.



(d) When the particle 'leaves' the boundary, there are two possible paths.

Figure 2: A particle, in these pictures represented by the red dot, moving on the double of some manifold Q with boundary.

manifolds with boundary using local extensions of the metric obtained by applying the method described in Lemma 5.1.3, and that this notion is independent of the particular extension.

Secondly, we have already seen in part (4) of Proposition 4.1.7 that the boundary of the n-dimensional manifold Q can be endowed with a canonical differentiable structure turning $\mathrm{Bound}(Q)$ into an $n-1$-dimensional manifold with empty boundary. The Riemannian metric on Q can be restricted to a Riemannian metric on $\mathrm{Bound}(Q)$. More precisely, the metric on $\mathrm{Bound}(Q)$ is the pull-back of the metric g under the inclusion map $\mathrm{Bound}(Q) \hookrightarrow Q$, so $\mathrm{Bound}(Q)$ can be given the structure of a Riemannian manifold.

5.1.12 Definition. Let $(Q, \mathcal{D}_Q, g)$ be a smooth, n-dimensional Riemannian manifold with boundary, let $(\mathrm{Bound}(Q), \mathcal{D}_{\mathrm{Bound}}(Q), g_{\mathrm{Bound}}(Q))$ be the boundary with the induced structure of a Riemannian manifold, and let $q_0 \in \mathrm{Bound}(Q)$. We say that the boundary of Q is *totally geodesic at* q_0 iff for each $v \in T_{q_0}\mathrm{Bound}(Q)$, there exists an $\varepsilon > 0$ such that the geodesic $\gamma\colon]-\varepsilon, \varepsilon[\to \mathrm{Bound}(Q)$ in $\mathrm{Bound}(Q)$ with $\gamma(0) = q_0$, and $\gamma'(v)$ is also a geodesic in Q.
We say that the boundary $\mathrm{Bound}(Q)$ of Q is *totally geodesic* iff it is totally geodesic at each point of $\mathrm{Bound}(Q)$.

5.1.13 Proposition. *Let $(Q, \mathcal{D}_Q, g)$ be a smooth, n-dimensional Riemannian manifold with boundary, let $(\mathrm{Double}(Q), \mathcal{D}_{\mathrm{Double}(Q)}, g_{\mathrm{Double}(Q)})$ be its Riemannian double, let $q_0 \in \mathrm{Bound}(Q)$, and suppose that $\mathrm{Bound}(Q)$ is totally geodesic at q_0. Then $g_{\mathrm{Double}(Q)}$ is C^2 at $[1, q_0]$.*

Proof. In view of our construction of the differentiable structure on $\mathrm{Double}(Q)$ that we outlined in the proof of Theorem 5.1.4, we can assume without loss of generality that Q is a convex open subset of H_n, and that $q_0 = 0$. Moreover, each curve of the form $t \mapsto (x, t) \in Q$, where $x \in \mathbb{R}^n$ and $t \geq 0$, is a geodesic. Finally, writing $g_{jk}(q) := g_q(e_j, e_k)$, where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$, we have $g_{jn}(q) = g_{nj}(q) = \delta_{jn}$ for $j = 1, 2, \ldots, n$ and for each $q \in \mathrm{Bound}(Q) = Q \cap (\mathbb{R}^{n-1} \times \{0\})$, and therefore also $g^{jn}(q) = g^{nj}(q) = \delta_{jn}$.

The double of Q can now be identified with the set $Q \cup r(Q)$, where $r\colon \mathbb{R}^n \to \mathbb{R}^n$ is the reflection in the hyperplane perpendicular to the n-th standard basis vector (with respect to the standard inner product on $\mathbb{R}^n$). The Riemannian metric g_{Double} is now an extension of g. Slightly abusing notation, its components will be denoted by $g_{jk}(q)$. Note that

$$g_{jk}(r(q)) = g_{jk}(q), \quad g_{jn}(r(q)) = -g_{jn}(q), \quad g_{nn}(r(q)) = g_{nn}(q),$$

for $j, k = 1, 2, \ldots, n-1$ and $q \in Q$. From the smoothness of g on Q, it is clear that each of the components g_{jk} of the Riemannian metric is partially differentiable up to any order in the directions $e_1, \ldots, e_{n-1}$ at 0. We must show that they are differentiable at 0 in the direction e_n, i.e. their normal derivatives exist up to second order.

Let $\gamma\colon]-\varepsilon, \varepsilon[\to \mathrm{Double}(Q) = Q \cup r(Q)$ be the curve $t \mapsto (0, 0, \ldots, 0, t)$, where $\varepsilon > 0$ is sufficiently small. Define the following one-sided derivatives for $j, k = 1, 2, \ldots, n$:

$$\frac{\partial g_{jk}}{\partial x_n^{\pm}}(0) := \lim_{t \to 0^{\pm}} \frac{g_{jk}(\gamma(t)) - g_{jk}(0)}{t}.$$

Note that these derivatives exist, since the Riemannian metric g on Q is smooth. The function $]-\varepsilon, \varepsilon[\ni t \mapsto g_{jk}(\gamma(t))$ is odd for $j = 1, 2, \ldots, n-1$ and $k = n$, or $j = n$ and

$k = 1, 2, \ldots, n-1$, so in these cases, $\frac{\partial g_{jk}}{\partial x_n^+}(0) = \frac{\partial g_{jk}}{\partial x_n^-}(0)$, which implies that $\frac{\partial g_{jk}}{\partial x_n}(0)$ exists. The function $]-\varepsilon, \varepsilon[\ni t \mapsto g_{jk}(\gamma(t))$ is even for $j, k = 1, 2 \ldots, n-1$, or $(j, k) = (n, n)$, so

$$\frac{\partial g_{jk}}{\partial x_n^+}(0) = -\frac{\partial g_{jk}}{\partial x_n^-}(0).$$

which implies that the limit

$$\lim_{t \to 0} \frac{g_{jk}(\gamma(t)) - g_{jk}(0)}{t},$$

exists and is equal to 0 if and only if one of the corresponding one-sided limits vanishes.

Let us begin by showing that $\frac{\partial g_{nn}}{\partial x_n^+}(0)$ vanishes. We will denote the canonical local frame on Q by $e_1, \ldots, e_n$. In view of Lemma 5.1.3, we can find a metric $\tilde{g}$ on an open neighbourhood of 0 that coincides with g where their domains overlap. We now have

$$\frac{\partial g_{nn}}{\partial x_n^+}(0) = \frac{\partial}{\partial x_n}(\tilde{g}(e_n, e_n))(0) = \tilde{g}_0(\nabla_{e_n} e_n(0), e_n(0)) + \tilde{g}_0(e_n(0), \nabla_{e_n} e_n(0)) = 0,$$

where ∇ is the Levi-Civita connection associated to $\tilde{g}$. Here, in the second step, we used the compatibility of this connection with the Riemannian metric and in the third step, we used the fact that γ is a geodesic such that $\gamma'(0) = e_n(0)$. Thus $\frac{\partial g_{nn}}{\partial x_n^+}(0) = 0$.

