QUANTUM MECHANICS AND QUANTUM FIELD THEORY

A MATHEMATICAL PRIMER

JONATHAN DIMOCK

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9781107005099

This page intentionally left blank

Quantum Mechanics and Quantum Field Theory

Explaining the concepts of quantum mechanics and quantum field theory in a precise mathematical language, this textbook is an ideal introduction for graduate students in mathematics, helping to prepare them for further studies in quantum physics.

The textbook covers topics that are central to quantum physics: non-relativistic quantum mechanics, quantum statistical mechanics, relativistic quantum mechanics, and quantum field theory. There is also background material on analysis, classical mechanics, relativity, and probability. Each topic is explored through a statement of basic principles followed by simple examples. Around 100 problems throughout the textbook help readers develop their understanding.

Jonathan Dimock is Professor of Mathematics, SUNY at Buffalo. He has carried out research in various areas of mathematical physics, including constructive quantum field theory, quantum field theory on manifolds, renormalization group methods, and string theory.

Quantum Mechanics and Quantum Field Theory

A Mathematical Primer

JONATHAN DIMOCK

SUNY at Buffalo

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9781107005099

First published 2011

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data
Dimock, Jonathan, 1945–
Quantum mechanics and quantum field theory :
a mathematical primer / Jonathan Dimock.
p. cm.
ISBN 978-1-107-00509-9 (hardback)
1. Quantum theory – Mathematics. I. Title.
QC174.17.M35D56 2011
530.12–dc22
2010041723

ISBN 978-1-107-00509-9 Hardback

for Benjamin, Christina, Gregory, and Ann

Contents

Preface

This is a book on mathematical physics for a reader with a good background in mathematics, but possibly a minimal knowledge of physics. The subject matter is quantum physics and includes non-relativistic quantum mechanics, quantum statistical mechanics, relativistic quantum mechanics, and quantum field theory. The book only contains material which meets the twin criteria of being basic physics and being treatable with complete mathematical rigor. For each topic there is a straightforward statement of basic principles followed by simple examples. There is also background material in analysis, classical mechanics, relativity, and probability.

The book does not prove deep mathematical theorems. The book does not consider the complicated models of mathematical physics. The book does not enter into the fascinating speculative topics on the frontiers of physics, for example string theory. Finally the book does not consider questions concerning the foundations or philosophy of quantum physics. However the book does help prepare the reader for a journey in any of these directions.

The book assumes knowledge of elementary analysis, measure theory, linear algebra, some group theory, and some knowledge of differential equations. Some reference is made to manifolds, differential geometry, and Lie groups. Not much knowledge of physics is assumed beyond an introductory course. However one probably needs more than this to really appreciate the material.

The book is suitable for a graduate course in mathematics. In this connection there are problems scattered throughout the text. These serve the dual function of further developing the material and providing a study aid. The level of difficulty is quite variable.

Books which cover similar ground are Gustafson and Sigal (2003) and Takhtajan (2008). The mathematical level is about the same, but they have different points of emphasis.

Introduction

At the end of the nineteenth century most macroscopic phenomena could be explained in terms of a few basic equations. For the behavior of matter there was Newton's equation which said that the location of an object, modeled by a point $x \in \mathbb{R}^3$, evolves in time according to the equation

$$m\frac{d^2x}{dt^2} = F \tag{0.1}$$

Here $m > 0$ is the mass of the object and $F = F(t, x, dx/dt)$ is the sum of all the forces on the object. Forces were either gravitational or electromagnetic. In the electromagnetic case the force due to an electric field $E : \mathbb{R}^3 \to \mathbb{R}^3$ and a magnetic field $B : \mathbb{R}^3 \to \mathbb{R}^3$ on a particle with charge e was given by the Lorentz force $F = eE + e/c(dx/dt \times B)$. Here c is the speed of light, approximately 3×10^{10} cm/sec. In this case Newton's equations were

$$m\frac{d^2x}{dt^2} = eE + \frac{e}{c}\left(\frac{dx}{dt} \times B\right) \tag{0.2}$$

The electric and magnetic force fields (E, B) themselves might depend on time, and were determined by Maxwell's equations

$$\begin{aligned} \nabla \cdot E &= \rho \\ \nabla \cdot B &= 0 \\ \nabla \times E &= -\frac{1}{c}\frac{\partial B}{\partial t} \\ \nabla \times B &= \frac{1}{c}\left(\frac{\partial E}{\partial t} + j\right) \end{aligned} \tag{0.3}$$

where $\rho : \mathbb{R}^3 \to \mathbb{R}$ and $j : \mathbb{R}^3 \to \mathbb{R}^3$ are specified charge densities and current densities which necessarily obey the conservation law

$$\frac{1}{c}\frac{\partial \rho}{\partial t} + \nabla \cdot j = 0 \tag{0.4}$$

If ρ, j are expressed in terms of the positions of a number of particles obeying (0.2), the system of equations (0.2),(0.3) provide a model for an enormous range of phenomena.

However large velocity and large-scale gravitational phenomena were not accurately explained and it took the invention of special relativity (1905) and general

relativity (1915) by Einstein to rectify matters. Furthermore microscopic phenomena such as the structure of atoms were not accurately described and it took the invention of quantum mechanics by deBroglie, Schrödinger, Heisenberg, and others in the 1920s to rectify the situation. For phenomena involving both large velocities and tiny particles there is a synthesis known as quantum field theory which is still undergoing development.

Quantum mechanics does not itself contain physical laws. Rather it is a general framework in which physical laws should be formulated. As such it has a certain mysterious and *ad hoc* character; there is not much insight into why it is the way it is. However it is not ambiguous or inconsistent, and it has been very successful in describing microscopic phenomena.

In this book we explain quantum mechanics with particular mathematical care. In the first part of the book it is quantum mechanics without relativity. Here we take a historical, empirical approach to the subject and develop the theory as an extension of the classical equations (0.1), (0.2). After a discussion of general principles, the theme here is increasing complexity as the number of particles is increased, culminating in an introduction to quantum statistical mechanics.

In the second part of the book we add relativity to the mix studying quantum fields obeying various linear field equations such as (0.3) (which is already relativisitic, although its formulation predated relativity). A theme here is to develop the complementary field-particle aspects of the various cases. In this part we also make some attempt at understanding why the basic equations are natural from a mathematical point of view.

In the third part of the book we introduce some stochastic processes useful for analyzing various quantum problems. These are in fact essential for treating quantum fields obeying a nonlinear field equation. This is the interesting case since the nonlinearity corresponds to particle interactions. We illustrate the key role of stochastic processes by developing a two-dimensional model at some length.

Part I

Non-relativistic

1 Mathematical prelude

We begin with a survey of some of the mathematics we will need. The reader may wish to read it lightly and come back for details as needed.

Vector spaces can be real or complex, usually complex. A Banach space is a complete normed vector space. A Hilbert space is a Banach space in which the norm comes from an inner product. We review some basic facts about Hilbert spaces and Banach spaces in appendix A.

We are particularly interested in linear operators on a Hilbert space. Many of the results we present also hold for linear operators on a Banach space, but we will not need the more general result, and the proofs are sometimes easier for a Hilbert space.

1.1 Bounded operators

1.1.1 Definitions

A *linear operator* T from a Hilbert space $\mathcal{H}_1$ to a Hilbert space $\mathcal{H}_2$ is a mapping $T : \mathcal{H}_1 \to \mathcal{H}_2$ such that

$$T(af + bg) = aT(f) + bT(g) \qquad f, g \in \mathcal{H}_1, \quad a, b \in \mathbb{C} \tag{1.1}$$

The operator is *injective* or *one-to-one* (that is $Tf = Tg$ implies $f = g$) iff T has kernel $\{0\}$ (that is $Tf = 0$ implies $f = 0$). The operator is *surjective* or *onto* if the range is $\mathcal{H}_2$. The operator is *bijective* if it is injective and surjective and then there is an inverse $T^{-1} : \mathcal{H}_2 \to \mathcal{H}_1$ which is also linear. (If $\mathcal{H}_1 = \mathcal{H}_2$ is finite dimensional, then T is injective iff it is surjective, but not in general.)

A linear operator is *bounded* if there is a constant M such that

$$\|Tf\| \leq M\|f\| \tag{1.2}$$

for all $f \in \mathcal{H}_1$. Linear operators are continuous iff they are bounded.

The set of all bounded operators $T : \mathcal{H}_1 \to \mathcal{H}_2$ is itself a vector space with $(aT)f = a(Tf)$ and $(T_1 + T_2)f = T_1 f + T_2 f$. It is denoted $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ or $\mathcal{B}(\mathcal{H})$ if $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$. We define the norm of a bounded operator by

$$\|T\| = \sup_{f\neq 0} \frac{\|Tf\|}{\|f\|} = \sup_{\|f\|=1} \|Tf\| \tag{1.3}$$

Then we have

$$\|Tf\| \leq \|T\|\|f\| \tag{1.4}$$

and $\|T\|$ is the smallest possible constant here. With this norm, $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a normed vector space. We will see that it is complete and hence is a Banach space.

If $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, then by the Schwarz inequality $|(f, Tg)| \leq \|T\|\|f\|\|g\|$. Hence for $f \in \mathcal{H}_2$ the mapping $g \to (f, Tg)$ is a bounded linear functional on $\mathcal{H}_1$ and by the Riesz representation theorem (theorem A.3) there is a unique vector $f^* \in \mathcal{H}_1$ so that $(f, Tg) = (f^*, g)$. We define $T^*f = f^*$. Then $T^* : \mathcal{H}_2 \to \mathcal{H}_1$ is a linear operator called the *adjoint* of T and we have

$$(T^*f, g) = (f, Tg) \tag{1.5}$$

Then $|(T^*f, g)| \leq \|T\|\|f\|\|g\|$ and hence also by Riesz

$$\|T^*f\| = \sup_{\|g\|=1} |(T^*f, g)| \leq \|T\|\|f\| \tag{1.6}$$

Thus T^* is bounded, $T^* \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$, and $\|T^*\| \leq \|T\|$.

Now we look at some special classes of bounded operators.

1. Let $\mathcal{M} \subset \mathcal{H}$ be a closed linear subspace. Then any vector $f \in \mathcal{H}$ can be uniquely written $f = f_1 + f_2$ where $f_1 \in \mathcal{M}$ and $f_2 \in \mathcal{M}^\perp$ (theorem A.2). Define $P_{\mathcal{M}} f = f_1$. This is a bounded linear operator with norm 1 called the projection onto $\mathcal{M}$. It satisfies $P^2_{\mathcal{M}} = P_{\mathcal{M}}$ and $P^*_{\mathcal{M}} = P_{\mathcal{M}}$ and has range $\mathcal{M}$.

   More generally any bounded operator satisfying $P^2 = P$ and $P = P^*$ is called an *orthogonal projection*. One can show that any orthogonal projection has closed range $\mathcal{M}$ and that $P = P_{\mathcal{M}}$.
2. A linear operator T is an *isometry* if it is norm preserving, that is $\|Tf\| = \|f\|$. Since the norm determines the inner product by the polarization identity (A.5), it is equivalent to say that it is inner product preserving $(Tf, Tg) = (f, g)$. An isometry is bounded and injective.

   The range of an isometry is always closed. To see this suppose $Tf_n \to g$. Then $\|f_n - f_m\| = \|Tf_n - Tf_m\| \to 0$. Hence f_n is Cauchy and has a limit f. Then $Tf_n \to Tf$ by the continuity of T and hence $g = Tf$.

   An isometry satisfies $(f, (T^*T - I)g) = 0$ for any f, g, hence it satisfies $(T^*T - I)g = 0$ for any g, and hence

$$T^*T = I \tag{1.7}$$

   It follows that $P = TT^*$ is an orthogonal projection. The range is the same as the range of T (since $Tf = (TT^*)Tf$) and so

$$TT^* = P_{\mathrm{Ran}(T)} \tag{1.8}$$

3. If an isometry is also surjective, then the operator is called *unitary* or a *Hilbert space isomorphism*. In this case

$$T^*T = TT^* = I \tag{1.9}$$

and $T^* = T^{-1}$.

Problem 1.1 Let (M, μ) be a measure space and suppose $k(x, y)$ is an element of $L^2(M \times M, \mu \times \mu)$. Show that

$$(Kf)(x) = \int k(x, y) f(y) d\mu(y) \tag{1.10}$$

defines a bounded operator on $\mathcal{L}^2(M, \mu)$.

Problem 1.2 Let $T, S \in \mathcal{B}(\mathcal{H})$.

1. Show that TS is bounded and $\|TS\| \leq \|T\|\|S\|$.
2. Show that $\|T^*\| = \|T\|$.
3. Show that $\|T^*T\| = \|T\|^2$.

1.1.2 Sequences

A sequence of bounded operators $\{T_n\}$ converges strongly if $T_n f$ converges for all $f \in \mathcal{H}$. The sequence converges in norm if $\|T_n - T_m\| \to 0$ as $n, m \to \infty$. This is uniform convergence on the unit sphere. Norm convergence implies strong convergence since $\|T_n f - T_m f\| \leq \|T_n - T_m\| \|f\|$.

Theorem 1.1 *Let $T_n \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$.*

1. *If T_n converges strongly, then it has a strong limit, that is there is a bounded operator T such that $Tf = \lim_{n\to\infty} T_n f$.*
2. *If T_n converges in norm and T is the strong limit, then $\|T_n - T\| \to 0$ as $n \to \infty$.*

Remark We will use the principle of uniform boundedness which says that if $\|T_n f\|$ is bounded for each f, then there is a constant M such that $\|T_n\| \leq M$.

Proof For the first point define $Tf = \lim_{n\to\infty} T_n f$ and check that T is linear. Since $T_n f$ converges for each f, it is bounded for each f and so by the remark $\|T_n\| \leq M$. Now by the reverse triangle inequality, we can take the limit of $\|T_n f\| \leq M\|f\|$ and get $\|Tf\| \leq M\|f\|$. Hence T is bounded.

For the second point given ϵ choose N so that if $n, m \geq N$, then $\|T_n - T_m\| < \epsilon$. Then $\|T_n f - T_m f\| \leq \epsilon$ for all $\|f\| \leq 1$. Take the limit $m \to \infty$ and conclude that for $n \geq N$ we have $\|T_n f - Tf\| \leq \epsilon$ for all $\|f\| \leq 1$ and hence $\|T_n - T\| \leq \epsilon$. □

The second part of the theorem shows that $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ is a Banach space. If $\mathcal{H}_1 = \mathcal{H}_2 = \mathcal{H}$, then $\mathcal{B}(\mathcal{H})$ has even more structure. Since we can multiply operators and $\|TS\| \leq \|T\|\|S\|$, we say that $\mathcal{B}(\mathcal{H})$ is a Banach algebra. Since also $\|T\| = \|T^*\|$, we have a Banach $*$-algebra. Since also $\|T^*T\| = \|T\|^2$, we have a so-called C^*-algebra.

Theorem 1.2 *Let $T \in \mathcal{B}(\mathcal{H})$ satisfy $\|T\| < 1$. Then $I - T$ is bijective and the inverse is also a bounded operator.*

Proof Start by defining

$$S_n = \sum_{k=0}^{n} T^k \tag{1.11}$$

Then for $n > m$

$$\|S_n - S_m\| = \| \sum_{k=m+1}^{n} T^k \| \leq \sum_{k=m+1}^{n} \|T\|^k \to 0 \tag{1.12}$$

as $n, m \to \infty$. Thus S_n is a Cauchy sequence and since $\mathcal{B}(\mathcal{H})$ is complete, it has a limit which is a bounded operator

$$S = \lim_{n\to\infty} S_n \equiv \sum_{k=0}^{\infty} T^k \tag{1.13}$$

We then compute

$$(I - T)S_n = \sum_{k=0}^{n} T^k - \sum_{k=1}^{n+1} T^k = I - T^{n+1} \tag{1.14}$$

Now take the limit $n \to \infty$ and use $\|T^{n+1}\| \leq \|T\|^{n+1} \to 0$. Then $(I - T)S = I$ so $I - T$ is surjective. Similarly $S(I - T) = I$ so $I - T$ is injective and S is the inverse. □

1.1.3 Extensions

A subspace $\mathcal{D} \subset \mathcal{H}$ has a closure $\overline{\mathcal{D}}$, which is also a subspace. $\mathcal{D}$ is *dense* if $\overline{\mathcal{D}} = \mathcal{H}$. We consider the problem of extending a linear operator defined on a dense subspace.