Next, we show that $\frac{\partial g_{jk}}{\partial x_n^+}(0) = 0$ for $j, k = 1, 2, \ldots, n-1$. Let $\tilde{g}$ be as in the previous paragraph. Suppose $\tilde{\gamma}\colon]-\tilde{\varepsilon}, \tilde{\varepsilon}[$ is a geodesic in Bound(Q) such that $\tilde{\gamma}(0) = 0$. Then it satisfies the geodesic equation for Bound(Q), so for $l = 1, 2, \ldots, n-1$, we have

$$\tilde{\gamma}_l''(0) + \sum_{j=1}^{n-1} \sum_{k=1}^{n-1} \Gamma_{jk}^l(0) \tilde{\gamma}_j'(0) \tilde{\gamma}_k'(0) = 0.$$

The manifold Q was assumed to be geodesically complete, so $\tilde{\gamma}$ also satisfies the geodesic equation for Q, which reads

$$\tilde{\gamma}_l''(0) + \sum_{j=1}^{n} \sum_{k=1}^{n} \Gamma_{jk}^l(0) \tilde{\gamma}_j'(0) \tilde{\gamma}_k'(0) = 0,$$

for $l = 1, \ldots, n$. The curve $\tilde{\gamma}$ satisfies both equations, and $\tilde{\gamma}_n(t) = 0$ for each $t \in]-\tilde{\varepsilon}, \tilde{\varepsilon}[$, so

$$\sum_{j=1}^{n} \sum_{k=1}^{n} \Gamma_{jk}^n(0) \tilde{\gamma}_j'(0) \tilde{\gamma}_k'(0) = 0.$$

Now, for each vector $v \in \mathbb{R}^n$, we can find a geodesic $\tilde{\gamma}$ in Bound(Q) such that $\tilde{\gamma}(0) = 0$ and $\tilde{\gamma}'(0) = v$, so $\Gamma_{jk}^n(0) = 0$ for $j, k = 1, \ldots, n-1$, since $\Gamma_{jk}^n = \Gamma_{kj}^n$. On the other hand, we have

$$\begin{aligned}
\Gamma_{jk}^n(0) &= \frac{1}{2} \sum_{l=1}^{n} \tilde{g}^{nl} \left(\frac{\partial \tilde{g}_{lj}}{\partial x_k}(0) + \frac{\partial \tilde{g}_{lk}}{\partial x_j}(0) - \frac{\partial \tilde{g}_{jk}}{\partial x_l}(0) \right) \\
&= \frac{1}{2} \left(\frac{\partial \tilde{g}_{nj}}{\partial x_k}(0) + \frac{\partial \tilde{g}_{nk}}{\partial x_j}(0) - \frac{\partial \tilde{g}_{jk}}{\partial x_n}(0) \right) \\
&= -\frac{1}{2} \frac{\partial \tilde{g}_{jk}}{\partial x_n}(0),
\end{aligned}$$

since $g^{nj}(q) = g_{nj}(q) = \delta_{nj}$ for each $q \in \mathrm{Bound}(Q)$. Thus

$$\frac{\partial g_{jk}}{\partial x_n^+}(0) = \frac{\partial \tilde{g}_{jk}}{\partial x_n}(0) = 0,$$

and it follows that the metric $g_{\mathrm{Double}(Q)}$ is C^1 at 0.

Next, we show that $g_{\mathrm{Double}(Q)}$ is actually C^2 at 0. Note that the mixed partial derivatives $\frac{\partial^2 g_{jk}}{\partial x_l \partial x_n}(0)$, with $j, k = 1, 2, \dots, n$ and $l = 1, 2, \dots, n-1$ exist, because the partial derivatives of the form $\frac{\partial \tilde{g}_{jk}}{\partial x_n}(q)$ exist for each $q \in \mathrm{Bound}(Q)$, and because the Riemannian metric g on Q is smooth. Therefore, it suffices to show that the partial derivatives $\frac{\partial^2 g_{jk}}{\partial x_n^2}(0)$ exist for $j, k = 1, 2, \dots, n$.

This is easily done if $j, k = 1, 2, \dots, n-1$ or $(j, k) = (n, n)$. Indeed, in that case, we observe that the function

$$]-\varepsilon, \varepsilon[\ni t \mapsto \frac{\partial \tilde{g}_{jk}}{\partial x_n}(\gamma(t)),$$

is odd, and repeat the argument that we used to show that $\frac{\partial \tilde{g}_{ln}}{\partial x_n}$ exists for $l = 1, 2, \dots, n-1$.

Now we show that $\frac{\partial^2 \tilde{g}_{jn}}{\partial x_n^2}(q) = 0$ for $j = 1, 2, \dots, n-1$ and for each $q \in \mathrm{Double}(Q)$. Clearly, this implies that $\frac{\partial^2 g_{jn}}{\partial x_n^2}(0)$ exists and is equal to 0. First observe that the argument showing that $\frac{\partial g_{nn}}{\partial x_n^+}(0) = 0$ can be easily modified to prove that $\frac{\partial g_{nn}}{\partial x_n}(q) = 0$ for each $q \in \mathrm{Double}(Q)$. Since $g_{nn}(q) = 1$ for each $q \in \mathrm{Bound}(Q)$, it follows that $g_{nn}(q) = 1$ for each $q \in \mathrm{Double}(Q)$, since Q is convex. Thus g_{nn} is constant, and we infer that

$$\frac{\partial g_{nn}}{\partial x_j}(q) = 0,$$

for each $q \in \mathrm{Double}(Q)$. On the other hand, we have

$$\frac{\partial g_{nn}}{\partial x_j}(q) = 2\tilde{g}_q(\nabla_{e_j} e_n(q), e_n(q)),$$

Moreover,

$$\nabla_{e_n} e_j(q) = \nabla_{e_j} e_n(q) + [e_n, e_j](q) = \nabla_{e_j} e_n(q),$$

so

$$\begin{aligned}\frac{\partial g_{jn}}{\partial x_n}(q) &= \tilde{g}_q(\nabla_{e_n} e_j(q), e_n(q)) + \tilde{g}_q(e_j(q), \nabla_{e_n} e_n(q)) = \tilde{g}_q(\nabla_{e_j} e_n(q), e_n(q)) \\ &= \frac{1}{2}\frac{\partial g_{nn}}{\partial x_j}(q) = 0.\end{aligned}$$

This proves the assertion, and we conclude that $g_{\mathrm{Double}(Q)}$ is C^2 at q_0. ■

5.1.14 Corollary. *Let* $(Q, \mathcal{D}_Q, g)$ *be a compact connected Riemannian manifold with nonempty totally geodesic boundary, and let* $(\mathrm{Double}(Q), \mathcal{D}_{\mathrm{Double}(Q)}, g_{\mathrm{Double}(Q)})$ *be its Riemannian double. Then* $\mathrm{Double}(Q)$ *is geodesically complete.*