Theorem 1.3 *Let $\mathcal{D}$ be a dense subspace of a Hilbert space $\mathcal{H}_1$ and let $T : \mathcal{D} \to \mathcal{H}_2$ be a bounded linear operator.*

1. *T has a unique extension to a bounded linear operator $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and the extension has the same bound.*

2. *If T is norm preserving, then so is the extension, that is it is an isometry.*
3. *If T is norm preserving and has dense range, the extension is unitary.*

Proof Our assumption is that $\|Tf\| \leq M\|f\|$ for all $f \in \mathcal{D}$. For $f \in \mathcal{H}_1$, choose a sequence $f_j \in \mathcal{D}$ so that $f_j \to f$. Then

$$\|Tf_j - Tf_k\| \leq M\|f_j - f_k\| \to 0 \tag{1.15}$$

as $j, k \to \infty$. Hence there is a limit $f^* = \lim_{j\to\infty} Tf_j$ and we define $Tf = f^*$. The definition is independent of the sequence since if f_j' is another sequence converging to f, we have

$$\|Tf_j - Tf_j'\| \leq M\|f_j - f_j'\| \to 0 \tag{1.16}$$

Now linearity and boundedness for T on $\mathcal{H}$ follow by taking limits of the same relations for T on $\mathcal{D}$.

The second assertion follows by taking the limit of $\|Tf_j\| = \|f_j\|$. The third assertion follows since the range of the extension is closed and dense and hence it is $\mathcal{H}$. □

1.1.4 Fourier transform

We want to define the Fourier transform as a unitary operator on $L^2(\mathbb{R}^n)$, but we start with a smaller space.

Let $\mathcal{S}(\mathbb{R}^n)$ be the Schwartz space of smooth rapidly decreasing functions on $\mathbb{R}^n$. These are complex valued $\mathcal{C}^\infty$ functions f on $\mathbb{R}^n$ with the property that for any multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n)$

$$\|x^\beta D^\alpha f\|_\infty < \infty \tag{1.17}$$

where

$$x^\beta = x_1^{\beta_1} \ldots x_n^{\beta_n} \qquad D^\alpha = D_1^{\alpha_1} \ldots D_n^{\alpha_n} \tag{1.18}$$

and $D_i = -i\partial/\partial x_i$. For $N > n/4$ we can write

$$f(x) = [(1 + |x|^2)^{-N}][(1 + |x|^2)^N f(x)] \tag{1.19}$$

This exhibits f as a product of an L^2 function and an L^∞ function and hence $f \in L^2(\mathbb{R}^n)$. Similarly $\|x^\beta D^\alpha f\|_2 < \infty$ for any multi-indices. Indeed we could have used these conditions to define the space. Examples of functions in $\mathcal{S}(\mathbb{R}^n)$ are those of the form $f(x) = P(x)e^{-a|x|^2}$ where P is a polynomial. Another example is the infinitely differentiable functions of compact support,[1] denoted $\mathcal{C}_0^\infty(\mathbb{R}^n)$. We have the inclusion of subspaces

$$\mathcal{C}_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \tag{1.20}$$

[1] The support of a function f, written $\operatorname{supp} f$, is the closure of $\{x \in \mathbb{R}^n : f(x) \neq 0\}$.

One can show $\mathcal{C}_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$ and so the same is true for $\mathcal{S}(\mathbb{R}^n)$.

Now for $f \in \mathcal{S}(\mathbb{R}^n)$ we define the *Fourier transform* $\tilde{f} : \mathbb{R}^n \to \mathbb{C}$ by

$$\tilde{f}(p) = (\mathcal{F}f)(p) = (2\pi)^{-n/2} \int e^{-ipx} f(x) dx \tag{1.21}$$

Since $f \in L^1(\mathbb{R}^n)$ by (1.19) with $N > n/2$, the integral converges and $\tilde{f}$ is bounded. The mapping $\mathcal{F}f = \tilde{f}$ is a linear operator.

More generally we compute

$$p^\alpha D_p^\beta (\mathcal{F}f) = (-1)^{|\beta|} \mathcal{F}(D^\alpha x^\beta f) \tag{1.22}$$

This exhibits $p^\alpha D_p^\beta \mathcal{F}f$ as the Fourier transform of a Schwartz function, hence it is bounded as well, and hence $\mathcal{F}f \in \mathcal{S}(\mathbb{R}^n)$. Thus the Fourier transform maps $\mathcal{S}(\mathbb{R}^n)$ to itself.

Next define

$$(\overline{\mathcal{F}}f)(p) = \tilde{f}(-p) = (2\pi)^{-n/2} \int e^{ipx} f(x) dx \tag{1.23}$$

which also maps $\mathcal{S}(\mathbb{R}^n)$ to itself. The basic inversion theorem says $\overline{\mathcal{F}}$ is the inverse of $\mathcal{F}$.

Theorem 1.4

1. $\overline{\mathcal{F}}\mathcal{F} = \mathcal{F}\overline{\mathcal{F}} = I$, *so* $\mathcal{F}$ *is a bijection on* $\mathcal{S}(\mathbb{R}^n)$.
2. $\mathcal{F}, \overline{\mathcal{F}}$ *extend to unitary operators on* $L^2(\mathbb{R}^n)$ *satisfying* $\overline{\mathcal{F}}\mathcal{F} = \mathcal{F}\overline{\mathcal{F}} = I$. *(So* $\overline{\mathcal{F}} = \mathcal{F}^{-1} = \mathcal{F}^*$*.)*

Proof For $f \in \mathcal{S}(\mathbb{R}^n)$ we compute

$$\begin{aligned}
(\overline{\mathcal{F}}\mathcal{F}f)(x) &= \lim_{\epsilon \to 0} (2\pi)^{-n/2} \int e^{ikx} e^{-\epsilon|k|^2/2} (\mathcal{F}f)(k) dk \\
&= \lim_{\epsilon \to 0} (2\pi)^{-n} \int f(y) e^{ik(x-y)} e^{-\epsilon|k|^2/2} dk\, dx \\
&= \lim_{\epsilon \to 0} (2\pi\epsilon)^{-n/2} \int f(y) e^{-|x-y|^2/2\epsilon} dx \\
&= f(x)
\end{aligned} \tag{1.24}$$

Here in the first step we regularize the $\overline{\mathcal{F}}$ integral. In the second step we insert the definition of $\mathcal{F}f$ and use Fubini's theorem to change the order of integration. In the third step we explicitly do the integral over k. The last step is a standard estimate using the facts that $(2\pi\epsilon)^{-n/2} e^{-|x-y|^2/2\epsilon}$ has integral one and peaks around $x = y$ as $\epsilon \to 0$. Details are left to the problems.

For the second point we compute that for $f, g \in \mathcal{S}(\mathbb{R}^n)$

$$(f, \mathcal{F}g) = (\overline{\mathcal{F}}f, g) \tag{1.25}$$

Then $(\mathcal{F}f, \mathcal{F}g) = (f, g)$ and the same for $\overline{\mathcal{F}}$. Hence they are norm preserving and by theorem 1.3 they extend to a unitary operators on $L^2(\mathbb{R}^n)$. The identity $\overline{\mathcal{F}}\mathcal{F} = \mathcal{F}\overline{\mathcal{F}} = I$ still holds since it holds on a dense set. □

Problem 1.3 Show that in $\mathbb{R}^n$

$$(2\pi)^{-n/2} \int e^{ikx} e^{-\epsilon|k|^2/2} dk = \epsilon^{-n/2} e^{-|x|^2/2\epsilon} \tag{1.26}$$

Problem 1.4 Show that if f is bounded and continuous on $\mathbb{R}^n$

$$\lim_{\epsilon \to 0} (2\pi\epsilon)^{-n/2} \int e^{-|x-y|^2/2\epsilon} f(y)\, dy = f(x) \tag{1.27}$$

Problem 1.5 If $f \in L^1 \cap L^2$, then the Fourier transform can be defined directly by (1.21). Show that this definition coincides with our definition on L^2.

Problem 1.6 For $f, g \in \mathcal{S}(\mathbb{R}^n)$ define the convolution

$$(f * g)(x) = \int f(x - y) g(y) dy \tag{1.28}$$

Show that $f * g \in \mathcal{S}(\mathbb{R}^n)$ and that

$$\mathcal{F}(f * g) = (2\pi)^{n/2} (\mathcal{F}f)(\mathcal{F}g) \tag{1.29}$$

1.2 Unbounded operators

1.2.1 Closed operators

We consider linear operators T from $\mathcal{H}_1$ to $\mathcal{H}_2$ defined on a subspace $D(T) \subset \mathcal{H}_1$. The operator is not necessarily bounded, but we would like it to be closed. An operator is *closed* if for any sequence $f_n \in D(T)$ we have that $f_n \to f$ and $Tf_n \to g$ imply $f \in D(T)$ and $Tf = g$. A bounded operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is easily seen to be closed. Furthermore if $D(T) = \mathcal{H}_1$, we have:

Theorem 1.5 *(Closed graph theorem) If $T : \mathcal{H}_1 \to \mathcal{H}_2$ is closed, then it is bounded.*

Thus if the operator is closed and unbounded, $D(T)$ must be a proper subspace of $\mathcal{H}_1$. Usually it will be a dense subspace. Here is an example:

Example 1.1 Let $\mathcal{H} = L^2(\mathbb{R})$ and let

$$D(T) = \{f \in \mathcal{H} : \int |x|^2 |f(x)|^2 dx < \infty\} \tag{1.30}$$

Define $T : D(T) \to \mathcal{H}$ by

$$(Tf)(x) = xf(x) \tag{1.31}$$

This is not bounded since if f_n is the characteristic function of $[n, n+1]$, then $\|Tf_n\| \geq n\|f_n\|$. However suppose f_n is a sequence such that $f_n \in D(T) \to f$ and $Tf_n \to g$. After passing to a subsequence we have $f_{n_j}(x) \to f(x)$ and $xf_{n_j}(x) \to g(x)$ for almost every x. Hence $xf(x) = g(x)$ for almost every x so $f \in D(T)$ and $Tf = g$. Thus T is closed.

The *graph* of an operator $T : D(T) \to \mathcal{H}_2$ is the subset of $\mathcal{H}_1 \times \mathcal{H}_2$ defined by

$$\Gamma(T) = \{< f, g >: f \in D(T), Tf = g\} \tag{1.32}$$

This is in fact a subspace of $\mathcal{H}_1 \times \mathcal{H}_2$. A subspace of $\mathcal{H}_1 \times \mathcal{H}_2$ is the graph of an operator iff it has no elements of the form $< 0, g >$ with $g \neq 0$.

An operator S is an extension of an operator T written $T \subset S$ if the domains satisfy $D(T) \subset D(S)$ and $Tf = Sf$ for $f \in D(T)$. Then $T \subset S$ iff $\Gamma(T) \subset \Gamma(S)$.

An operator is closed iff every sequence $< f_n, Tf_n >$ in $\Gamma(T)$ converging to $< f, g > \in \mathcal{H}_1 \oplus \mathcal{H}_2$ has $< f, g > \in \Gamma(T)$. Thus an operator is closed iff its graph is closed. It follows that an injective operator T is closed iff T^{-1} is closed.

An operator T that is not closed may have a closed extension. This is true iff the closure of the graph is the graph of an operator, that is $\overline{\Gamma(T)} = \Gamma(\bar{T})$. In this case we say the operator is *closable* and call $\bar{T}$ the *closure* of T. If T is a closed operator, a subspace $D \subset D(T)$ is a *core* for T if $\overline{T \mid D} = T$. This is the same as the statement that for every $\psi \in D(T)$ there is a sequence $\psi_n \in D$ so that $\psi_n \to \psi$ and $T\psi_n \to T\psi$.

Example 1.2 In the Hilbert space $\mathcal{H} = L^2[-1, 1]$, consider the subspace $D(T)$ of bounded continuous functions and the operator $T : D(T) \to \mathcal{H}$ defined by

$$(Tf)(x) = f(0) \tag{1.33}$$

Let $f_n \in D(T)$ be the "tent function" which is piecewise linear and satisfies $f_n(\pm 1) = 0, f_n(\pm 1/n) = 0, f_n(0) = 1$. Then as $n \to \infty, f_n \to 0$ and $Tf_n = 1 \to 1$, so $< 0, 1 >$ is in the closure of the graph, which is therefore not the graph of an operator. Thus T is not closable.

1.2.2 Spectrum of a closed operator

For $D(T) \subset \mathcal{H}$ let $T : D(T) \to \mathcal{H}$ be a closed operator. The *resolvent set* $\rho(T)$ is all complex z such that $T - z : D(T) \to \mathcal{H}$ is a bijection and the inverse is bounded

$$\rho(T) = \{z \in \mathbb{C} : (T - z)^{-1} \in \mathcal{B}(\mathcal{H})\} \tag{1.34}$$

(This would be empty if T were not closed.) The *spectrum* $\sigma(T)$ is the complement of the resolvent set.

A complex number z is an *eigenvalue* for T if $(T - z)f = 0$ for some $f \neq 0$. Then $T - z$ is not injective and so z is in the spectrum. The set of all eigenvalues is a subset of the spectrum called the *point spectrum*. If z is not an eigenvalue, then $T - z$ is injective, but z still may not be in the resolvent set since the range of $T - z$ may not be all of $\mathcal{H}$. In this case we make a further distinction and specify that z is in the *continuous spectrum* if the range of $T - z$ is dense and otherwise z is in the *residual spectrum*.

Theorem 1.6 *$\rho(T)$ is open and $\sigma(T)$ is closed.*

Proof Suppose $z_0 \in \rho(T)$. Then on the domain of T

$$(T - z) = (I - (z - z_0)(T - z_0)^{-1})(T - z_0) \tag{1.35}$$

By theorem 1.2 the operator $I - (z - z_0)(T - z_0)^{-1}$ is a bijection and has a bounded inverse if

$$\|(z - z_0)(T - z_0)^{-1}\| < 1 \tag{1.36}$$

Hence $T - z$ has a bounded inverse under the same condition which we write

$$|z - z_0| < \|(T - z_0)^{-1}\|^{-1} \tag{1.37}$$

Hence this disc is in the resolvent set which is therefore open. □

Problem 1.7 Show that if T is bounded, $\sigma(T) \subset \{z \in \mathbb{C} : |z| \leq \|T\|\}$.

Problem 1.8 Show that if U is unitary, $\sigma(U) \subset \{z \in \mathbb{C} : |z| = 1\}$.