Proof. The manifold Q is connected and its boundary is nonempty, so Double(Q) is connected by part (1) of Proposition 5.1.11. Thus we can define the Riemannian distance function d on Double(Q), as well as the corresponding distance function $\tilde{d}$ on Q. The metric space $(Q, \tilde{d})$ is compact by part (2) of Proposition 5.1.11, so it is a complete space. It follows from part (3) of the same proposition that $(\text{Double}(Q), d)$ is a complete metric space. Now Q has a totally geodesic boundary, so the metric $g_{\text{Double}(Q)}$ is C^2 by Proposition 5.1.13. Hence $(\text{Double}(Q), \mathcal{D}_{\text{Double}(Q)}, g_{\text{Double}(Q)})$ is a geodesically complete Riemannian manifold by the Hopf-Rinow theorem. ■

5.1.15 Example. In the following examples, we endow the subsets of $\mathbb{R}^n$ with the Riemannian metric inherited from the canonical metric on $\mathbb{R}^n$.

(1) If $I \subseteq \mathbb{R}$ is a bounded, closed interval, then its boundary consists of two points and is therefore trivially totally geodesic. Consequently, the double is geodesically complete. Of course, the double is isomorphic (as a Riemannian manifold) to a circle whose circumference is twice the length of I.

(2) If $I \subseteq \mathbb{R}$ is a closed half-line, say $I = [0, \infty[$, then its boundary is of course also totally geodesic. Although I is not compact, the double of I is isomorphic to $\mathbb{R}$ and hence geodesically complete.

(3) Let $Q := S^2 \cap H_3$, that is, Q is the closed northern hemisphere in $\mathbb{R}^3$. Its boundary is a great circle, and curves traversing great circles with constant velocity are known to be geodesics, so Q has a totally geodesic boundary. Consequently, its double is geodesically complete. Obviously, the double of Q can be identified with S^2.

5.2 Phase space as an orbifold

We shall now further investigate the relation between the configuration space of a particle and its phase space. Let $(Q, \mathcal{D}_Q, g)$ be a Riemannian manifold with boundary. We regard Q as the configuration space of some particle. One would expect that its phase space is $M = T^*Q$. However, if one wants the motion of the particle to be complete, then we have already seen that it is better to consider the motion of the particle on Q as the projection onto Q of some curve on Double(Q). This poses the question whether the same can be done for the associated curve in phase space. Note that, although there exists a canonical continuous, open map $P\colon \text{Double}(Q) \to Q$ projecting the double of Q onto Q, we do not have such a nice map $\text{PDouble}(Q) \to T^*Q$. Thus we must try something else.

Observe that the group $\mathbb{Z}_2 = \{\pm 1\}$ acts canonically on the space PDouble(Q): the nontrivial element -1 interchanges $[1, (p, q)]$ and $[2, (p, q)]$ for each $(p, q) \in T^*Q$. Let $M_{\text{collision}} := \text{PDouble}(Q)/\mathbb{Z}_2$ be the associated space of orbits. Then $M_{\text{collision}}$ is very similar to T^*Q, except that for each point $q \in \text{Bound}(Q)$ of the boundary, one 'loses' half of the cotangent space T_q^*Q. In particular, the resulting object is in general no longer a manifold, but belongs to a larger class of objects known as *orbifolds*, which, informally speaking, are topological spaces that are locally homeomorphic to $\mathbb{R}^n$ modulo a finite group action. For a more precise definition, we refer to [20, pp. 6-7].

Take for instance $Q = [0, 1]$, and identify T^*Q with $\mathbb{R} \times Q$. Then $[1, (p, q)]$ is identified with $[1, (-p, q)] = [2, (p, q)]$ for $q = 0, 1$. While studying the motion of a particle on $[0, 1]$ colliding elastically with the boundary, we saw that when a particle reaches the boundary, its momentum was ill-defined. The identification of $[1,(p,q)]$ with $[1,(-p,q)]$ resolves precisely this problem! In this way, we can regard $M_{\text{collision}}$ as the phase space of a particle with configuration space Q that elastically collides with the boundary of Q. Note that this does not just work for $Q=[0,1]$, but for any arbitrary configuration space Q.

In the case of $Q=[0,1]$, the action of $\mathbb{Z}_2$ defined above is not the only possible action. The configuration space possesses a symmetry, namely the reflection $r_{\frac{1}{2}}\colon [0,1]\to[0,1]$, given by $q\mapsto 1-q$, which induces a map $T^*Q\to T^*Q$ given by

$$(p,q)\mapsto (p\circ T_q r_{\frac{1}{2}}^{-1}, r_{\frac{1}{2}}(q)),$$

which corresponds to the map $\mathbb{R}\times Q\to\mathbb{R}\times Q$ given by

$$(p,q)\mapsto(-p,1-q).$$

The isometry $r_{\frac{1}{2}}$ gives rise to another action of $\mathbb{Z}_2$ on $\mathrm{PDouble}(Q)$, where -1 sends $[1,(p,q)]$ to $[2,(-p,1-q)]$ and $[2,(p,q)]$ to $[1,(-p,1-q)]$. Let M_{periodic} be the corresponding quotient space. Under the canonical projection $\mathrm{PDouble}(Q)\to M_{\text{periodic}}$, the point $[1,(p,0)]$ is identified with $[1,(p,1)]=[2,(-p,0)]$ for each $p\in\mathbb{R}$. But this identification reflects exactly the discontinuity that we saw in our study of periodic motion on the interval, so M_{periodic} can be understood as the phase space of a particle on Q exhibiting periodic motion.

However, not every type of dynamical behaviour corresponding to some self-adjoint extension of $-\frac{\hbar^2}{2m}\Delta$ as $\hbar\to 0$ can be obtained by modding out $\mathrm{PDouble}(Q)$ to some action of the group $\mathbb{Z}_2$ on that space in the way we did for $M_{\text{collision}}$ and M_{periodic}. Indeed, we shall now construct a counterexample.

Let $I_1 :=]a_1,b_1[$ and $I_2 :=]a_2,b_2[$ be two intervals, and assume that $b_1<a_2$. Let $I:=I_1\cup I_2$, let $I' :=]a_1+a_2, b_1+b_2[$, let $\alpha\colon \overline{I}\to\overline{I'}$ be the map given by

$$x\mapsto\begin{cases} x+a_2 & x\in\overline{I_1},\\ x+b_1 & x\in\overline{I_2},\end{cases}$$

and let $F\colon L^2(I')\to L^2(I)$ be the map given by $\psi\mapsto\psi\circ\alpha|_I$.

5.2.1 Lemma. *For each $m\in\mathbb{N}_0$, the restriction of F to $H^m(I')$ is a unitary map onto*

$$V_m := \{\psi\in H^m(I)\colon \psi^{(k)}(a_2)=\psi^{(k)}(b_1) \text{ for } k=0,1,\ldots,m-1\},$$

with respect to the Sobolev norm on both spaces. Furthermore, we have $F(\psi^{(k)})=F(\psi)^{(k)}$ for each $\psi\in H^m(I')$ and $k=0,1,\ldots,m-1$.

Proof. One readily verifies that F is a unitary map from $L^2(I')$ to $L^2(I)$. Now let $\psi\in H^1(I')$, and let $\varphi\in C_0^\infty(I)$. From the fact that F is unitary and that

$$\varphi'\circ\alpha^{-1}|_{I'}=(\varphi\circ\alpha^{-1}|_{I'})',$$

we infer that

$$\begin{aligned}\int_I F(\psi)\varphi'\,dx &= \int_{I'}\psi F^{-1}(\varphi')\,dx=\int_{I'}\psi F^{-1}(\varphi)'\,dx=-\int_{I'}\psi' F^{-1}(\varphi)\,dx\\ &= -\int_I F(\psi')\varphi\,dx,\end{aligned}$$

so $F(\psi) \in H^1(I)$, and $F(\psi)' = F(\psi')$. Since $\psi \in H^1(I') \subseteq C(\overline{I'})$, we have $F(\psi) \in V_1$.

Now let $\psi \in V_1$, and let $\varphi \in C_0^\infty(I')$. Using the same facts as in the previous paragraph, and applying the integration by parts formula (Lemma 1.2.5), we obtain

$$\begin{aligned}\int_{I'} F^{-1}(\psi)\varphi'\,dx = \int_I \psi F(\varphi)'\,dx &= \int_{a_1}^{b_1} \psi F(\varphi)'\,dx + \int_{a_2}^{b_2} \psi F(\varphi)'\,dx \\ &= -\int_{a_1}^{b_1} \psi' F(\varphi)\,dx + \psi(b_1)F(\varphi)(b_1) - \int_{a_2}^{b_2} \psi' F(\varphi)\,dx - \psi(a_2)F(\varphi)(a_2) \\ &= -\int_I \psi' F(\varphi)\,dx = -\int_{I'} F^{-1}(\psi')\varphi\,dx.\end{aligned}$$

Thus F^{-1} maps V_1 into $H^1(I')$, so the restriction of F to $H^1(I')$ is an injection with image V_1. It can now be shown with a straightforward induction argument that F restricted to $H^m(I')$ is an injection with image V_m, and that $F(\psi)^{(k)} = F(\psi^{(k)})$ for each $\psi \in H^m(I')$ and $k = 0, 1, \ldots, m-1$. Finally, using this fact and the fact that F is a unitary map, it is readily verified that $F|_{H^m(I')}$ is a unitary map onto V_m with respect to the Sobolev norms on both spaces. ∎

Let us employ this lemma to study the behaviour of a particular self-adjoint extension of $H = D^2$ on I:

5.2.2 Proposition. *Let*

$$V' := \{\phi \in H^2(I') \colon \phi(a_1 + a_2) = \phi(b_1 + b_2),\ \phi'(a_1 + a_2) = \phi'(b_1 + b_2)\},$$

be the domain of the realisation H_{periodic} of $H = D^2$ on I' corresponding to periodic boundary conditions, and let

$$V := \{\phi \in H^2(I) \colon \phi(a_2) = \phi(b_1), \phi'(a_2) = \phi'(b_1), \phi(a_1) = \phi(b_2), \phi'(a_1) = \phi'(b_2)\} \subset L^2(I).$$

Then V is the domain of a self-adjoint realisation $\widetilde{H}$ of $H = D^2$ on I. Furthermore, we have $F(V') = V$, $\widetilde{H}F = FH_{\text{periodic}}$, and

$$e^{-it\widetilde{H}}F = Fe^{-itH_{\text{periodic}}},$$

for each $t \in \mathbb{R}$.

Proof. Using Theorem 2.3.7 and Lemma 2.3.14, one readily verifies that V is the domain of a self-adjoint realisation $\widetilde{H}$ of $H = D^2$ on I. The equalities $F(V') = V$ and $\widetilde{H}F = FH_{\text{periodic}}$ are immediate consequences of Lemma 5.2.1.

To prove the final identity, let $(\phi_j)_{j\in\mathbb{N}}$ be an orthonormal basis of $L^2(I')$ consisting of eigenvectors of H_{periodic} with corresponding eigenvalues $(E_j)_{j\in\mathbb{N}}$. Because F is unitary and $\widetilde{H}F = FH_{\text{periodic}}$, the sequence $(F(\phi_j))_{j\in\mathbb{N}}$ is an orthonormal basis of $L^2(I)$ consisting of eigenvectors of $\widetilde{H}$ with corresponding eigenvalues $(E_j)_{j\in\mathbb{N}}$. Hence, if $\phi = \sum_{j\in\mathbb{N}} c_j\phi_j \in L^2(I')$, then for each $t \in \mathbb{R}$ we have

$$e^{-it\widetilde{H}}F\phi = \sum_{j\in\mathbb{N}} e^{-it\widetilde{H}}F(c_j\phi_j) = \sum_{j\in\mathbb{N}} e^{-itE_j}c_jF(\phi_j) = \sum_{j\in\mathbb{N}} Fe^{-itH_{\text{periodic}}}(c_j\phi_j) = Fe^{-itH_{\text{periodic}}}\phi,$$

which proves the identity. ∎

We infer that to obtain an appropriate model for the configuration space of the particle carrying out the classical motion corresponding to $\widetilde{H}$, we would have to remove the boundary points of I by identifying a_1 and b_1 with b_2 and a_2, respectively. If we set $Q := I$, and take two copies $Q_j := \{(j,q) : q \in Q\}$, $j = 1,2$ of Q, then we can obtain that space by defining an equivalence relation $\sim$ on $Q_1 \sqcup Q_2$ by setting

$$(1,a_1) \sim (2,b_2), \quad (2,a_1) \sim (1,b_2), \quad (1,a_2) \sim (2,b_1), \text{ and } (2,a_2) \sim (1,b_1),$$

taking the quotient $Q_1 \sqcup Q_2/\sim$ and subsequently taking the quotient of the cotangent bundle of this object with respect to the canonical $\mathbb{Z}_2$-action on that bundle. If I_1 and I_2 do not have the same length, we cannot obtain this space by modding out $\mathrm{PDouble}(Q)$ to an action of the group $\mathbb{Z}_2$ on $\mathrm{PDouble}(Q)$ involving some isometry of Q.