1.2.3 Adjoints

We generalize the notion of adjoint to unbounded operators. For $D(T) \subset \mathcal{H}$ let $T : D(T) \to \mathcal{H}$ be densely defined, but not necessarily closed. Let $D(T^*)$ be all vectors $g \in \mathcal{H}$ such that the function $f \in D(T) \to (g, Tf)$ is continuous. Equivalently

$$D(T^*) = \{g \in \mathcal{H} : \exists C \text{ so } |(g, Tf)| \leq C\|f\| \text{ for all } f \in D(T)\} \tag{1.38}$$

Then the linear functional has an extension to all of $\mathcal{H}$ and by the Riesz representation theorem (theorem A.3) if $g \in D(T^*)$, then there is a unique g^* such that $(g^*,f) = (g,Tf)$. We define a new operator $T^* : D(T^*) \to \mathcal{H}$ by $T^*g = g^*$. Then T^* is a linear operator called the *adjoint* of T. It is defined so that

$$(T^*g,f) = (g,Tf) \qquad f \in D(T), g \in D(T^*) \tag{1.39}$$

Theorem 1.7 *Define V on $\mathcal{H} \times \mathcal{H}$ by $V < f,g >=< -g,f >$. Then*

$$\Gamma(T^*) = V[\Gamma(T)]^\perp \tag{1.40}$$

Proof $< f,g > \in V[\Gamma(T)]^\perp$ is equivalent to $(-g,h) + (f,Th) = 0$ for all $h \in D(T)$. But by definition this is equivalent to $f \in D(T^*)$ and $T^*f = g$, that is $< f,g > \in \Gamma(T^*)$. □

Corollary 1.1 *T^* is closed.*

Proof $\Gamma(T)^\perp$ is closed and V is unitary. □

Corollary 1.2 *If $T \subset S$, then $S^* \subset T^*$.*

Proof $\Gamma(T) \subset \Gamma(S)$ implies $\Gamma(S)^\perp \subset \Gamma(T)^\perp$. □

Corollary 1.3 *If $D(T^*)$ is dense so T^{**} exists, then T is closable and $T^{**} = \bar{T}$.*

Proof

$$\Gamma(T^{**}) = V[V\Gamma(t)^\perp]^\perp = [\Gamma(T)]^{\perp\perp} = \overline{\Gamma(T)} \tag{1.41}$$

Since $\overline{\Gamma(T)}$ is the graph of an operator, T is closable and the closure is that operator, that is $\bar{T} = T^{**}$. □

Corollary 1.4

$$\text{Ker } T^* = [\text{Ran}(T)]^\perp \tag{1.42}$$

Proof The kernel of T^* is all g so $(g,0) \in \Gamma(T^*)$. By the theorem this is the same as $(0,g) \in \Gamma(T)^\perp$, that is $g \in [\text{Ran}(T)]^\perp$. □

1.3 Self-adjoint operators

1.3.1 Definitions

A densely defined operator is *symmetric* if

$$(g,Tf) = (Tg,f) \tag{1.43}$$

for all $f, g \in D(T)$. Then $g \in D(T^*)$ and $T^*g = Tg$ so that $T \subset T^*$. If the domains are the same, that is if $T = T^*$, then the operator is said to be *self-adjoint* . A self-adjoint operator is necessarily closed.

Self-adjoint operators have nice properties not shared by symmetric operators as we will see. A symmetric operator T fails to be self-adjoint because its domain is too small. Indeed if S is a symmetric extension of T, then $T \subset S \subset S^* \subset T^*$ so S is closer to being self-adjoint. A general problem is to find a large enough symmetric extension of T so that it is self-adjoint.

If T is symmetric, then $D(T^*)$ is dense so the closure T^{**} exists. Then $T \subset T^*$ implies $T^{**} \subset T^*$ and we have the situation

$$T \subset T^{**} \subset T^* \tag{1.44}$$

Now T^{**} is always symmetric ($T^{**} \subset T^{***}$). The simplest possibility for a self-adjoint extension for T is that $T^{**} = \bar{T}$ is self-adjoint, that is $T^{**} = T^{***}$. We say that T is *essentially self-adjoint*. Since T^* is closed $T^{***} = T^*$ and an equivalent statement is that $T^{**} = T^*$, that is T^* is self-adjoint.

Example 1.3 Let $\mathcal{H} = L^2[0, 1]$ and let $T = id/dx$ defined on $\mathcal{C}^1$ functions with compact support in (0,1). Then integration by parts shows that T is symmetric. However if we let $h(x) = e^x$, then integration by parts also gives $(h, Tf) = (ih, f)$. Hence $h \in D(T^*)$ and $T^*h = ih$. The imaginary eigenvalue means that T^* is not symmetric and hence T is not essentially self-adjoint.

We quote without proof some further results.[2] Consider the subspace

$$AC[0, 1] = \{f \in L^2[0, 1] : f \text{ is absolutely continuous and } f' \in L^2[0, 1]\}. \tag{1.45}$$

(Absolutely continuous implies $f' \in L^1[0, 1]$. Since $L^2[0, 1] \subset L^1[0, 1]$, we are assuming a bit more.) Let $S = id/dx$ now defined on the larger domain

$$D(S) = \{f \in AC[0, 1] : f(0) = f(1) = 0\} \tag{1.46}$$

One can still integrate by parts and show that this operator is symmetric. It turns out it is also closed, but it is not self-adjoint. The adjoint has the domain $D(S^*) = AC[0, 1]$. A further extension is a family of operators S_α indexed by a complex number α with $|\alpha| = 1$. We have $S_\alpha = id/dx$ with domain

$$D(S_\alpha) = \{f \in AC[0, 1] : f(0) = \alpha f(1)\} \tag{1.47}$$

These turn out to be self-adjoint. Thus there is a family of self-adjoint extensions and we have

$$T \subset S \subset S_\alpha = S_\alpha^* \subset S^* \subset T^* \tag{1.48}$$

[2] See Reed and Simon (1975: 141) for details.

This example has a property typical of differential operators in a region with a boundary, namely the choice of a self-adjoint extension corresponds to a choice of boundary conditions.

Problem 1.9 Show that if T is essentially self-adjoint and S is a symmetric extension, then S is essentially self-adjoint and $S^{**} = T^{**}$. (Thus if T is essentially self-adjoint, then it has a unique self-adjoint extension.)

1.3.2 Properties

If T is symmetric, then for $f \in D(T)$ the quantity (f, Tf) is real since $\overline{(f, Tf)} = (Tf, f) = (f, Tf)$. It follows that any eigenvalue must be real. For self-adjoint operators we have the following stronger statement.

Theorem 1.8 *The spectrum of a self-adjoint operator is a subset of the real line.*

Proof We have to show that a complex number z with $\mathrm{Im}z \neq 0$ is in the resolvent set. First we note that for $f \in D(T)$

$$|\mathrm{Im}z|\|f\|^2 = |\mathrm{Im}((T-z)f, f)| \leq |((T-z)f, f)| \leq \|(T-z)f\|\|f\| \tag{1.49}$$

and so

$$|\mathrm{Im}z|\|f\| \leq \|(T-z)f\| \tag{1.50}$$

Hence $(T-z)f = 0$ implies $f = 0$ so T is injective.

The inequality also implies that $T - z$ has a closed range since if $(T-z)f_n$ is a sequence in the range converging to g, then

$$\|f_n - f_m\| \leq |\mathrm{Im}z|^{-1}\|(T-z)(f_n - f_m)\| \to 0 \tag{1.51}$$

as $n, m \to \infty$. Then f_n is Cauchy and so converges to some f. Since T is closed, $T - z$ is closed. Then $f_n \to f$ and $(T-z)f_n \to g$ imply $f \in D(T-z) = D(T)$ and $(T-z)f = g$. Hence g is in the range and so the range is closed.

Now by (1.42) we have

$$[\mathrm{Ran}(T-z)]^{\perp} = \mathrm{Ker}(T^* - \bar{z}) = \mathrm{Ker}(T - \bar{z}) = \{0\} \tag{1.52}$$

Hence $\mathrm{Ran}(T-z) = [\mathrm{Ran}(T-z)]^{\perp\perp} = \mathcal{H}$.

Thus $T - z$ is a bijection from $D(T)$ to $\mathcal{H}$. Now in (1.50) let $f = (T-z)^{-1}g$ for any $g \in \mathcal{H}$. This gives

$$\|(T-z)^{-1}g\| \leq |\mathrm{Im}z|^{-1}\|g\| \tag{1.53}$$

which shows that the inverse is bounded. □

The following is a test for self-adjointness:

Theorem 1.9 *A symmetric operator T is self-adjoint iff* $\mathrm{Ran}(T \pm i) = \mathcal{H}$.

Proof If T is self-adjoint, then $\pm i$ is in the resolvent set by the previous theorem and hence the result.

For the converse suppose T is symmetric and the range of $T \pm i$ is $\mathcal{H}$. We must show $D(T^*) \subset D(T)$. If $g \in D(T^*)$, choose $f \in D(T)$ so that

$$(T^* - i)g = (T - i)f \tag{1.54}$$

Since also $f \in D(T^*)$ and $T^* f = Tf$, this says

$$(T^* - i)(f - g) = 0 \tag{1.55}$$

But by (1.42) we have $\mathrm{Ker}(T^* - i) = [\mathrm{Ran}(T + i)]^\perp = \{0\}$ so $f = g$ and hence $g \in D(T)$. □

Problem 1.10 A self-adjoint operator is *positive* if $(f, Tf) \geq 0$ for all $f \in D(T)$. Show that in this case $\sigma(T) \subset [0, \infty)$.

Problem 1.11 Let T be self-adjoint. Show that $z \in \sigma(T)$ iff for every $\epsilon > 0$ there exists a $f \in \mathcal{H}$ with $\|f\| = 1$ such that $\|(T - z)f\| < \epsilon$.

Problem 1.12 Show that if T is symmetric and $\mathrm{Ran}(T \pm i)$ is dense, then T is essentially self-adjoint.

Problem 1.13 Let T be symmetric and suppose the domain contains a complete set of eigenvectors $e_1, e_2, \ldots$ with eigenvalues $\lambda_1, \lambda_2, \ldots$. Show that T is essentially self-adjoint and that the spectrum of the closure is the closure of the set of eigenvalues.

Problem 1.14 Let $(\mathcal{M}, \mu)$ be a measure space and let $\tau : \mathcal{M} \to \mathbb{R}$ be a measurable function. Define an operator $[\tau]$ on $L^2(\mathcal{M}, \mu)$ by $([\tau]f)(x) = \tau(x)f(x)$ with domain

$$D([\tau]) = \{f \in L^2(\mathcal{M}, \mu) : \tau f \in L^2(\mathcal{M}, \mu)\} \tag{1.56}$$

1. Show that $[\tau]$ is self-adjoint.
2. Show that the spectrum is the essential range of τ. (The essential range of τ is all $\lambda \in \mathbb{R}$ such that $\mu\left(\tau^{-1}(\lambda - \epsilon, \lambda + \epsilon)\right)$ is positive for all $\epsilon > 0$.)

1.3.3 Spectral theorem

Suppose that T is a bounded self-adjoint operator on a Hilbert space $\mathcal{H}$ and suppose T has a complete set of eigenvectors $e_1, e_2, \ldots$ with eigenvalues $\lambda_1, \lambda_2, \ldots$, which are bounded but not necessarily distinct. (For example suppose T is a compact operator which we study in the next section.) Then we have for any $f \in \mathcal{H}$

$$\begin{aligned} f &= \sum_{i=1}^{\infty} (e_i, f) e_i \\ Tf &= \sum_{i=1}^{\infty} (e_i, f) \lambda_i e_i \end{aligned} \tag{1.57}$$

The operator $V : \mathcal{H} \to \ell^2$ defined by

$$(Vf)_i = (e_i, f) \tag{1.58}$$

is unitary. Define a multiplication operator $[\lambda] : \ell^2 \to \ell^2$ by

$$([\lambda] f)_i = \lambda_i f_i \tag{1.59}$$

Then $VT = [\lambda]V$ or

$$T = V^{-1}[\lambda]V \tag{1.60}$$

Thus T is unitarily equivalent to a multiplication operator.

The content of the spectral theorem is that something similar is true for any bounded self-adjoint operator.

Theorem 1.10 *(Spectral theorem – bounded operator) Let T be a bounded self-adjoint operator on a Hilbert space $\mathcal{H}$. Then there exists a measure space $(\mathcal{M}, \mu)$, a bounded measurable function $\tau : \mathcal{M} \to \mathbb{R}$, and a unitary operator $V : \mathcal{H} \to L^2(\mathcal{M}, d\mu)$ such that $T = V^{-1}[\tau]V$ where $[\tau]$ is the operator multiplication by τ.*

In the example, $(\mathcal{M}, \mu)$ is the integers with counting measure. There is also a version for unbounded operators:

Theorem 1.11 *(Spectral theorem – unbounded operators) Let T be a self-adjoint operator on a Hilbert space $\mathcal{H}$. Then there exists a measure space $(\mathcal{M}, \mu)$, a measurable function $\tau : \mathcal{M} \to \mathbb{R}$, and a unitary operator $V : \mathcal{H} \to L^2(\mathcal{M}, d\mu)$ such that $VD(T) = D([\tau])$ as defined in (1.56) and $T = V^{-1}[\tau]V$.*

For proofs see Reed and Simon (1980). The representation as a multiplication operator is not unique.

The spectral theorem allows us to define functions of a self-adjoint operator. Suppose T is self-adjoint, bounded or not, and let $h : \mathbb{R} \to \mathbb{C}$ be a bounded Borel

function.[3] Then $h \circ \tau$ is measurable and we define $h(T)$ by

$$h(T) = V^{-1}[h \circ \tau]V \tag{1.61}$$

Then $h(T)$ is a bounded operator for we have

$$\|h(T)f\| = \|Vh(T)f\|_2 = \|[h \circ \tau]Vf\|_2 \leq \|h \circ \tau\|_\infty \|Vf\|_2 \leq \|h\|_\infty \|f\| \tag{1.62}$$

If h is unbounded, we can still use (1.61) to define $h(T)$ but now restrict the domain to $V^{-1}D([h \circ \tau])$.

The definition $h(T)$ has the following properties (known as the functional calculus):

1. $(h_1 + h_2)(T) = h_1(T) + h_2(T)$
2. $(\lambda h)(T) = \lambda h(T) \qquad \lambda \in \mathbb{C}$
3. $(h_1 \cdot h_2)(T) = h_1(T)h_2(T)$
4. $1(T) = I$
5. $h(T)^* = \overline{h}(T)$.

An important case is when h is the characteristic function of a Borel set $B \subset \mathbb{R}$. We define

$$E(B) = \chi_B(T) \tag{1.63}$$

Then we have by the functional calculus

$$E(B)^2 = E(B) \qquad E(B)^* = E(B) \tag{1.64}$$

Thus $E(B)$ is an orthogonal projection. The $E(B)$ are called the *spectral projections* for T. The projections $E(\lambda) \equiv E((-\infty, \lambda])$ are increasing and satisfy $E(-\infty) = 0$ and $E(\infty) = I$.

For $f \in \mathcal{H}$ we define Borel measures $(f, E(B)f)$ of total mass $\|f\|^2$. These are called the *spectral measures* for the operator. The integral of a function h with respect to such a measure is denoted

$$\int h(\lambda) d(f, E(\lambda)f) \tag{1.65}$$

Problem 1.15 Let T be a self-adjoint operator with spectral projections $E(B)$.

1. Show that $\mu(B) = (f, E(B)f)$ defines a Borel measure.
2. Show that for bounded h

$$\begin{aligned} (f, h(T)f) &= \int h(\lambda) d(f, E(\lambda)f) \\ \|h(T)f\|^2 &= \int |h(\lambda)|^2 d(f, E(\lambda)f) \end{aligned} \tag{1.66}$$

[3] The Borel sets in $\mathbb{R}$ are the smallest σ-algebra of subsets which contains the open sets. A function $h : \mathbb{R} \to \mathbb{R}$ is Borel measurable if $h^{-1}(\mathcal{O})$ is a Borel set for any open $\mathcal{O}$. If f is a measurable real-valued function on any measure space, and h is Borel, then $h \circ f$ is also measurable.

3. Show that $f \in D(T)$ iff $\int \lambda^2 d(f, E(\lambda)f)$ is finite in which case

$$\begin{aligned}(f, Tf) &= \int \lambda \, d(f, E(\lambda)f) \\ \|Tf\|^2 &= \int \lambda^2 d(f, E(\lambda)f)\end{aligned} \tag{1.67}$$

1.3.4 One-parameter groups

A *one-parameter unitary group* is defined to be a representation of the additive group $\mathbb{R}$ by unitary operators. More precisely a one-parameter unitary group is a function U from $\mathbb{R}$ to unitary operators on a Hilbert space such that $U(t)U(s) = U(t+s)$ and $U(0) = I$. Then $U(t)^* = U(t)^{-1} = U(-t)$.

A self-adjoint operator determines a one-parameter unitary group as follows:

Theorem 1.12 *Let T be self-adjoint on $\mathcal{H}$ and define $U(t) = \exp(itT)$.*

1. *$U(t)$ is a one-parameter unitary group.*
2. *$t \to U(t)f$ is strongly continuous from $\mathbb{R}$ to $\mathcal{H}$ for any $f \in \mathcal{H}$.*
3. *If $f \in D(T)$, then $f(t) \equiv U(t)f \in D(T)$ and solves the differential equation*

$$\frac{df}{dt} = iTf \tag{1.68}$$

Proof $U(t)$ is defined via the spectral theorem as $U(t) = V^{-1}[e^{i\tau}]V$. It is straightforward to check that this defines a one-parameter unitary group. For the continuity we use the dominated convergence theorem to show that

$$\begin{aligned}\|U(t+h)f - U(t)f\|^2 &= \|(U(h) - I)f\|^2 \\ &= \int |e^{i\tau(m)h} - 1|^2 |(Vf)(m)|^2 d\mu(m) \\ &\to 0 \text{ as } h \to 0\end{aligned} \tag{1.69}$$

For the last point $U(t)f \in D(T)$ by the spectral theorem and we have

$$\begin{aligned}&\left\| \left(\frac{U(t+h) - U(t)}{h} \right) f - iTU(t)f \right\|^2 \\ =&\left\| \left[\frac{U(h) - I}{h} - iT \right] f \right\|^2 \\ =&\int \left| \frac{e^{i\tau(m)h} - 1}{h} - i\tau(m) \right|^2 |(Vf)(m)|^2 d\mu(m) \\ &\to 0 \text{ as } h \to 0\end{aligned} \tag{1.70}$$

Here again we use the dominated convergence theorem, using $|e^{ix} - 1| \leq |x|$ and the fact that $|\tau|^2 |Vf|^2$ is integrable. □

The operator T is called the *generator* of the unitary group. The last theorem has a converse known as Stone's theorem.

Theorem 1.13 *Let $U(t)$ be a strongly continuous one-parameter unitary group. Then there is a unique self-adjoint operator T such that $U(t) = e^{iTt}$.*

Proof (sketch) For the existence part of Stone's theorem follow the steps below:

1. Define $D(T)$ to be all $f \in \mathcal{H}$ such that $t^{-1}(U(t) - I)f$ converges as $t \to 0$ and for $f \in D(T)$ define
$$Tf = \lim_{t\to 0} \frac{(U(t) - I)}{it} f$$
Show that $D(T)$ is a subspace and that T is a linear operator.
2. For $\lambda > 0$ define
$$R_\lambda f = \frac{1}{i} \int_0^\infty e^{-\lambda s} U(s)f \, ds$$
Show that the integral exists as a Hilbert space valued Riemann integral and defines a bounded operator.
3. Show that $R_\lambda f \in D(T)$ and that
$$(T + i\lambda)R_\lambda f = f$$
4. Show that for any $f \in \mathcal{H}$ we have $\lambda R_\lambda f \to f$ as $\lambda \to \infty$ and conclude that $D(T)$ is dense.
5. Show that if $f \in D(T)$, then $U(t)f \in D(T)$ and
$$\frac{1}{i}\frac{d}{dt} U(t)f = TU(t)f = U(t)Tf$$
6. Show that T is symmetric by verifying for $f \in D(T)$
$$0 = \frac{1}{i}\frac{d}{dt} \|U(t)f\|^2 = (Tf, f) - (f, Tf)$$
7. Show that T is self-adjoint by verifying $\mathrm{Ran}(T \pm i\lambda) = \mathcal{H}$. (We already know this with the plus sign.)
8. Show that $U(t) = e^{iTt}$ by defining $h(t) = (U(t) - e^{iTt})f$ and showing that
$$\frac{d}{dt} \|h(t)\|^2 = 0$$

□

Problem 1.16 Supply the details in the above argument.

Here is a variation of the above. Let H be a positive self-adjoint operator. Then $V(t) = e^{-tH}$ is bounded and self-adjoint for $t \geq 0$ and gives a representation of the additive semi-group $\mathbb{R}^+ = [0, \infty)$ ("semi-group" since there are no inverses). Conversely we have as in Stone's theorem:

Theorem 1.14 *Let $V(t)$ be a semi-group of bounded self-adjoint operators defined for $t \geq 0$ satisfying:*

1. $\|V(t)\| \leq 1$
2. $V(0) = I$
3. $V(t)V(s) = V(t+s)$
4. $t \to V(t)f$ *is continuous for all* $f \in \mathcal{H}$.

Then there is a unique positive self-adjoint operator H such that $V(t) = e^{-tH}$.

1.4 Compact operators

1.4.1 Properties

A Hilbert space is a metric space and so a subset $K \subset \mathcal{H}$ is compact iff every sequence in K has a convergent subsequence. In a finite-dimensional Hilbert space, a subset is compact iff it is closed and bounded. But in an infinite-dimensional Hilbert space, closed and bounded is not sufficient. For example an orthonormal basis $\{\phi_1, \phi_2, \dots\}$ is closed and bounded, but since $\|\phi_i - \phi_j\|^2 = 2$ for all $i \neq j$, there can be no convergent subsequence.

An bounded operator $T : \mathcal{H}_1 \to \mathcal{H}_2$ is *compact* if it maps bounded sequences in $\mathcal{H}_1$ into sequences in $\mathcal{H}_2$ with convergent subsequences. An operator is *finite rank* if the range is finite dimensional. By the above remarks a finite rank operator is compact. More generally a norm limit of finite rank operators is compact. This follows from:

Lemma 1.1 *The compact operators form a closed subspace of $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$.*

Proof It is straightforward to show that sums of compacts are compact, and it is trivial to show that multiplication by a scalar preserves compactness. Thus they form a subspace.

To show that it is a closed subspace let T_n be a sequence of compact operators such that $\|T_n - T\| \to 0$ as $n \to \infty$. We must show that T is compact. Let f_n be a bounded sequence in $\mathcal{H}_1$. Then there is a subsequence $f_n^{(1)}$ such that $T_1 f_n^{(1)}$ is Cauchy. Then this sequence has a subsequence $f_n^{(2)}$ so that $T_2 f_n^{(2)}$ is Cauchy (as is $T_1 f_n^{(2)}$). Continuing in this fashion we get for each k a subsequence $f_n^{(k)}$ such that $T_k f_n^{(k)}$ is Cauchy. The diagonal sequence $g_n = f_n^{(n)}$ is then a subsequence of each $f_n^{(k)}$ and so $T_k g_n$ is Cauchy for all k.

Now given $\epsilon > 0$, choose k so that $\|T - T_k\| < \epsilon$, and choose N so that for $n, m \geq N$ we have $\|T_k g_n - T_k g_m\| \leq \epsilon$. Let $M = \sup_n \|f_n\|$. Then for $n, m \geq N$

$$\begin{aligned} \|Tg_n - Tg_m\| \leq & \|(T - T_k)g_n\| + \|T_k(g_n - g_m)\| + \|(T_k - T)g_m\| \\ \leq & (2M + 1)\epsilon \end{aligned} \tag{1.71}$$

Since ϵ is arbitrary, Tg_n is Cauchy, and since g_n is a subsequence of f_n, this shows T is compact. □

Lemma 1.2 *The compact operators on $\mathcal{H}$ are a (two-sided) $*$-ideal in the Banach algebra $\mathcal{B}(\mathcal{H})$. That is:*

1. *The compacts are a subspace of $\mathcal{B}(\mathcal{H})$.*
2. *If T is compact and S is bounded, then TS and ST are compact.*
3. *If T is compact, then so is T^*.*

We already noted the first. The second is easy. We will not need the third, so we omit it.

Lemma 1.3 *Let $T \in \mathcal{B}(\mathcal{H})$ be compact. If λ is not an eigenvalue and $\lambda \neq 0$, then* $\mathrm{Ran}(T - \lambda)$ *is closed.*

Proof Suppose $f_n \in D(T)$ and $(T - \lambda)f_n \to g$. We show $g \in \mathrm{Ran}(T - \lambda)$.

If f_n is not bounded, then there is a subsequence going to infinity, so we may as well assume $\|f_n\| \to \infty$. Then $h_n = f_n/\|f_n\|$ is bounded and so there is a subsequence h_{n_j} so that Th_{n_j} converges. We also have that $(T - \lambda)h_{n_j} = (T - \lambda)f_{n_j}/\|f_{n_j}\|$ converges to zero. Combining these statements and the fact that $\lambda \neq 0$ we conclude that $h_{n_j} \to h$ with $\|h\| = 1$. But also $(T - \lambda)h = 0$, which contradicts the assumption that λ is not an eigenvalue.

Thus we may assume f_n is bounded. Then Tf_n has a convergent subsequence Tf_{n_j}. Since $\lambda \neq 0$, it follows that $f_{n_j} \to f$. Then $g = \lim_j (T - \lambda)f_{n_j} = (T - \lambda)f$ as required. □

The next result characterizes the spectrum of a compact operator:

Theorem 1.15 (Riesz–Schauder) *Let $T \in \mathcal{B}(\mathcal{H})$ be compact.*

1. *Complex $\lambda \neq 0$ is either an eigenvalue or else is in the resolvent set.*
2. *Eigenvalues $\lambda \neq 0$ have finite multiplicity (*that is $\dim \mathrm{Ker}(T - \lambda) < \infty$*).*
3. *Eigenvalues have no limit point except possibly zero.*

Proof We give the proof with the simplification that T is self-adjoint so the spectrum is real. Suppose λ is real, is not an eigenvalue, and $\lambda \neq 0$. Then $T - \lambda$ is injective. $\mathrm{Ran}(T - \lambda)$ is closed by the lemma, and since $\mathrm{Ran}(T - \lambda)^\perp = \mathrm{Ker}(T - \lambda) = \{0\}$ by (1.42), it follows that $\mathrm{Ran}(T - \lambda) = \mathcal{H}$. The inverse $(T - \lambda)^{-1}$ is bounded by the closed graph theorem and hence λ is in the resolvent set. This proves the first point.

For the second point suppose $\dim \mathrm{Ker}(T - \lambda) = \infty$. Let ϕ_n be an orthonormal basis for this space. Then there is a subsequence so $T\phi_{n_j}$ converges and it follows that $\phi_{n_j} = \lambda^{-1} T\phi_{n_j}$ converges. But this is impossible for an orthonormal set.

For the third point suppose λ_n is a sequence of distinct eigenvalues such that $\lambda_n \to \lambda \neq 0$. Choose eigenfunctions ϕ_n such that $T\phi_n = \lambda_n \phi_n$. These are necessarily orthogonal and we may assume they are orthonormal. Then $\lambda_n^{-1}\phi_n$ is bounded and hence $T(\lambda_n^{-1}\phi_n) = \phi_n$ has a convergent subsequence. But this is impossible. □

Problem 1.17 Let T be a bounded operator with a complete set of eigenfunctions $\{\phi_n\}$ and eigenvalues λ_n so $\lambda_n \to 0$ as $n \to \infty$. Show that T is compact.

1.4.2 Hilbert–Schmidt operators

A bounded operator $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ is said to be *Hilbert–Schmidt* if

$$\|T\|_{HS}^2 \equiv \sum_{i=1}^{\infty} \|T\phi_i\|^2 < \infty \tag{1.72}$$

for some orthonormal basis $\{\phi_i\}$ in $\mathcal{H}_1$. The condition is independent of basis since if $\{\psi_j\}$ is an orthonormal basis for $\mathcal{H}_2$, then

$$\sum_i \|T\phi_i\|^2 = \sum_{i,j} |(\psi_j, T\phi_i)|^2 = \sum_{i,j} |(T^*\psi_j, \phi_i)|^2 = \sum_{i,j} \|T^*\psi_j\|^2 \tag{1.73}$$

This also shows that if T is Hilbert–Schmidt, then so is T^*. In fact restricting to the case $\mathcal{H}_1 = \mathcal{H}_2$ we have:

Lemma 1.4 *The Hilbert–Schmidt operators on $\mathcal{H}$ are a $*$-ideal in $\mathcal{B}(\mathcal{H})$:*

1. *If T, S are Hilbert–Schmidt, then so is $aT + bS$ for $a, b \in \mathbb{C}$.*
2. *If T is Hilbert–Schmidt and S is bounded, then TS and ST are Hilbert–Schmidt.*
3. *If T is Hilbert–Schmidt, then so is T^*.*

The proof is straightforward.

Lemma 1.5 *A Hilbert–Schmidt operator is compact.*

Proof Given a Hilbert–Schmidt operator T, let $\{\psi_j\}$ be an orthonormal basis for $\mathcal{H}_2$. Then

$$Tf = \sum_{j=1}^{\infty} (\psi_j, Tf)\psi_j \tag{1.74}$$

We also define

$$T_n f = \sum_{j=1}^{n} (\psi_j, Tf)\psi_j \tag{1.75}$$

Each T_n has finite-dimensional range and hence is compact. We have

$$\|(T - T_n)f\|^2 = \sum_{j=n+1}^{\infty} |(\psi_j, Tf)|^2 \leq \left(\sum_{j=n+1}^{\infty} \|T^*\psi_j\|^2 \right) \|f\|^2 \tag{1.76}$$

Hence

$$\|T - T_n\|^2 \leq \left(\sum_{j=n+1}^{\infty} \|T^* \psi_j\|^2 \right) \to 0 \tag{1.77}$$

as $n \to \infty$. Since T is a norm limit of compacts, it is compact by lemma 1.1. □

Now suppose T is a bounded operator on a Hilbert space $\mathcal{H}$, which has a basis of eigenfunctions $\{\phi_i\}$ with eigenvalues λ_i. Taking this basis in (1.72) we see that T is Hilbert–Schmidt iff

$$\sum_i |\lambda_i|^2 < \infty \tag{1.78}$$

Here is another test for Hilbert–Schmidt.

Lemma 1.6 *Let (M, μ) be a measure space and suppose $k(x, y)$ is an element of $L^2(M \times M, \mu \times \mu)$. Then*

$$(Kf)(x) = \int k(x, y) f(y) d\mu(y) \tag{1.79}$$

defines a Hilbert–Schmidt operator on $L^2(M, \mu)$ and

$$\|K\|_{HS} = \|k\|_2 \tag{1.80}$$

Proof We have seen in problem 1.1 that K is a bounded operator. Let $\{\phi_j\}$ be an orthonormal basis for $L^2(M, \mu)$. Then $\bar{\phi}_i \otimes \phi_j$ is an orthonormal basis for $L^2(M \times M, \mu \times \mu)$; see the proof of theorem B.1 in the appendix. We compute

$$\begin{aligned} \|K\|_{HS}^2 &= \sum_j \|K\phi_j\|^2 = \sum_{i,j} |(\phi_i, K\phi_j)|^2 \\ &= \sum_{i,j} |(k, \bar{\phi}_i \otimes \phi_j)|^2 = \|k\|_2^2 \end{aligned} \tag{1.81}$$

□

1.4.3 Trace class

Let T be a bounded positive self-adjoint operator on a Hilbert space $\mathcal{H}$. (Recall that positive means $(f, Tf) \geq 0$ for all $f \in \mathcal{H}$.) We say that T is *trace class* if

$$Tr(T) = \sum_i (\phi_i, T\phi_i) < \infty \tag{1.82}$$

for some orthonormal basis $\{\phi_i\}$. By the spectral theorem, $T = V^{-1}[\tau]V$ for a function τ satisfying $\tau \geq 0$ almost everywhere. Then T has a positive self-adjoint square root $T^{1/2} = V^{-1}[\tau^{1/2}]V$ and the condition is equivalent to the statement that $T^{1/2}$ is Hilbert–Schmidt. It follows that the sum is independent of the basis. If T

has a complete set of eigenfunctions $\{\phi_i\}$ with nonnegative eigenvalues λ_i, then the condition for trace class is equivalent to

$$\sum_i \lambda_i < \infty \tag{1.83}$$

More generally if T is a bounded operator on $\mathcal{H}$, then T^*T is a positive operator and we say that T is trace class if $|T| = (T^*T)^{1/2}$ is trace class. This reduces to the previous definition if T is positive self-adjoint.