Thus the notion of the double of a manifold with boundary is not general enough to construct all of the phase spaces on which we can adequately describe the complete classical motion associated to a certain realisation of the Hamiltonian. Our discussion in the previous paragraph suggests a way to generalise the construction of the double of a manifold with boundary. The boundary of a smooth Riemannian manifold $(Q, \mathcal{D}_Q, g)$ with boundary is again a smooth manifold by part (4) of Proposition 4.1.7, and the pullback of g under the inclusion map is a Riemannian metric, turning $\mathrm{Bound}(Q)$ in a Riemannian manifold. Now suppose that f is an isometry of this Riemannian manifold. Then we can define a relation $\sim$ on $Q_1 \sqcup Q_2$ analogous to the one used in our construction of $\mathrm{Double}(Q)$, by $(1,q) \sim (2, f(q))$ for each $q \in \mathrm{Bound}(Q)$. Since we want to define a $\mathbb{Z}_2$-action on the quotient $Q \cup_f Q := Q_1 \sqcup Q_2/\sim$, we demand that $f^2 = \mathrm{Id}_{\mathrm{Bound}(Q)}$.

Using the fact that f is an isometry, one can now construct a differentiable structure on $Q \cup_f Q$ in a way analogous to the construction of the differentiable structure on $\mathrm{Double}(Q)$ described in Theorem 5.1.4. We leave it to the reader to formulate and prove Theorems 5.1.4 and 5.1.7, and Propositions 5.1.11 and 5.1.13 for $Q \cup_f Q$ instead of $\mathrm{Double}(Q)$. We now define the phase space corresponding to f as $T^*(Q \cup_f Q)$ modded out to the canonical $\mathbb{Z}_2$-action on this cotangent bundle.

It is clear that if we take $f = \mathrm{Id}_{\mathrm{Bound}(Q)}$, then $Q \cup_f Q = \mathrm{Double}(Q)$. If Q is a union of closed intervals, then f is a permutation of order 2 of a finite set. Equivalently, f is a product of disjoint transpositions. If Q is a single closed interval, then the transposition of the two endpoints corresponds to periodic boundary conditions. In our discussion of the realisation $\widetilde{H}$ of $H = D^2$ on the interval I, the constructed space would correspond to $Q \cup_f Q$, where f is the map that sends a_1 to b_2 and vice versa, and a_2 to b_1 and vice versa.

Of course, we can try other permutations as well:

5.2.3 Example.

(1) Let f be the map that transposes a_1 and a_2, and b_1 and b_2. Particles moving in the corresponding phase space on I_1 in the positive direction would appear at b_2 the moment that they reach b_1, and continue to move across I_2 with the same magnitude of velocity, but in the negative direction. Upon reaching a_2, they would appear at a_1 and again move in the positive direction across I_1. In this case, there is a self-adjoint realisation that generates this type of dynamics as well, with domain

$$\{\phi \in H^2(I) : \phi(a_1) = \phi(a_2),\ \phi'(a_1) = -\phi'(a_2),\ \phi(b_1) = \phi(b_2),\ \phi'(b_1) = -\phi'(b_2)\}.$$

One can relate the unitary evolution group to the unitary evolution group of H_{periodic} on I' in a way much like the one presented in Lemma 5.2.1 and Proposition 5.2.2, except one has to replace the map α with the map $\beta\colon \overline{I} \to \overline{I'}$ given by

$$x \mapsto \begin{cases} a_2 + x & x \in \overline{I_1}, \\ b_1 + b_2 + a_2 - x & x \in \overline{I_2}, \end{cases}$$

and modify the map F accordingly.

(2) Let f be the map that transposes a_2 and b_1, and fixes a_1 and b_2. Particles moving in the corresponding phase space on I_1 in the positive direction would appear at a_2 the moment that they reach b_1, and continue to move across I_2 until they reach b_2, where they collide elastically with the boundary. Similarly, if a particle was moving in the negative direction towards a_1, then it would also collide elastically with the boundary upon arrival at a_1. One of the self-adjoint realisations that generate this type of dynamics has domain

$$\{\phi \in H^2(I)\colon \phi(a_2) = \phi(b_1),\ \phi'(a_2) = \phi'(b_1),\ \phi'(a_1) = \phi'(b_2) = 0\}.$$

One can relate the unitary evolution group to the unitary evolution group corresponding to the extension of H on I' associated to Neumann boundary conditions in the same way as in Lemma 5.2.1 and Proposition 5.2.2.

(3) If f sends a_1 to b_1 and vice versa, and leaves a_2 and b_2 fixed, then particles on I_1 will move periodically across that interval, whereas particles on I_2 will collide elastically with the boundary. There is also a self-adjoint realisation $\widetilde{H}$ of the Hamiltonian on I that realises the quantum mechanical equivalent of this behaviour, with domain

$$\{\phi \in H^2(I)\colon \phi(a_1) = \phi(b_1),\ \phi'(a_1) = \phi'(b_1),\ \phi'(a_2) = \phi'(b_2) = 0\}.$$

To see that this self-adjoint extension does indeed generate the described type of dynamics, we cannot use the same approach as in the previous examples, however. Instead, note that the domain of $\widetilde{H}$ is the direct sum of two domains of self-adjoint extensions $\widetilde{H_1}$ and $\widetilde{H_2}$ of H on I_1 and I_2, respectively. The realisation $\widetilde{H_1}$ corresponds to periodic boundary conditions, whereas $\widetilde{H_2}$ corresponds to Neumann boundary conditions. For $k = 1, 2$, let $(\phi_{k,j})_{j\in\mathbb{N}}$ be orthonormal bases of $L^2(I_k)$ of eigenvectors of $\widetilde{H_k}$. Then $(\phi_{k,j})_{(k,j)\in\{1,2\}\times\mathbb{N}}$ is an orthonormal basis of $L^2(I)$ eigenvectors of $\widetilde{H}$, where we have identified $\phi_{k,j} \in L^2(I_k)$ with its extension by zero to I. It follows that for each $t \in \mathbb{R}$, we have

$$e^{-it\widetilde{H}} = e^{-it\widetilde{H_1}} p_1 + e^{-it\widetilde{H_2}} p_2,$$

where p_k is the orthogonal projection onto $L^2(I_k)$ for $k = 1, 2$.

Similarly, one can examine all other possible permutations and find corresponding self-adjoint realisations, and do the same thing for configuration spaces consisting of any finite number of disjoint closed intervals. The following theorem shows how to define these self-adjoint realisations:

5.2.4 Theorem. *Let $m \in \mathbb{N}$, fix $a_j, b_j \in \mathbb{R}$ with $a_j < b_j$ for $j = 1, 2, \ldots, m$, and $b_j < a_{j+1}$ for $j = 1, 2, \ldots, m-1$, and let $I := \bigcup_{j=1}^{m}]a_j, b_j[$. Moreover, let*

$$\underline{2m} := \{k \in \mathbb{N} \colon k \leq 2m\},$$

let h be the unique monotone increasing bijection from $\underline{2m}$ to ∂I, and let f be a permutation of ∂I of order 2. Then the extension H_f of $H = D^2$ on I with domain $\mathcal{D}(H_f)$ given by

$$\{\phi \in H^2(I) \colon (-1)^{kj} \phi^{(k)}(h(j)) = (-1)^{k(h^{-1} \circ f \circ h(j)+1)} \phi^{(k)}(f \circ h(j)) \text{ for } j \in \underline{2m} \text{ and } k = 0, 1\},$$

is a self-adjoint realisation of H.