To investigate this concept we need to know more about the relation between T and $|T|$. A bounded operator U is a *partial isometry* if it is an isometry when restricted to $(\text{Ker } U)^\perp$.

Lemma 1.7 *(Polar decomposition) If T is a bounded operator, there are partial isometries U, V such that $T = U|T|$ and $|T| = VT$.*

Proof Define $U : \text{Ran}(|T|) \to \text{Ran}(T)$ by $U(|T|f) = Tf$. Since

$$\||T|f\|^2 = (f, |T|^2 f) = (f, T^*Tf) = \|Tf\|^2 \tag{1.84}$$

this is well-defined (that is $|T|f = |T|g$ implies $Tf = Tg$) and norm preserving. It extends to an isometry from $\overline{\text{Ran}|T|}$ to $\overline{\text{Ran}T}$. Define U to be zero on $(\text{Ran}|T|)^\perp = \text{Ker}|T| = \text{Ker } T$. Then U is a partial isometry and $U|T| = T$. For V, reverse the roles of T and $|T|$. □

Lemma 1.8 *A bounded operator is trace class iff it is the product of two Hilbert–Schmidt operators. In particular a trace class operator is Hilbert–Schmidt and hence compact.*

Proof If T is trace class, then $T = (U|T|^{1/2})(|T|^{1/2})$ exhibits the operator as a product of Hilbert–Schmidt operators. On the other hand suppose $T = A^*B$ with A, B Hilbert–Schmidt. First if T is positive, then

$$\begin{aligned} \sum_i (\phi_i, T\phi_i) = \sum_i (A\phi_i, B\phi_i) &\le \sum_i \|A\phi_i\| \|B\phi_i\| \\ &\le \left(\sum_i \|A\phi_i\|^2 \right)^{1/2} \left(\sum_i \|B\phi_i\|^2 \right)^{1/2} < \infty \end{aligned} \tag{1.85}$$

and hence T is trace class. In the general case if $T = A^*B$, then $|T| = (VA^*)(B)$ exhibits $|T|$ as a product of two Hilbert–Schmidts. Thus $|T|$ is trace class and hence T is trace class. □

Lemma 1.9 *The trace class operators are a $*$-ideal in $\mathcal{B}(\mathcal{H})$:*

1. *If T, S are trace class, then so is $aT + bS$ for $a, b \in \mathbb{C}$.*
2. *If T is trace class and S is bounded, then TS and ST are trace class.*
3. *If T is trace class, then so is T^*.*

Proof The only tricky part is showing that $S+T$ is again trace class. To see this write $S = A_1^* B_1$ and $T = A_2^* B_2$ with A_i, B_i Hilbert–Schmidt. Then $A = A_1 \oplus A_2$ and $B = B_1 \oplus B_2$ are Hilbert–Schmidt operators from $\mathcal{H}$ to $\mathcal{H} \oplus \mathcal{H}$ and $T+S = A^*B$. This is sufficient to conclude that $T+S$ is trace class as in the proof of lemma 1.8. □

If T is trace class but not positive, we again define the trace by (1.82). This makes sense because:

Lemma 1.10 *Let T be trace class.*

1. *$Tr(T) \equiv \sum_i (\phi_i, T\phi_i)$ is absolutely convergent and independent of the choice of basis.*
2. *Tr is a linear functional on the trace class operators.*
3. *(cyclicity) If $T = AB$ with A, B Hilbert–Schmidt, then*

$$Tr(AB) = Tr(BA) \tag{1.86}$$

Proof Write $T = A^*B$ as a product of Hilbert–Schmidt operators. Then the absolute convergence of the sum is demonstrated as above in (1.85). If $\{\psi_j\}$ is another basis, we have

$$\begin{aligned} Tr(A^*B) &= \sum_i (A\phi_i, B\phi_i) = \sum_{i,j} (A\phi_i, \psi_j)(\psi_j, B\phi_i) \\ &= \sum_{i,j} (B^*\psi_j, \phi_i)(\phi_i, A^*\psi_j) = \sum_j (B^*\psi_j, A^*\psi_j) = Tr(BA^*) \end{aligned} \tag{1.87}$$

This shows the independence of the basis and also establishes the cyclicity. □

Problem 1.18 Show that if T is trace class and S is bounded, then $Tr(ST) = Tr(TS)$.

Notes on chapter 1: A good general reference is the four volume set Reed and Simon (1980), Reed and Simon (1975), Reed and Simon (1979), Reed and Simon (1978). Other books that cover this material are Kato (1966), Yosida (1966), and Taylor (1996).

2 Classical mechanics

2.1 Hamiltonian mechanics

We start by reviewing some classical physics, specifically mechanics. Classical mechanics is the study of the motions of macroscopic bodies under the influence of certain specified forces. Mathematically it is formulated in terms of ordinary differential equations. These equations can be presented in one of three general forms: Newtonian, Hamiltonian, or Lagrangian. Here we emphasize the Hamiltonian form, which is most easily connected with quantum mechanics.

In Hamiltonian mechanics the states of the system are specified by points in a phase space which we take to be $\mathcal{P} = \mathcal{O} \times \mathbb{R}^n$ for some integer n and some open set $\mathcal{O}$ in $\mathbb{R}^n$. (More generally the phase space is a $2n$-dimensional manifold.) Points in the phase space have the form (x, p) where the point $x \in \mathcal{O}$ describes the location or configuration of the various objects in the system and the point $p \in \mathbb{R}^n$ describes the momenta of the various objects in the system.

The evolution of the system in time is specified by a function $(x(t), p(t))$ from (an interval in) $\mathbb{R}$ to $\mathcal{P}$ called a trajectory. The fundamental dynamical principle is that the allowed trajectories obey Hamilton's equations. These are a system of $2n$ ordinary differential equations of the form

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial p_i} \\ \frac{dp_i}{dt} &= -\frac{\partial H}{\partial x_i} \end{aligned} \tag{2.1}$$

where $H = H(x, p)$ is a function on phase space called the *Hamiltonian*. The choice of the Hamiltonian depends on the system we are trying to describe, and we will see a number of examples shortly.

An immediate advantage of formulating the dynamics in this way is that the Hamiltonian $H(t) = H(x(t), p(t))$ is constant in time. Indeed we have

$$\frac{dH}{dt} = \sum_{i=1}^{n} \frac{\partial H}{\partial x_i}\frac{dx_i}{dt} + \frac{\partial H}{\partial p_i}\frac{dp_i}{dt}$$

$$\begin{aligned} &= \sum_{i=1}^{n} \frac{\partial H}{\partial x_i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial x_i} \\ &= 0 \end{aligned} \tag{2.2}$$

The usual interpretation is that the Hamiltonian is the energy of the system, and this property represents conservation of energy.

More generally let $F(p, q)$ be an arbitrary smooth function on the phase space, sometimes called a "classical observable." Let $(x(t), p(t))$ be a solution of Hamilton's equations, and let $F(t) = F(x(t), p(t))$ be the time evolution of the quantity F. Then we have

$$\begin{aligned} \frac{dF}{dt} &= \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}\frac{dx_i}{dt} + \frac{\partial F}{\partial p_i}\frac{dp_i}{dt} \\ &= \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}\frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial x_i} \end{aligned} \tag{2.3}$$

We write this as

$$\frac{dF}{dt} = \{F, H\} \tag{2.4}$$

where the right side is evaluated at $(x(t), p(t))$ and where we define the *Poisson bracket* of F and G to be the function on $\mathcal{P}$

$$\{F, G\} = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial x_i} \tag{2.5}$$

For future reference we note that the Poisson bracket is anti-symmetric and satisfies the Jacobi identity

$$\begin{aligned} \{F, G\} + \{G, F\} &= 0 \\ \{\{F, G\}, H\} + \{\{H, F\}, G\} + \{\{G, H\}, F\} &= 0 \end{aligned} \tag{2.6}$$

Also the coordinate functions satisfy

$$\begin{aligned} \{x_i, x_j\} &= 0 \\ \{p_i, p_j\} &= 0 \\ \{x_i, p_j\} &= \delta_{ij} \end{aligned} \tag{2.7}$$

where $\delta_{ij} = 1$ if $i = j$ and is zero otherwise.

2.2 Examples

Example 2.1 **Single particle in an external field** We consider a single particle in an external force field. The particle is considered small enough that its state can

be described by its position which is a point $x \in \mathbb{R}^d$. The force is modeled by a vector field, that is a function $F : \mathcal{O} \to \mathbb{R}^d$. The particle at x feels a force $F(x)$. The time evolution of the system is given by Newton's second law

$$m\frac{d^2x}{dt^2} = F(x) \tag{2.8}$$

Suppose further that F is a gradient, that is $F = -\nabla V$ for some function $V : \mathcal{O} \to \mathbb{R}$. We say that the force is conservative and call V a potential for the problem. For example in $\mathbb{R}^3$ if the particle has charge q_1 and the force is the electrostatic force due to another particle of charge q_2 at the origin, then $F(x) = q_1q_2x/|x|^3$ and $V(x) = q_1q_2/|x|$. For a conservative force we can write Newton's law as a Hamiltonian system. We define $p = m(dx/dt)$ and then (2.8) is equivalent to

$$\begin{aligned} \frac{dx}{dt} &= \frac{p}{m} \\ \frac{dp}{dt} &= -\nabla V \end{aligned} \tag{2.9}$$

This is a Hamiltonian system on $\mathcal{P} = \mathcal{O} \times \mathbb{R}^d$ with

$$H(x,p) = \frac{|p|^2}{2m} + V(x) \tag{2.10}$$

Example 2.2 Single particle in electric and magnetic fields In $\mathcal{O} \subset \mathbb{R}^3$ suppose that the magnetic field B is static, $\partial B/\partial t = 0$. Then Maxwell's third equation (0.3) says that $\nabla \times E = 0$. If $\mathcal{O}$ is simply connected, it follows that there exists a scalar function Φ called the electrostatic potential so that $E = -\nabla\Phi$. Furthermore Maxwell's second equation says that $\nabla \cdot B = 0$ and it follows that there is a vector field A called the the magnetic potential so that $B = \nabla \times A$. (These potentials are not unique. Indeed we could replace A by $A + \nabla\lambda$ for any scalar function λ. This is called a gauge transformation, about which more later.)

For a single particle of charge e in such a field the Lorentz force equation (0.2) becomes

$$m\frac{d^2x}{dt^2} = -e\nabla\Phi + \frac{e}{c}\left(\frac{dx}{dt} \times \nabla \times A\right) \tag{2.11}$$

This can be written as the system of equations ($r, s = 1, 2, 3$)

$$\begin{aligned} \frac{dx_r}{dt} &= \left(p_r - \frac{e}{c}A_r\right)/m \\ \frac{dp_r}{dt} &= \frac{e}{mc}\sum_s \left(p_s - \frac{e}{c}A_s\right)\frac{\partial A_s}{\partial x_r} - e\frac{\partial\Phi}{\partial x_r} \end{aligned} \tag{2.12}$$

This is a Hamiltonian system on $\mathcal{P} = \mathcal{O} \times \mathbb{R}^3$ with

$$H(x,p) = \frac{1}{2m}\left|p - \frac{e}{c}A(x)\right|^2 + e\Phi(x) \tag{2.13}$$

Note that if $A = 0$, then this example reduces to the previous example with $V = e\Phi$.

Example 2.3 **Many particles** Next consider a collection of n particles in $\mathbb{R}^d$ interacting with each other. The location of the particles is given by a point $(x_1, ..., x_n) \in \mathbb{R}^{nd}$ and the momenta is given by a point $(p_1, ..., p_n) \in \mathbb{R}^{nd}$. The phase space is $\mathbb{R}^{nd} \times \mathbb{R}^{nd}$.

The force of the jth particle on the ith particle is assumed to depend only on the relative positions and have the form $F(x_i - x_j)$ for some force F as in the first example. We assume that $F(-x) = -F(x)$ so that the force of the ith particle on the jth particle is minus the force of the jth particle on the ith particle (Newton's third law). Newton's equations for this problem take the form

$$m_i \frac{d^2 x_i}{dt^2} = \sum_{j \neq i} F(x_i - x_j) \tag{2.14}$$

where m_i is the mass of the ith particle.

If we further assume that the force F is conservative with $F = -\nabla V$ and $V(-x) = (x)$, then this equation can be written

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{p_i}{m_i} \\ \frac{dp_i}{dt} &= -\nabla_i \left(\sum_{j \neq i} V(x_i - x_j) \right) \end{aligned} \tag{2.15}$$

This is a Hamiltonian system with

$$H(x_1, \ldots, x_n, p_1, \ldots, p_n) = \sum_i \frac{|p_i|^2}{2m_i} + \frac{1}{2} \sum_{j \neq i} V(x_i - x_j) \tag{2.16}$$

2.3 Canonical transformations

Let $\xi = (x,p)$ be a point in a phase space $\mathcal{P} = \mathcal{O} \times \mathbb{R}^n$ and let J be the $2n \times 2n$ matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{2.17}$$

Then Hamilton's equations can be written in the form

$$\frac{d\xi}{dt} = J\nabla H \tag{2.18}$$

We want to investigate transformations which preserve the form of this equation.

First define a $2n \times 2n$ matrix M to be *symplectic* if $M^T JM = J$. Products of symplectic matrices are symplectic. Symplectic matrices are non-singular since $|\det M| = 1$, and the inverse is also symplectic. Thus symplectic matrices form a group. The inverse and the transpose are related by $M^{-1} = -JM^T J$ or $M^T = -JM^{-1}J$. Thus M^T is symplectic as well.

A smooth mapping $\xi' = \phi(\xi)$ is a *canonical transformation* if the derivative $D\phi = \{d\xi_i'/d\xi_j\}$ is symplectic, that is if

$$(D\phi)^T J(D\phi) = J \tag{2.19}$$

Since $|\det D\phi| = 1$, a canonical transformation is volume preserving. By the inverse function theorem, a canonical transformation is at least locally invertible, and can be thought of as a change of coordinates in phase space.

Canonical transformations preserve the form of Hamilton's equations as the following result shows:

Theorem 2.1 *Let ϕ be a canonical transformation and suppose $\xi(t)$ solves Hamilton's equations (2.18). Then the transformed solution $\xi'(t) = \phi(\xi(t))$ solves*

$$d\xi'/dt = J\nabla H' \tag{2.20}$$

where $H' = H \circ \phi^{-1}$.

Proof Since $H = H' \circ \phi$, we have

$$\nabla H = (\nabla H' \circ \phi)D\phi = (D\phi)^T(\nabla H' \circ \phi) \tag{2.21}$$

Thus for $\xi'(t) = \phi(\xi(t))$ we have

$$\begin{aligned} \frac{d\xi'}{dt} &= (D\phi)(\xi(t))\frac{d\xi}{dt} \\ &= (D\phi)(\xi(t))(J\nabla H)(\xi(t)) \\ &= ((D\phi)J(D\phi)^T)(\xi(t))\nabla H'(\xi'(t)) \\ &= J\nabla H'(\xi'(t)) \end{aligned} \tag{2.22}$$

where we use that $(D\phi)^T$ is symplectic. □

Theorem 2.2 *Let $\phi^* F = F \circ \phi$ be the pull-back of F. A smooth function ϕ is canonical iff*

$$\phi^*\{F, G\} = \{\phi^* F, \phi^* G\} \tag{2.23}$$

for all smooth functions F, G.

Proof We have the representation

$$\{F, G\} = -(\nabla F) J \nabla G \tag{2.24}$$

Since $\nabla(\phi^* G) = (D\phi)^T (\nabla G) \circ \phi$, the equation (2.23) can be written as

$$((\nabla F) J (\nabla G)) \circ \phi = ((\nabla F) \circ \phi)(D\phi) J (D\phi)^T ((\nabla G) \circ \phi) \tag{2.25}$$

If ϕ is canonical, then $(D\phi)J(D\phi)^T = J$ and the identity holds. On the other hand suppose (2.25) holds. Take $F(\xi) = \xi_i, G(\xi) = \xi_j$ so that $(\nabla F)_k = \delta_{ik}, (\nabla G)_k = \delta_{jk}$. Then we get $[(D\phi)J(D\phi)^T]_{ij} = J_{ij}$ so that ϕ is canonical. □

Before continuing we introduce some additional concepts. Let X be a vector field on $\mathcal{P}$, that is a function from $\mathcal{P}$ to $\mathbb{R}^{2n}$. Let ϕ_t be the flow of X. That is $\xi(t) = \phi_t(\xi)$ is the solution of $d\xi/dt = X(\xi)$ starting at ξ and defined for t sufficiently small. For any smooth function F on $\mathcal{P}$ define the *Lie derivative* $\mathcal{L}_X F$ to be the function

$$(\mathcal{L}_X F)(\xi) = \frac{d}{dt} F(\phi_t(\xi))|_{t=0} \tag{2.26}$$

Applying the chain rule in (2.26) we have the representation

$$\mathcal{L}_X F = \sum_i X_i(\xi) \frac{\partial F}{\partial \xi_i} \tag{2.27}$$

This is the vector field X regarded as a differential operator. The flow satisfies $\phi_t \circ \phi_s = \phi_{t+s} = \phi_s \circ \phi_t$, and from this it is straightforward to deduce that $\phi_t^*(\mathcal{L}_X F) = \mathcal{L}_X(\phi_t^* F)$ and that

$$\frac{d}{dt}(\phi_t^* F) = \mathcal{L}_X(\phi_t^* F) \tag{2.28}$$

Now our solution of Hamilton's equations is the flow ϕ_t of the vector field $X_H = J\nabla H$, called a *Hamiltonian vector field*. The evolution equation (2.4) can be written

$$\frac{d}{dt}(\phi_t^* F) = \phi_t^* \{F, H\} \tag{2.29}$$

and specializing to $t = 0$ we have

$$\mathcal{L}_{X_H} F = \{F, H\} \tag{2.30}$$

Then $\phi_t^* \{F, H\} = \{\phi_t^* F, H\}$ and either (2.28) or (2.29) becomes

$$\frac{d}{dt}(\phi_t^* F) = \{\phi_t^* F, H\} \tag{2.31}$$

Theorem 2.3 *The flow ϕ_t of Hamilton's equations is canonical for each t.*

Proof Let $F_t = \phi_t^* F = F \circ \phi_t$ so $dF_t/dt = \{F_t, H\}$. Using the Jacobi identity we compute

$$\begin{aligned} \frac{d}{dt}\{F_t, G_t\} &= \left\{\frac{dF_t}{dt}, G_t\right\} + \left\{F_t, \frac{dG_t}{dt}\right\} \\ &= \{\{F_t, H\}, G_t\} + \{F_t, \{G_t, H\}\} \\ &= \{\{F_t, G_t\}, H\} = \mathcal{L}_{X_H}\{F_t, G_t\} \end{aligned} \tag{2.32}$$

Thus $U(t, \xi) = \{F_t, G_t\}(\xi)$ satisfies the first-order linear partial differential equation with initial condition

$$\left(\frac{\partial}{\partial t} - \mathcal{L}_{X_H}\right) U = 0 \qquad U(0, \xi) = \{F, G\}(\xi) \tag{2.33}$$

This is also satisfied by $U(t, \xi) = \{F, G\}_t(\xi)$. Solutions are unique and thus $\{F, G\}_t = \{F_t, G_t\}$. Since the Poisson bracket is preserved, the flow is canonical. □

Remark Suppose the phase space is the vector space $\mathbb{R}^{2n}$ and suppose that the Hamiltonian flow ϕ_t is linear, that is suppose that the Hamiltonian is a quadratic polynomial. Then $D\phi_t = \phi_t$ and the statement that $D\phi_t$ is symplectic becomes $\phi_t^T J \phi_t = J$. Another way to formulate it is to define a skew-symmetric bilinear form (a symplectic form) by $\sigma(\xi_1, \xi_2) = \xi_1 \cdot J\xi_2$. Then σ is invariant under time evolution: $\sigma(\phi_t \xi_1, \phi_t \xi_2) = \sigma(\xi_1, \xi_2)$. An infinite-dimensional version of this will be of interest when we study quantum field theory.

Problem 2.1 On $\mathbb{R}^2$ consider the Hamiltonian $H(x, p) = \frac{1}{2}(p^2 + \omega^2 x^2)$ where ω is a constant. Find an explicit expression for the flow ϕ_t and verify directly that it is canonical.

2.4 Symmetries

We continue to let ϕ_t be the flow of Hamilton's equations. For any smooth function F on phase space $\mathcal{P}$ we have $d(\phi_t^* F)/dt = \phi_t^* \{F, H\}$ and from this it follows that

$$\phi_t^* F = F \iff \{F, H\} = 0 \tag{2.34}$$

In this case we say that F is a conserved quantity or a constant of the motion.

There may be other Hamiltonian flows occurring naturally in the problem. Suppose G is a smooth function on $\mathcal{P}$ and let ψ_t be the flow of $X_G = J\nabla G$. We say that G is the *generator* of ψ_t. Then $\psi_t^* F = F$ iff $\{F, G\} = 0$. Combining these facts we have

$$\phi_t^* G = G \iff \{G, H\} = 0 \iff \psi_t^* H = H \tag{2.35}$$

We can paraphrase this by saying that a function G is a constant of motion iff the flow ψ_t that it generates leaves the Hamiltonian invariant. This suggests that we look

for constants of the motion by looking for symmetries of the Hamiltonian. We now proceed to look at some special cases.

In our models we will generally have an action of translations on the phase space, that is an action of the additive group $\mathbb{R}^3$. Translations in a particular direction are a Hamiltonian flow and the generator of this flow will be called the *total momentum* in that direction. If the Hamiltonian is invariant under translations in that direction, then the associated total momentum is conserved.

Example 2.4 We continue with a single particle in an external field (example 2.1). The phase space is $\mathcal{P} = \mathbb{R}^3 \times \mathbb{R}^3$ and the Hamiltonian is $H = |p|^2/2m + V(x)$. A translation by $a \in \mathbb{R}^3$ acts on $\mathcal{P}$ by $T_a(x,p) = (x + a, p)$. We consider one-parameter subgroups of the form $T_{tn}(x,p) = (x + tn, p)$ with $|n| = 1$. These are the flow of the vector field $(n, 0)$ and this is a Hamiltonian vector field since it can be written $J\nabla(p \cdot n)$. Thus the total momentum in direction n is the particle momentum $p \cdot n$ in the direction n. The translation takes the Hamiltonian to $H \circ T_{tn} = |p|^2/2m + V(x + tn)$. The Hamiltonian is invariant if V is constant in the direction n, that is there is no force in the direction n. In this case the momentum $p \cdot n$ is conserved.

Example 2.5 Now consider again n particles interacting with each other (example 2.3). The phase space is $\mathcal{P} = \mathbb{R}^{3n} \times \mathbb{R}^{3n}$ and the Hamiltonian is given by (2.16). Translations by $a \in \mathbb{R}^3$ act on $\mathcal{P}$ by

$$T_a(x_1, ..., x_n, p_1, ..., p_n) = (x_1 + a, ..., x_n + a, p_1, ..., p_n) \tag{2.36}$$

and we consider the one-parameter subgroups of the form T_{tn} for $|n| = 1$. The action of T_{tn} is the flow of the vector field $(n, \ldots, n, 0, \ldots, 0)$. This is a Hamiltonian vector field since it can be written $J\nabla(P \cdot n)$ where

$$P = \sum_{i=1}^{n} p_i \tag{2.37}$$

is identified as the total momentum. The Hamiltonian is invariant under translations since the potentials $V(x_i - x_j)$ only depend on the differences $x_i - x_j$. Hence $P \cdot n$ is conserved for any n and so P conserved.

Also we will generally have an action of the rotation group on the phase space. Rotations around a particular axis are Hamiltonian flows and the generator will be called the *total angular momentum* for the system around that axis. If the Hamiltonian is invariant under rotations, then the total angular momentum is conserved.

Before getting into examples we review some facts about the rotation group on $\mathbb{R}^3$. The orthogonal group $O(3)$ is all 3×3 matrices R, preserving lengths or equivalently so that $R^T R = I$. It inherits a topology as a subset of $\mathbb{R}^9$ and is in fact a Lie group, that is a manifold. Such R have $\det R = \pm 1$ and this divides $O(3)$ into two connected components. The component with $\det R = 1$ is a subgroup called the rotation group or special orthogonal group and is denoted $SO(3)$. One-parameter subgroups are rotations about a fixed axis $n \in \mathbb{R}^3$, $|n| = 1$. For example a rotation by an angle θ around the axis $e_3 = (0, 0, 1)$ is

$$R(e_3, \theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.38}$$

Then

$$X_3 \equiv \frac{dR(e_3, \theta)}{d\theta}\Big|_{\theta=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.39}$$

is an element of the Lie algebra (tangent space to the group at the identity). The matrix X_3 is the generator of the subgroup in the sense that $R(e_3, \theta) = \exp(\theta X_3)$. Similarly we have generators X_1, X_2 for the rotations $R(e_1, \theta), R(e_2, \theta)$ around the other axes. They are

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \tag{2.40}$$

The matrices X_1, X_2, X_3 are a basis for the Lie algebra of $SO(3)$, the skew-symmetric matrices. The Lie algebra has a bracket operation inherited from the group structure and determined by the commutators[1]

$$[X_1, X_2] = X_3 \quad [X_2, X_3] = X_1 \quad [X_3, X_1] = X_2 \tag{2.41}$$

Example 2.6 We consider again a single particle in an external field (example 2.4 continued). The phase space is $\mathcal{P} = \mathbb{R}^3 \times \mathbb{R}^3$ and the Hamiltonian is $H = |p|^2/2m + V(x)$. A rotation by $R \in SO(3)$ acts on $\xi = (x, p) \in \mathcal{P}$ by $\hat{R}\xi = (Rx, Rp)$. The one-parameter subgroup $R(e_3, \theta)$ acts by

$$\hat{R}(e_3, \theta)\xi = (R(e_3, \theta)x, R(e_3, \theta)p) \tag{2.42}$$

This is the flow of the vector field

$$\hat{X}_3\xi = (X_3 x, X_3 p) \tag{2.43}$$

[1] $[A, B] = AB - BA$.

This is a Hamiltonian vector field since it has the form $J\nabla L_3$ where

$$L_3 = -\frac{1}{2}(\xi, J\hat{X}_3\xi) = x_1p_2 - x_2p_1 \tag{2.44}$$

The generator L_3 is identified as the angular momentum around e_3. Similarly associated with $R(e_1,\theta), R(e_2,\theta)$ we have

$$L_1 = x_2p_3 - x_3p_2 \qquad L_2 = x_3p_1 - x_1p_3 \tag{2.45}$$

Thus $L = (L_1, L_2, L_3)$ is the cross product $L = x \times p$. For rotations around a unit vector n the angular momentum turns out to be $L \cdot n$.

Is the Hamiltonian invariant under rotations? The term $|p|^2/2m$ is invariant. If the potential is also invariant, $V(Rx) = V(x)$, then the Hamiltonian is invariant, $L \cdot n$ is conserved for any n, and so the vector L is conserved.

For multiparticle systems as in example 2.5 one finds that the total angular momentum is the sum of the angular momenta for the individual particles.

Problem 2.2 On the phase space $\mathbb{R}^4$ consider the Hamiltonian

$$H(x_1,x_2,p_1,p_2) = a(p_1^2 + x_1^2) + b(p_2^2 + x_2^2) + c(p_1p_2 - x_1x_2) \tag{2.46}$$

Find a constant of motion by showing that H is invariant under the rotations

$$\begin{pmatrix} x_1(\theta) \\ p_1(\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ p_1 \end{pmatrix}$$
$$\begin{pmatrix} x_2(\theta) \\ p_2(\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_2 \\ p_2 \end{pmatrix} \tag{2.47}$$

and then finding a generator for this flow.

Notes on chapter 2: For a modern treatment of classical mechanics try Abraham and Marsden (1978), Gallavotti (1983), or Marsden and Ratiu (1994).

A good general reference for mathematical physics is the two volume set by Choquet-Bruhat *et al.* (1977) and Choquet-Bruhat and DeWitt-Morette (1989). In particular they discuss Lie groups. For a more elementary treatment of Lie groups see Miller (1972) or Hall (2003).

3 Quantum mechanics

3.1 Principles of quantum mechanics

We now begin with the main subject of the book: quantum mechanics. Quantum mechanics is a fundamental framework for describing physical phenomena. Although in principle it is needed for all phenomena, its features are especially evident in microscopic phenomena such as the structure of atoms.

The fundamental principle is that some attributes of a physical system, for example the position, cannot be specified exactly, but only by a probability density. Furthermore it is not the probability density itself which is fundamental, but rather a complex valued function ψ called a "probability amplitude" or a "wave function" whose modulus squared $|\psi|^2$ gives the probability density. These ideas are encompassed in the following axiom:

Axiom I: The states of the system are described by vectors of norm one in a complex Hilbert space $\mathcal{H}$.

Actually we do not distinguish states which differ by a phase. Two vectors $\psi_1, \psi_2 \in \mathcal{H}$ are considered as equivalent if $\psi_1 = e^{i\theta}\psi_2$ for some real θ. It is equivalence classes of unit vectors, called *rays*, which describe the states of the system.

As an example for a single particle, say an electron, the Hilbert space would be $\mathcal{H} = L^2(\mathbb{R}^3)$. A particle in the state $\psi \in \mathcal{H}$ with $\|\psi\| = 1$ is located with a probability distribution $|\psi(x)|^2$. In particular the probability of finding the particle in a measurable set $B \subset \mathbb{R}^3$ is $\int_B |\psi(x)|^2 dx$.

As a second example suppose we have two different particles, say a proton and an electron. In this case the Hilbert space would be $\mathcal{H} = L^2(\mathbb{R}^3 \times \mathbb{R}^3)$. If the particles are in a state ψ with $\|\psi\| = 1$, then $\int_{B_1 \times B_2} |\psi(x_1, x_2)|^2 dx_1 dx_2$ is the probability of finding the first particle in the set B_1 and the second particle in the set B_2.

The wave function contains information about all observable quantities, not just the position. This information is extracted according to the following principle.

Axiom II: Properties of physical measurements of a system correspond to projection operators on $\mathcal{H}$. Physically measurable quantities for a system correspond to self-adjoint operators on $\mathcal{H}$.