Proof. Let (V_H, ω_H) be the endpoint space of H, and define the linear isomorphism $\varrho_H \colon V_H \to \mathbb{C}^{4m}$ by

$$(\phi, H_{\max}\phi) \mapsto (\phi(a_1), \phi'(a_1), \phi(b_1), \phi'(b_1), \ldots, \phi(a_m), \phi'(a_m), \phi(b_m), \phi'(b_m)).$$

Furthermore, let

$$M := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

define the $4m \times 4m$-matrix B by

$$B := \begin{pmatrix} M & & & & & & \\ & -M & & & & & \\ & & M & & & & \\ & & & -M & & & \\ & & & & \ddots & & \\ & & & & & M & \\ & & & & & & -M \end{pmatrix},$$

and define the complex symplectic form ω on $\mathbb{C}^{4m}$ by

$$\omega(\mathbf{c}, \mathbf{d}) := \mathbf{c}^* B \mathbf{d}.$$

From Theorem 2.3.7 and Lemma 2.3.14, we infer that for each $u, v \in V_H$ we have

$$\omega(\varrho_H(u), \varrho_H(v)) = \omega_H(u, v),$$

so there is a bijective correspondence between Lagrangian subspaces U of $(\mathbb{C}^{4m}, \omega)$ and self-adjoint extensions H_U of H given by

$$\mathcal{D}(H_U) := \{\phi \in H^2(I) \colon \varrho_H \circ p_{V_H}(\phi, H_{\max}\phi) \in U\},$$

where p_{V_H} is the orthogonal projection of $L^2(I)^2$ onto V_H.

Now let $e_1, e_2, \ldots, e_{4m}$ be the standard basis of $\mathbb{C}^{4m}$. For each $j \in \underline{2m}$ and $k = 0, 1$, let

$$v_{j,k} := (-1)^{kj} e_{2(j-1)+k+1} - (-1)^{k(h^{-1} \circ f \circ h(j)+1)} e_{2(h^{-1} \circ f \circ h(j)-1)+k+1},$$

let

$$S_f := \{v_{j,k} \colon j \in \underline{2m},\ k = 0,1\},$$

and let U_f be the subspace of $\mathbb{C}^{4m}$ spanned by S_f, so that $H_{U_f} = H_f$. Then for each $j \in \underline{2m}$ and each $k = 0,1$ we have

$$(*) \qquad v_{j,k} = (-1)^{k+1} v_{h^{-1}\circ f\circ h(j),k},$$

and if $h(j)$ is a fixed point of f, then $v_{j,0} = 0$. It follows that $\dim(U_f) = 2m$. Therefore, to prove that U_f is Lagrangian, it suffices to show that U_f is isotropic. Since ω is sesquilinear, it suffices to show that for each $j_1, j_2 \in \underline{2m}$ and each $k_1, k_2 = 0,1$ we have $\omega(v_{j_1,k_1}, v_{j_2,k_2}) = 0$. If $k_1 = k_2$, then this is obviously true, so let us assume that $k_1 = 0$ and $k_2 = 1$. Then the only nontrivial cases left to check are $j_1 = j_2$ and $j_1 = h^{-1}\circ f\circ h(j_2)$. From equation $(*)$ we deduce that it suffices to investigate the case $j_1 = j_2 =: j$. Now there are three possibilities:

- $j = h^{-1} \circ f \circ h(j)$: then $v_{j,0} = 0$, so $\omega(v_{j,0}, v_{j,1}) = 0$;
- $j - h^{-1} \circ f \circ h(j) \in 2\mathbb{Z}\backslash\{0\}$: then

$$\begin{aligned} v_{j,0} &= e_{2(j-1)+1} - e_{2(h^{-1}\circ f\circ h(j)-1)+1}, \text{ and} \\ (-1)^j v_{j,1} &= e_{2(j-1)+2} + e_{2(h^{-1}\circ f\circ h(j)-1)+2}, \end{aligned}$$

  so

$$\begin{aligned} \omega(v_{j,0}, v_{j,1}) &= (-1)^j(\omega(e_{2(j-1)+1}, e_{2(j-1)+2}) - \omega(e_{2(h^{-1}\circ f\circ h(j)-1)+1}, e_{2(h^{-1}\circ f\circ h(j)-1)+2})) \\ &= (-1)^j((-1)^j - (-1)^j) = 0; \end{aligned}$$

- $j - h^{-1} \circ f \circ h(j) \in 2\mathbb{Z} + 1$: then

$$\begin{aligned} v_{j,0} &= e_{2(j-1)+1} - e_{2(h^{-1}\circ f\circ h(j)-1)+1}, \text{ and} \\ (-1)^j v_{j,1} &= e_{2(j-1)+2} - e_{2(h^{-1}\circ f\circ h(j)-1)+2}, \end{aligned}$$

  so

$$\begin{aligned} \omega(v_{j,0}, v_{j,1}) &= (-1)^j(\omega(e_{2(j-1)+1}, e_{2(j-1)+2}) + \omega(e_{2(h^{-1}\circ f\circ h(j)-1)+1}, e_{2(h^{-1}\circ f\circ h(j)-1)+2})) \\ &= (-1)^j((-1)^j + (-1)^{j+1}) = 0. \end{aligned}$$

We conclude that U_f is isotropic, hence U_f is Lagrangian and H_f is a self-adjoint realisation of H. ■

5.2.5 Remark. Of course, the idea behind this theorem is that the unitary evolution group associated to H_f corresponds to the complete classical motion on the orbifold $T^*(\overline{I} \cup_f \overline{I})/\mathbb{Z}_2$. Indeed, the three possibilities distinguished in the final part of the proof correspond to the following three cases, respectively:

- If a_l (b_l) is a fixed point of f for some $l \in \mathbb{N}$ with $l \leq m$, then we impose the Neumann boundary condition $\phi'(a_l) = 0$ ($\phi'(b_l) = 0$) at this point. As a result, particles moving towards this point will collide elastically with the boundary when they reach it.

- If f transposes a_{l_1} and a_{l_2} (b_{l_1} and b_{l_2}), then we impose the boundary conditions $\phi(a_{l_1}) = \phi(a_{l_2})$ and $\phi'(a_{l_1}) = -\phi'(a_{l_2})$ (likewise for b_{l_1} and b_{l_2}), so that any particle reaching either a_{l_1} or a_{l_2} will appear at the other point and continue to move in the opposite direction.
- If f transposes a_{l_1} and b_{l_2}, then we impose the boundary conditions $\phi(a_{l_1}) = \phi(b_{l_2})$ and $\phi'(a_{l_1}) = \phi'(b_{l_2})$, so that any particle reaching either a_{l_1} or b_{l_2} will appear at the other point and continue to move in the same direction.