We elaborate on the meaning of the first point. If the state of the system is ψ and one experimentally tests for a property corresponding to a projection operator P, then the probability that the result of the test is positive is

$$(\psi, P\psi) = \|P\psi\|^2 \tag{3.1}$$

A special case is that one tests whether the the system is in some other state ϕ. In this case the projection operator is the projection onto ϕ which is $P_\phi = \phi(\phi, \cdot)$. (Note that this depends only on the ray.) Thus if the system is in the state ψ, the probability of finding it in the state ϕ is

$$(\psi, P_\phi \psi) = |(\phi, \psi)|^2 \tag{3.2}$$

For the second point suppose that a physically measurable quantity (e.g. position, momentum, etc.) corresponds to a self-adjoint operator A. For short we say A is an *observable*. By the functional calculus, we have an associated family of projection operators $E(B) = \chi_B(A)$ indexed by Borel sets $B \subset \mathbb{R}$. The basic interpretation is that if the system is in the state ψ, then the probability that a measurement of the observable A yields a result in B is

$$(\psi, E(B)\psi) = \|E(B)\psi\|^2 \tag{3.3}$$

These quantities constitute the spectral measures introduced earlier. By the spectral theorem (see problem 1.15)

$$(\psi, A\psi) = \int \lambda \, d(\psi, E(\lambda)\psi) \tag{3.4}$$

As in classical probability theory, $(\psi, A\psi)$ is interpreted as the average value of repeated measurements. It is called the *expectation value* of the observable A.

If we have several commuting self-adjoint operators $A_1, ..., A_n$, then it turns out they have a joint spectral resolution, that is there is a unitary operator which transforms all of them to multiplication operators on the same L^2 space. Then we define projection operators $E(B) = \chi_B(A_1, ..., A_n)$ for $B \subset \mathbb{R}^n$ and $(\psi, E(B)\psi)$ is interpreted as the probability that a simultaneous measurement of observables $A_1, \ldots, A_n$ will yield a value in B. However if self-adjoint operators do not commute, there is no joint spectral resolution and no probability density for simultaneous measurements. Indeed arbitrarily precise simultaneous measurements are not possible. This is a complete departure from classical probability theory.

As an example, for a single particle with Hilbert space $L^2(\mathbb{R}^3)$ let $[x_r]$ be the operator of multiplication by the coordinate x_r. Then for $B \subset \mathbb{R}$ the projection operators are $\chi_B([x_r]) = \chi_B(x_r)$ and the probability of finding the rth coordinate in B is $\int_{x_r \in B} |\psi(x)|^2 dx$. For $B \subset \mathbb{R}^3$ the joint projection operators are $\chi_B([x_1], [x_2], [x_3]) = \chi_B(x_1, x_2, x_3)$ and the probability of finding the particle in B is $\int_B |\psi(x)|^2 dx$ just as before.

Now there is the question of how we associate physically measurable quantities with self-adjoint operators. This is one of the more obscure areas of the subject. However a few general principles suffice to cover most of the situations that arise in practice. One way of systematizing these principles is as follows. We know how to associate classical observables with actual physical attributes. Thus a correspondence between classical observables (that is functions on phase space) and quantum observables (that is self-adjoint operators) serves our purpose. This correspondence is known as "canonical quantization" and is discussed in the next section.

Next we turn to the question of how the system evolves in time. We assume that external influences on the system are independent of time. (Otherwise the following needs modification.)

Axiom III: The time evolution of a system is given by a one-parameter group of unitary operators $U(t)$ on $\mathcal{H}$ such that if $\psi \in \mathcal{H}$ is the state of the system at time zero, then $\psi_t = U(t)\psi$ is the state at time t.

The fact that time evolution is given by a family of linear operators (even for systems which are classically nonlinear) is perhaps mysterious, but once this is accepted we are more or less forced to admit that it is unitary to preserve the probabilistic interpretation.

We have formulated dynamics so that the states evolve in time and operators corresponding to observables are fixed in time. The expectation of an observable given by a self-adjoint operator A in a state ψ at time t is $(\psi_t, A\psi_t)$. This is known as the *Schrödinger picture*. There is also the *Heisenberg picture* in which the operators evolve in time and the states are fixed. For any operator A on the $\mathcal{H}$ we define the operator at time t by $A_t = U(t)^{-1}AU(t)$. Now the expectation of A in the state ψ at time t is $(\psi, A_t\psi)$. This is the same as the Schrödinger picture. The two pictures are equivalent in the sense that they have the same expectation values.

By Stone's theorem (Theorem 1.13) the time evolution $U(t)$ will be generated by a self-adjoint operator H. We write

$$U(t) = e^{-itH/h} \tag{3.5}$$

Here $\hbar$ is a small fundamental constant of nature, which sets the scale on which quantum effects are important. In CGS units it is $\hbar = 1.05 \times 10^{-27} erg \cdot seconds$. In the Schrödinger picture we have $\psi_t = e^{-itH/h}\psi$ and so if $\psi \in D(H)$, the state satisfies the Schrödinger equation,

$$i\hbar\frac{d\psi_t}{dt} = H\psi_t \tag{3.6}$$

In the Heisenberg picture we have $A_t = e^{itH/h}Ae^{-itH/h}$, which satisfies

$$-i\hbar\frac{dA_t}{dt} = [H, A_t] \tag{3.7}$$

on a suitable domain. Note the analogy with the time evolution of observables in a classical Hamiltonian system as in (2.4). Poisson brackets are replaced by commutators and the operator H plays the role of the Hamiltonian. Indeed H is called the (quantum) Hamiltonian and corresponds to the energy of the system. This correspondence is further developed in the framework of canonical quantization in the next section.

In addition to this continuous unitary time evolution the system also changes in a discontinuous way when a measurement is made upon it. Roughly the state jumps to a state which is specified by the results of the measurement. This is known as "reduction of the wave function." The question of which physical processes constitute measurements in this sense is rather unsettled, as well as the question of finding a correct mathematical description. Nevertheless it turns out that one can solve most practical problems without entering into these issues.

This completes our survey of the basic principles in the form of three axioms which list the kind of mathematical structures we are interested in and how they are supposed to model the physical world. They are not meant to be exhaustive or inflexible, but only a general point of reference.

3.2 Canonical quantization

Canonical quantization is a recipe for passing from a classical Hamiltonian system to a quantum mechanical system. We should say at the outset that there are limits to how seriously one should take this procedure. We do not mean to say that the classical system is fundamental and it is somehow modified for microscopic phenomena. Rather it is the quantum system which is fundamental. The classical system is an approximation, which however is excellent for macroscopic bodies. Canonical quantization is just a sophisticated method for guessing the correct quantum description from its classical manifestation. As such it need not be too sharply drawn.

We begin with a classical Hamiltonian system with phase space $\mathcal{P} = \mathbb{R}^n \times \mathbb{R}^n$ with points $(x, p) = (x_1, \dots, x_n, p_1, \dots, p_n)$ and a Hamiltonian $H(x, p)$. The first step of quantization consists of associating with each of the coordinates (unbounded) linear operators $(\hat{x}, \hat{p}) = (\hat{x}_1, \dots, \hat{x}_n, \hat{p}_1, \dots, \hat{p}_n)$ on a Hilbert space $\mathcal{H}$ such that the commutators satisfy the relations

$$\begin{aligned} \left[\hat{x}_i, \hat{x}_j\right] &= 0 \\ \left[\hat{p}_i, \hat{p}_j\right] &= 0 \\ \left[\hat{x}_i, \hat{p}_j\right] &= i\hbar\delta_{ij} \end{aligned} \tag{3.8}$$

The relations (3.8) are known as the *canonical commutation relations* or CCR. Note the analogy with the Poisson brackets of the classical coordinates given in (2.7). Note also the presence of the fundamental constant $\hbar$.

The second step is to substitute $(\hat{x}, \hat{p})$ into the classical Hamiltonian $H(x, p)$ to form the quantum Hamiltonian $\hat{H} = H(\hat{x}, \hat{p})$. This is just a formal expression which we must give a meaning as a self-adjoint operator. The time evolution operator is then taken to be $U(t) = e^{-it\hat{H}/h}$ in accordance with our general principles. In the Heisenberg picture we then define canonical operators $(\hat{x}_t, \hat{p}_t)$ at time t by

$$\begin{aligned} \hat{x}_t &= e^{it\hat{H}/\hbar}\hat{x}e^{-it\hat{H}/h} \\ \hat{p}_t &= e^{it\hat{H}/\hbar}\hat{p}e^{-it\hat{H}/h} \end{aligned} \tag{3.9}$$

These also satisfy the CCR and obey the equations

$$\begin{aligned} -i\hbar\frac{d}{dt}\hat{x}_t &= [\hat{H}, \hat{x}_t] \\ -i\hbar\frac{d}{dt}\hat{p}_t &= [\hat{H}, \hat{p}_t] \end{aligned} \tag{3.10}$$

Example 3.1 As an example consider the quantization of the classical system in example 2.1. This is a single particle in $\mathbb{R}^3$ with mass m and under the influence of a potential V. The classical coordinates x_r, p_r with $1 \leq r \leq 3$ can be quantized as the operators $\hat{x}_r, \hat{p}_r$ defined by

$$\begin{aligned} \hat{x}_r &= [x_r] \equiv \text{multiplication by } x_r \\ \hat{p}_r &= -i\hbar\frac{\partial}{\partial x_r} \end{aligned} \tag{3.11}$$

When defined on a suitable domain in $L^2(\mathbb{R}^3)$, say the Schwartz space $\mathcal{S}(\mathbb{R}^3)$, these satisfy the canonical commutation relations (3.8). The classical Hamiltonian $H(x, p) = |p|^2/2m + V(x)$ becomes the quantum Hamiltonian

$$\begin{aligned} \hat{H} &= \frac{|\hat{p}|^2}{2m} + V(\hat{x}) \\ &= -\hbar^2\frac{\Delta}{2m} + [V(x)] \end{aligned} \tag{3.12}$$

In the next chapter we will define $\hat{H}$ as a self-adjoint operator and so specify the dynamics. In any case proceeding formally and using that $[H, A_t] = [H, A]_t$ we find that the equations (3.10) become in this case

$$\begin{aligned} \frac{d}{dt}\hat{x}_{r,t} &= \frac{\hat{p}_{r,t}}{m} \\ \frac{d}{dt}\hat{p}_{r,t} &= -\frac{\partial V}{\partial x_r}(\hat{x}_t) \end{aligned} \tag{3.13}$$

(Here we use the formal identity $[\partial V/\partial x_r]_t = \partial V/\partial x_r(\hat{x}_t)$.) Thus the quantum operators $(\hat{x}_r, \hat{p}_r)$ obey the classical Hamilton's equations. This is a general principle known as Ehrenfest's theorem.

Let us consider the momentum operators $\hat{p}_r$ in more detail. These satisfy on $\mathcal{S}(\mathbb{R}^3)$

$$\hat{p}_r = \mathcal{F}_h^{-1}[p_r]\mathcal{F}_h \tag{3.14}$$

where $\mathcal{F}_h$ is the Fourier transform defined with the exponent $\exp(-ipx/\hbar)$ rather than $\exp(-ipx)$. Since the multiplication operator $[p_r]$ is naturally a self-adjoint operator, this formula can be used to define $\hat{p}_r$ as a self-adjoint operator on $L^2(\mathbb{R}^3)$ with domain

$$D(\hat{p}_r) = \{\psi : \int p_r^2 |\tilde{\psi}(p)|^2 dp < \infty\} \tag{3.15}$$

The joint spectral projections for $\hat{p} = (\hat{p}_1, \hat{p}_2, \hat{p}_3)$ are the operators $E(B) = \mathcal{F}_\hbar^{-1}[\chi_B(p_1, p_2, p_3)]\mathcal{F}_h$ and so the probability of finding the momentum in the set $B \subset \mathbb{R}^3$ in the state ψ is

$$(\psi, E(B)\psi) = \int_B |\tilde{\psi}(p)|^2 dp \tag{3.16}$$

Thus $|\tilde{\psi}(p)|^2$ gives probability density for momentum. We say that the Fourier transform $\tilde{\psi}(p)$ is the wave function in momentum space. This interpretation of the Fourier transform is not special to this example, but pervades quantum physics.

Problem 3.1 In the above example let $\psi \in \mathcal{S}(\mathbb{R}^3)$ with $\|\psi\| = 1$ and define the expectations $\bar{x}_r = (\psi, \hat{x}_r\psi)$ and $\bar{p}_r = (\psi, \hat{p}_r\psi)$. Further define the variances $(\Delta x_r)^2 = (\psi, (\hat{x}_r - \bar{x}_r)^2\psi)$ and $(\Delta p_r)^2 = (\psi, (\hat{p}_r - \bar{p}_r)^2\psi)$. Show that

$$\Delta x_r \Delta p_r \geq \frac{\hbar}{2} \tag{3.17}$$

This result, known as the *Heisenberg uncertainty principle*, shows that position and momentum cannot simultaneously be constrained arbitrarily sharply.

The value of $\hbar$ depends on which system of units we are using. Hereafter we choose units in which $\hbar = 1$ so that $\hbar$ disappears from our equations.

3.3 Symmetries

We want to consider symmetries and conserved quantities for quantum systems. Our discussion parallels the classical discussion of section 2.4, but with canonical

flows replaced by one-parameter unitary groups and Poisson brackets replaced by commutators.

Suppose we have a quantum system with Hilbert space $\mathcal{H}$ and time evolution e^{-iHt}. Let

$$\alpha_t(A) = A_t = e^{iHt} A e^{-iHt} \tag{3.18}$$

be the time evolution of the observable A. Since $-i dA_t/dt = [H, A]_t$, we deduce that

$$\alpha_t(A) = A \iff [H, A] = 0 \tag{3.19}$$

Thus A is constant in time iff it commutes with H. We say that A represents a conserved quantity. To find such quantities we consider other group actions on $\mathcal{H}$.

Suppose e^{-iGt} is another one-parameter unitary group on $\mathcal{H}$ with self-adjoint generator G. The action on an observable A is

$$\beta_t(A) = e^{iGt} A e^{-iGt} \tag{3.20}$$

and the observable is invariant iff $[G, A] = 0$. Combining the above we have

$$\alpha_t(G) = G \iff [H, G] = 0 \iff \beta_t(H) = H \tag{3.21}$$

Thus an observable is conserved iff it generates a symmetry of the Hamiltonian.

Now we get more specific. The translation group $\mathbb{R}^3$ acts on $\mathbb{R}^3$ by $x \to x + a$. We suppose that we have a continuous representation of this group by unitary operators $U(a)$ on $\mathcal{H}$, that is $U(a)U(a') = U(a + a')$. For any state ψ the state $U(a)\psi$ is interpreted as the state translated by a. For any basis vector e_r, $U(te_r)$ is a one-parameter unitary group and by Stone's theorem there is a self-adjoint operator P_r such that

$$U(te_r) = \exp(-iP_r t) \tag{3.22}$$

Then P_r is interpreted as the rth component of the *total momentum* of the system, in analogy with the classical situation. If the Hamiltonian is invariant under translations in the direction e_r, then P_r is conserved.

We also suppose that we have a continuous representation of the rotation group, that is unitary operators $U(R)$ on $\mathcal{H}$ for each $R \in SO(3)$ such that $U(R)U(R') = U(RR')$. For a state ψ the state $U(R)\psi$ is the state rotated by R. If $R(e_r, \theta)$ is the rotation by θ around a basis vector e_r, then $U(R(e_r, \theta))$ is a one-parameter unitary group and by Stone's theorem there is a self-adjoint operator denoted J_r such that

$$U(R(e_r, \theta)) = \exp(-i\theta J_r) \tag{3.23}$$

Then J_r is interpreted as the r-component of the *total angular momentum* of the system, again in analogy with the classical situation. If the Hamiltonian is invariant under rotations around e_r, then J_r is conserved.

Example 3.2 We continue the example of a single particle in an external potential, example 3.1. The translation group on $L^2(\mathbb{R}^3)$ is represented by

$$(U(a)\psi)(x) = \psi(x-a) \tag{3.24}$$

The generator in the direction e_r is for suitable ψ

$$(P_r\psi)(x) = i\frac{d}{dt}(U(te_r)\psi)(x)|_{t=0} = -i\frac{\partial\psi}{\partial x_r}(x) \tag{3.25}$$

Thus the total momentum P_r is just the momentum $\hat{p}_r$ of the single particle as we might expect. The Hamiltonian H is invariant under $U(te_r)$ if the potential satisfies $V(x+te_r) = V(x)$. In this case P_r is conserved.