Using the same methods as in Example 5.2.3, one can check for each isometry f that the above boundary conditions yield the described dynamical behaviour.

Up to now, our investigation of the relation between the orbifolds of the form $T^*(Q \cup_f Q)/\mathbb{Z}_2$ and the self-adjoint extensions of the test Hamiltonian has been motivated by physical considerations only. It is worth noting that from a mathematical point of view, there is also a reason why this relation is plausible. We have seen in the proof of part (5) of Proposition 2.2.2 that, given a hermitian operator T on a Hilbert space $\mathcal{H}$, then

$$\mathcal{H}^2 = \mathcal{G}(T_{\min}) \oplus V_T \oplus J(\mathcal{G}(T_{\min})),$$

and that self-adjoint extensions T_U of T correspond to decompositions of V_T of the form

$$V_T = U \oplus J(U),$$

where U is a Lagrangian subspace of (V_T, ω_T). Now note that the map iJ satisfies $(iJ)^2 = \mathrm{Id}_{\mathcal{H}^2}$, so it defines a $\mathbb{Z}_2$-action on $\mathcal{H}^2$ that interchanges the subspace $\mathcal{G}(T_U) = \mathcal{G}(T_{\min}) \oplus U$ and its image under J.

Now let us return to the case $\mathcal{H} = L^2(I)$ and $T = H = D^2$ the test Hamiltonian on a union I of bounded open intervals. Since $\mathcal{G}(H_{\min})$ is the closure of $C_0^\infty(I)$ with respect to the inner product on $L^2(I)^2$, we think of $\mathcal{G}(H_{\min})$ as the part of $\mathcal{G}(T_U)$ corresponding to $I = \mathrm{Int}(\overline{I})$. In addition, we have seen that there is a natural relation between V_H and the boundary $\partial I = \mathrm{Bound}(\overline{I})$, suggesting that the choice of a Lagrangian subspace $U \subseteq V_H$ is related to the choice of some bijection of ∂I of order 2. These observations reinforce the idea that orbifolds of the form $T^*(Q \cup_f Q)/\mathbb{Z}_2$ are appropriate models of phase space.

Conclusion and further research

Let us examine what can be said regarding the two problems formulated in the Introduction, beginning with the matter of completeness of the motion. Although we have only studied a couple of the unitary evolution groups associated to the self-adjoint extensions of an operator, when applied to our modified coherent states, none of the resulting states displayed any form of incomplete motion. Instead, the dynamical behaviour was in some sense an approximation to the behaviour of a classical particle moving on our modified versions of phase space. Therefore, it seems more likely that the incompleteness of the classical motion of a free particle is only apparent, rather than being a phenomenon emerging in the classical limit.

As for the second problem, the non-uniqueness of physics, we have seen that different self-adjoint extensions may indeed correspond to different types of physics. Even so, not every self-adjoint extension produces completely different behaviour; in our MATLAB simulations, we have only observed two radically different types of dynamics, and both of them had classical analogues with different modified phase spaces; see our discussion in section 5.2.

This of course prompts two questions. Firstly, *does every self-adjoint extension of the Hamiltonian yield a time evolution that in the limit* $\hbar \to 0$ *can be understood as a classical motion on some orbifold* $T^*(Q \cup_f Q)/\mathbb{Z}_2$ *for some isometry* f *of order* 2 *of* $\mathrm{Bound}(Q)$*?* According to Theorem 5.2.4 and the subsequent remark, if Q consists of finitely many bounded intervals, then the answer to the converse of the above question is yes.

If the first question is answered postively, then the second question that comes up is: *which self-adjoint extensions exhibit the same limiting behaviour as* $\hbar \to 0$*, and what determines their limiting behaviour?* We believe that there are two different approaches that one could take to answering this question. One would be to study the set of Lagrangian subspaces of the endpoint space (V_H, ω_H) of the Hamiltonian, and see whether there exists a 'natural' partition that is in bijective correspondence with the set of isometries of order 2 of the boundary of Q. Another could be to examine the orthonormal bases $(\phi)_{j=1}^{\infty}$ of eigenfunctions associated to the self-adjoint extensions, and see whether there is a correlation between the eigenfunctions ϕ_j corresponding to high values of j, and the associated classical behaviour of the extension. In either case, one would most likely have to resort to numerical simulations to study the problem.

Many other question still remain open. For instance, is it possible to give a complete characterisation of the self-adjoint extensions of H when Q is a bounded subset of dimension $n \geq 2$? Furthermore, we have only shown how to define complete classical motion on Q when Q is a smooth Riemannian manifold with totally geodesic boundary. The case in which the boundary of Q is not totally geodesic, or the more general case where Q has corners, have been left untouched. In the latter case, Q will no longer have a Riemannian double that is a smooth manifold with empty boundary, making it much harder if not impossible to define classical motion properly. In the former case, there is some hope of salvaging some sort of classical mechanics on $T^*(Q \cup_f Q)/\mathbb{Z}_2$, although one has to abandon the notion of point particle. One possible way to achieve this is outlined in the appendix.

Appendix: the Koopman-von Neumann formalism

Let $(Q, \mathcal{D}_Q, g)$ be a Riemannian manifold with boundary. We have seen in Proposition 5.1.13 that the metric of its Riemannian double is C^2 if the manifold has a totally geodesic boundary. This is a useful result, because it implies that the geodesic equation has a unique solution given an initial position and velocity, and thus has a well-defined Hamiltonian flow. Requiring that Q has a totally geodesic boundary is, however, a strong condition, especially if one realises that the set of points in T^*Q for which the Hamiltonian flow is potentially ill-defined, is rather 'small'. For example, if Q is the closed unit disk in $\mathbb{R}^2$, then we have already seen that particles at the boundary whose velocity is parallel to the boundary do not have a well-defined trajectory; but particles with a different position or velocity could be given a well-defined trajectory, namely by demanding that they collide elastically with the boundary.

This leads us to ask whether there is a way of defining classical motion that is unaffected by this small set of 'bad points'. In the disk example, the set of bad points has measure zero, so objects that are defined up to a set of measure zero, such as elements of $L^1_{\mathrm{loc}}(Q)$, may be suited for setting up a theory that provides us with some notion of classical mechanics on Q. It is here that the Koopman-von Neumann formalism (cf. [15, Section X.14]) becomes interesting. The starting point of this theory is the differential equation

$$\frac{df}{dt} = \{H, f\},$$

with initial condition $f(0) = f_0$ on phase space, where H is the classical Hamiltonian, f is some classical observable of a particle, and $\{H, f\}$ their Poisson bracket. Assume for the moment that $(Q, \mathcal{D}_Q, g)$ is a smooth n-dimensional Riemannian manifold with boundary. We can now take the same functional analytic approach that we used to examine the Schrödinger equation: First, we multiply the above equation by i. Now, let L be the differential operator on $L^2(\Omega)$ with domain $C_0^\infty(T^*Q)$, given by

$$L\phi := \{H, \phi\} = \sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\frac{\partial \phi}{\partial q_j} - \frac{\partial H}{\partial q_j}\frac{\partial \phi}{\partial p_j}.$$

The operator L is called the *Liouville operator*. Notice that iL is formally self-adjoint. By Corollary 2.2.13, it has a self-adjoint realisation $i\widetilde{L}$. By Theorem 1.5.3, there exists a unitary evolution group $(U(t))_{t\in\mathbb{R}}$ with infinitesimal generator $i\widetilde{L}$. Now suppose that $\phi_0 \in \mathcal{D}(i\widetilde{L})$. Then, we have

$$\frac{d}{dt}(U(t)\phi_0) = -i(i\widetilde{L})(U(t)\phi_0) = \widetilde{L}(U(t)\phi_0),$$

where the derivative is taken with respect to the norm on $L^2(Q)$, hence the solution to our differential equation is $\mathbb{R} \ni t \mapsto \phi_t := U(t)\phi_0 \in \mathcal{D}(i\widetilde{L})$. If we now regard ϕ_0 as a mass distribution or probability density function, then the time evolution of ϕ_0 is given by $t \mapsto \phi_t$, which is well-defined and complete!

What remains of this idea if the smooth Riemannian manifold with boundary $(Q, \mathcal{D}_Q, g)$ is replaced by $Q \cup_f Q$, where f is an isometry of the boundary? The main problem is the derivative $\frac{\partial H}{\partial q_j}$ in the definition of the Liouville operator. The classical Hamiltonian H is defined in terms of the metric, which, as we know, is in general not differentiable on the embedding of the boundary in $Q \cup_f Q$, but merely continuous. However, this set is a set of measure zero, so we may define the function $L\phi$ at the boundary in any way we see fit, since the Liouville operator is an operator on $L^2(T^*(Q \cup_f Q))$, whose elements are sets of square integrable functions that are equal almost everywhere on $T^*(Q \cup_f Q)$.

Thus we can employ the Koopman-von Neumann formalism to examine the time evolution of an element $\phi_0 \in L^2(T^*(Q \cup_f Q))$. Subsequently, we can define a version of complete classical mechanics on $T^*(Q \cup_f Q)/\mathbb{Z}_2$. Indeed, consider the map

$$P\colon L^2(T^*(Q \cup_f Q)) \to L^2(T^*(Q \cup_f Q)/\mathbb{Z}_2)$$

that sends a function $\phi \in L^2(T^*(Q \cup_f Q))$ to the function $T^*(Q \cup_f Q)/\mathbb{Z}_2 \to \mathbb{C}$,

$$[p, [1, q]] \mapsto \phi(p, [1, q]) + \phi(p, [2, q]).$$

Then, given a function $\phi_0 \in T^*(Q \cup_f Q)/\mathbb{Z}_2$, we can obtain a function $\psi_0 \in L^2(T^*(Q \cup_f Q))$ that satisfies $P(\psi_0) = \phi_0$, for example, by setting

$$\psi_0(p, [j, q]) := \begin{cases} \phi_0([p, [1, q]]) & j = 1, \\ 0 & j = 2, \end{cases}$$

study its time-evolution to obtain a function $\mathbb{R} \ni t \mapsto \psi_t \in L^2(T^*(Q \cup_f Q))$ and finally compose this function with the projection onto $L^2(T^*(Q \cup_f Q)/\mathbb{Z}_2)$, i.e., $\phi_t := P(\psi_t)$ for each $t \in \mathbb{R}$.

Nevertheless, there is one major drawback. We have shown that the operator iL has a self-adjoint extension, but we do not know whether it is unique or not, and (to the best of the author's knowledge) no results regarding this specific operator have been established so far. Moreover, we could have used the Koopman-von Neumann formalism to immediately construct a complete version of classical mechanics on T^*Q, without taking the detour of constructing the space $Q \cup_f Q$. In that case, however, iL most likely has many more self-adjoint extensions, and one would probably encounter problems similar to the ones that we have been trying to solve for the operator(!) H in this thesis.

References

[1] S. T. Ali, J.-P. Antoine, J.-P. Gazeau, and U. Mueller. Coherent states and their generalisations: a mathematical overview. *Reviews in Mathematical Physics*, 7(7):1013–1104, 1995.

[2] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Universitext. Springer, 2011.

[3] A. Cannas da Silva. *Lectures on Symplectic Geometry.* Springer-Verlag, second edition, 2008.

[4] C. R. de Oliveira. *Intermediate spectral theory and quantum dynamics*, volume 54 of *Progress in Mathematical Physics.* Birkhäuser Verlag AG, 2009.

[5] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, second edition, 2010.

[6] W. N. Everitt and L. Markus. Complex symplectic geometry with applications to ordinary differential operators. *Transactions of the American Mathematical Society*, 351(12):4905–4945, 1999.

[7] W. N. Everitt and L. Markus. Complex symplectic spaces and boundary value problems. *Bulletin of the American Mathematical Society*, 42(4):461–500, 2005.

[8] M. P. Gaffney. The harmonic operator for exterior differential forms. *Proceedings of the National Academy of Sciences of the United States of America*, 37(1):48–50, 1951.

[9] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry.* Springer-Verlag, third edition, 2004.

[10] G. Grubb. *Distributions and operators.* Springer-Verlag, 2009.

[11] V. Guillemin and A. Pollack. *Differential Topology.* Prentice-Hall, 1974.

[12] N. P. Landsman. *Between Classical and Quantum.* Handbook of the Philosophy of Science. Philosophy of Physics. Elsevier, 2007. pp. 417 - 553. Edited by J. Butterfield and J. Earman.

[13] R. B. Melrose. Differential analysis on manifolds with corners, 1996.

[14] J. R. Munkres. *Elementary Differential Topology.* Princeton Univerity Press, 1966.

[15] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, volume II: Fourier Analysis, Self-Adjointness. Academic Press, Inc., 1975.

[16] W. Roelcke. Über den laplace-operator auf riemannschen mannigfaltigkeiten mit diskontinuierlichen gruppen. *Mathematische Nachrichten*, 21(3-5):131–149, 1960.

[17] W. Rudin. *Functional analysis.* International Series in Pure and Applied Mathematics. Tata McGraw-Hill, second edition, 2006.

[18] B. P. Rynne and M. A. Youngson. *Linear Functional Analysis*. Springer Undergraduate Mathematics Series. Springer-Verlag, second edition, 2008.

[19] K. Schmüdgen. *Unbounded Self-adjoint Operators on Hilbert Space*. Graduate Texts in Mathematics. Springer, 2012.

[20] W. P. Thurston. *The Geometry and Topology of Three-Manifolds*. Princeton Univerity Press, 1997.