The rotation group is represented by

$$(U(R)\psi)(x) = \psi(R^{-1}x) \tag{3.26}$$

which is unitary since $\det R = 1$. The third component of the angular momentum, denoted L_3 rather than J_3 in this example, is

$$(L_3\psi)(x) = i\frac{d}{d\theta}\psi(R(e_3,\theta)^{-1}x)|_{\theta=0} = \left(x_1\left(-i\frac{\partial}{\partial x_2}\right) - x_2\left(-i\frac{\partial}{\partial x_1}\right)\right)\psi(x) \tag{3.27}$$

If H is invariant under $R(e_3,\theta)$, that is if $V(R(e_3,\theta)x) = V(x)$, then L_3 is conserved. Rotations around other axes are treated similarly.

Note that $L_3 = \hat{x}_1\hat{p}_2 - \hat{x}_2\hat{p}_1$. Thus we can get the quantum angular momentum by making the substitution $x_r \to \hat{x}_r, p_r \to \hat{p}_r$ in the classical angular momentum (2.44), just as we did with the Hamiltonian. However it has proved difficult to elevate this substitution rule into a general principle for generating quantum observables. Canonical quantization is not a universal recipe.

Problem 3.2 It is generally true that $-iJ_r$ give a representation of the Lie algebra of $SO(3)$ as in (2.41), that is

$$[J_1,J_2] = iJ_3 \quad [J_2,J_3] = iJ_1 \quad [J_3,J_1] = iJ_2 \tag{3.28}$$

Check that this is true for $J_r = L_r$ in the above example.

3.4 Perspectives and problems

If one wants to give a quantum mechanical model for a physical system one proceeds as follows. First one selects a Hilbert space of states $\mathcal{H}$ and a Hamiltonian H. Possibly this would be by applying the method of canonical quantization to a classical model or possibly by experience or guesswork. This is the job of the physicist.

Then there are three mathematical problems:

1. (Self-adjointness) The Hamiltonian will typically be given as a formal operator and the first task, at least for a mathematician, is to give it a rigorous meaning as a self-adjoint operator. Then one has the existence of the time evolution operator $U(t) = e^{-itH}$.
2. (Spectrum) The next task is to study the spectrum of H. In particular one looks for eigenvalues $H\psi = E\psi$. These are the states of definite energy. Such a state would evolve in time according to $\psi_t = e^{-iEt}\psi$. Since the phase factor e^{-iEt} does not change the ray, the state is stationary. These are the states one would look for in nature. Differences in energy levels can often be observed directly since if the system changes from one state to another, it usually emits light with exactly the energy difference.
3. (Scattering) For states which are not eigenvectors or linear combinations of eigenvectors, we ask instead for the long time behavior of the state. Thus we ask for the asymptotic behavior of $e^{-iHt}\psi$ as $t \to \pm\infty$. This leads to the scattering problem: given a state with specified asymptotic behavior as $t \to -\infty$, find the asymptotic behavior as $t \to \infty$.

In the next chapter we take up these problems for the case of a single particle in an external potential.

Notes on chapter 3: The original mathematical treatment of quantum mechanics was von Neumann (1955). Other books on the mathematical foundations are Jauch (1968) and Isham (1995).

4 Single particle

4.1 Free particle

We start with the case of a single particle with no forces – a free particle. To begin we work in $\mathbb{R}^d$ so the Hilbert space is $L^2(\mathbb{R}^d)$ and the Hamiltonian from (3.12) with $V = 0$ is

$$H_0 = -\frac{\Delta}{2m} \tag{4.1}$$

Our first task is to choose a domain for this operator so it is self-adjoint. On $\mathcal{S}(\mathbb{R}^d)$ we have

$$H_0 = \mathcal{F}^{-1}\left[\frac{|p|^2}{2m}\right]\mathcal{F} \tag{4.2}$$

The multiplication operator $|p|^2/2m$ has a natural domain of self-adjointness and we just define $D(H_0)$ to be the transform of this domain. Thus H_0 is defined by (4.2) with domain

$$D(H_0) = \{\psi : \int |p|^4 |\tilde{\psi}(p)|^2 dp < \infty\} \tag{4.3}$$

As the unitary transform of a self-adjoint operator, it is self-adjoint. The operator is defined by its spectral representation.

Time evolution as defined by the spectral theorem is given by

$$e^{-iH_0 t}\psi = \mathcal{F}^{-1}[e^{-i(|p|^2/2m)t}]\mathcal{F} \tag{4.4}$$

or if $\psi \in \mathcal{S}(\mathbb{R}^d)$

$$(e^{-iH_0 t}\psi)(x) = (2\pi)^{-d/2} \int e^{ipx} e^{-i(|p|^2/2m)t} \tilde{\psi}(p) dp \tag{4.5}$$

Lemma 4.1 *For $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $t \neq 0$*

$$(e^{-iH_0 t}\psi)(x) = \left(\frac{m}{2\pi i t}\right)^{d/2} \int e^{i|x-y|^2 m/2t} \psi(y) dy \tag{4.6}$$

and so as $|t| \to \infty$

$$\|e^{-iH_0 t}\psi\|_\infty \leq \mathcal{O}(|t|^{-(d/2)}) \tag{4.7}$$

Remark The estimate says that amplitude of the wave function goes to zero at $t \to \pm\infty$. Since the L^2 norm is conserved, this means that the region of concentration of the wave function must increase. This loss of localization is known as *spreading of the wave function*.

Proof Suppose instead of the Schrödinger equation $d\psi/dt = -iH_0\psi$ we were solving the heat equation $d\psi/d\tau = -H_0\psi$. Then we would have as the solution for $\tau > 0$

$$\begin{aligned}(e^{-H_0\tau}\psi)(x) &= (2\pi)^{-d/2}\int e^{ipx}e^{-(|p|^2/2m)\tau}\tilde{\psi}(p)dp \\ &= \Big(\frac{m}{2\pi\tau}\Big)^{d/2}\int e^{-|x-y|^2m/2\tau}\psi(y)dy\end{aligned} \tag{4.8}$$

Here in the second step we insert the definition of $\tilde{\psi}(p)$ and do the integral over p, see problem 1.3. This calculation holds equally well for τ complex, Re $\tau > 0$. Now $\exp(-H_0(\epsilon + it))\psi$ converges to $\exp(-iH_0t))\psi$ in $L^2(\mathbb{R}^d)$ as $\epsilon \to 0$. Hence we have pointwise convergence for a subsequence ϵ_n. Take $\tau = \epsilon_n + it$ above and let $\epsilon_n \to 0$. Using the dominated convergence theorem on the right we obtain (4.6). (And we also see that the right side is in L^2.) □

Problem 4.1 Define for $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $t \neq 0$

$$(V_t\psi)(x) = \Big(\frac{m}{it}\Big)^{d/2}\exp\Big(i\frac{m|x|^2}{2t}\Big)\tilde{\psi}\Big(\frac{mx}{t}\Big) \tag{4.9}$$

Show that

$$\lim_{t\to\pm\infty}\|e^{-iH_0t}\psi - V_t\psi\|_2 = 0 \tag{4.10}$$

This result shows that in spite of the spreading of the wave function, some localization is preserved. Suppose that $|\tilde{\psi}(p)|$ is peaked at some momentum $p_0 \in \mathbb{R}^d$. Then the asymptotic form of $|(e^{-iH_0t}\psi)(x)|$ is peaked at points where $mx/t = p_0$ or $x = (p_0/m)t$. This special point moves with velocity p_0/m just as for the classical trajectory.

Problem 4.2 Show that the spectrum of H_0 is $[0, \infty)$ and that it is all continuous spectrum.

4.2 Particle in a potential

Now we specialize to $d = 3$ and study a single particle in a potential. As explained in the previous section this means our Hamiltonian has the form

$$H = H_0 + V = \frac{-\Delta}{2m} + V \tag{4.11}$$

where V is the potential function. The first task is to define it as a self-adjoint operator. For this we need:

Theorem 4.1 *(Kato's theorem) Let T be a self-adjoint operator on a Hilbert space and let S be symmetric. Suppose that $D(T) \subset D(S)$ and for some constants $0 \leq a < 1$ and $0 \leq b$ and all $f \in D(T)$ we have*

$$\|Sf\| \leq a\|Tf\| + b\|f\| \tag{4.12}$$

Then $T + S$ is self-adjoint on $D(T)$.

Proof The operator $T + S$ is symmetric on $D(T)$ so it suffices to show that $\mathrm{Ran}(T + S \pm i\mu) = \mathcal{H}$ for some $\mu > 0$ by theorem 1.9. (Actually theorem 1.9 is stated for $\mu = 1$, but it holds as well for any μ.)

For any $g \in \mathcal{H}$ we have that $(T \pm i\mu)^{-1}g \in D(T)$ and by the inequality

$$\|S(T \pm i\mu)^{-1}g\| \leq a\|T(T \pm i\mu)^{-1}g\| + b\|(T \pm i\mu)^{-1}g\| \tag{4.13}$$

On the right side we have certain bounded functions of T, which we estimate by the spectral theorem using $\|h(T)\| \leq \|h\|_\infty$. In particular

$$\begin{aligned} \|T(T \pm i\mu)^{-1}\| &\leq \sup_{\lambda \in \mathbb{R}} |\lambda(\lambda \pm i\mu)^{-1}| \leq 1 \\ \|(T \pm i\mu)^{-1}\| &\leq \sup_{\lambda \in \mathbb{R}} |(\lambda \pm i\mu)^{-1}| \leq \mu^{-1} \end{aligned} \tag{4.14}$$

Therefore

$$\|S(T \pm i\mu)^{-1}g\| \leq \left(a + \frac{b}{\mu}\right)\|g\| \tag{4.15}$$

Since $a + b/\mu < 1$ for μ sufficiently large, we conclude that $\|S(T \pm i\mu)^{-1}\| < 1$. Hence by theorem 1.2

$$\mathrm{Ran}(I + S(T \pm i\mu)^{-1}) = \mathcal{H} \tag{4.16}$$

Now for any $f \in D(T)$ we have

$$(T + S \pm i\mu)f = (I + S(T \pm i\mu)^{-1})(T \pm i\mu)f \tag{4.17}$$

On the right side we have the composition of two surjective operators, hence $T + S \pm i\mu$ is surjective as required. □

Theorem 4.2 *$H = H_0 + V$ on $L^2(\mathbb{R}^3)$ is self-adjoint on $D(H_0)$ in any of the following circumstances:*

1. $V \in L^\infty(\mathbb{R}^3)$
2. $V \in L^2(\mathbb{R}^3)$
3. $V = V_1 + V_2$ *where* $V_1 \in L^2(\mathbb{R}^3)$ *and* $V_2 \in L^\infty(\mathbb{R}^3)$.

Proof Take $m = 1/2$ for simplicity. If $V \in L^\infty$, then it determines a bounded operator and $\|Vf\| \leq \|V\|_\infty\|f\|$. The hypotheses of Kato's theorem are satisfied with $a = 0, b = \|V\|_\infty$, hence the result.

Now suppose $V \in L^2(\mathbb{R}^3)$. If $\psi \in D(H_0)$, then $|p|^2\tilde{\psi}(p)$ is in $L^2(\mathbb{R}^3)$. We write for any $\alpha > 0$

$$\tilde{\psi}(p) = (\alpha^2 + |p|^2)^{-1}\big((\alpha^2 + |p|^2)\tilde{\psi}(p)\big) \tag{4.18}$$

This exhibits $\tilde{\psi}(p)$ as the product of two L^2 functions and hence it is in L^1 as well as L^2. Hence the Fourier inversion formula $\psi(x) = (2\pi)^{-3/2}\int e^{ipx}\tilde{\psi}(p)dp$ holds pointwise. By a Schwarz inequality we get[1]

$$\begin{aligned} |\psi(x)| &\le (2\pi)^{-3/2}\left(\int (\alpha^2 + |p|^2)^{-2}dp\right)^{1/2}\left(\int (\alpha^2 + |p|^2)^2|\tilde{\psi}(p)|^2dp\right)^{1/2} \\ &= c\alpha^{-1/2}\|(H_0 + \alpha^2)\psi\| \\ &\le c\alpha^{-1/2}\|H_0\psi\| + c\alpha^{3/2}\|\psi\| \end{aligned} \tag{4.19}$$

for some constant c. Hence $|\psi(x)|$ is bounded and so is in the domain of V. Thus $D(H_0) \subset D(V)$. Furthermore for $\psi \in D(H_0)$ we have

$$\|V\psi\| \le \|V\|_2\|\psi\|_\infty \le \|V\|_2\left(c\alpha^{-1/2}\|H_0\psi\| + c\alpha^{3/2}\|\psi\|\right) \tag{4.20}$$

For α sufficiently large Kato's inequality holds and we conclude that H is self-adjoint on $D(H_0)$.

For the last result treat V_2 as a bounded perturbation of $H_0 + V_1$. □

Example 4.1 Consider the Yukawa potential

$$V(x) = \frac{c}{|x|}e^{-\mu|x|} \tag{4.21}$$

This is supposed to provide a crude model of nuclear forces transmitted by a particle of mass $\mu > 0$. We have $V \in L^2(\mathbb{R}^3)$ and hence $H = H_0 + V$ is self-adjoint on $D(H_0)$ by the theorem.

Example 4.2 Consider the Coulomb potential

$$V(x) = \frac{-e^2}{|x|} \tag{4.22}$$

This is supposed to describe an electron of charge $-e$ in the electrostatic field of a proton of charge e. This is a simple model of the Hydrogen atom. If B is the unit ball, then $V = V\,1_B + V\,1_{B^c}$ exhibits the potential as the sum of an L^2 function and an L^∞ function. Hence $H = H_0 + V$ is self-adjoint on $D(H_0)$ by the theorem.

We also consider a particle in a magnetic field $B = \nabla \times A$. Applying our canonical quantization procedure to the Hamiltonian (2.13) with $\Phi = 0$ we find the quantum Hamiltonian

[1] This is an example of a Sobolev inequality.

$$\begin{aligned} H &= \frac{1}{2m}\left(-i\nabla - \frac{e}{c}A\right)^2 \\ &\equiv H_0 + \frac{e}{2mc}(2iA\cdot\nabla + i\nabla\cdot A) + \frac{e^2}{2mc^2}|A|^2 \end{aligned} \tag{4.23}$$

Problem 4.3 Suppose A and all its first derivatives are bounded functions. Show that the Hamiltonian (4.23) is self-adjoint on $D(H_0)$.

4.3 Spectrum

We study the spectrum of $H = H_0 + V$. First we need a variation of the Riesz–Schauder theorem, theorem 1.15.

A function $F(z)$ from an open set $\mathcal{O} \subset \mathbb{C}$ to a Banach space is said to be *analytic* if the derivative $F'(z) = \lim_{h\to 0}(F(z+h) - F(z))/h$ exists for all $z \in \mathcal{O}$. Analytic functions in this sense enjoy many of the same properties of complex-valued analytic functions, e.g Cauchy's theorem, power series representation, etc.

Theorem 4.3 *(Analytic Fredholm theorem) Let $F(z)$ be an analytic function from a connected open set $\mathcal{O} \subset \mathbb{C}$ to $\mathcal{B}(\mathcal{H})$ such that $F(z)$ is compact for all $z \in \mathcal{O}$. Then one of the following holds:*

1. *$(I + F(z))^{-1}$ does not exist for any $z \in \mathcal{O}$.*
2. *$(I + F(z))^{-1}$ exists except for a discrete set $S \subset \mathcal{O}$ with no limit points in $\mathcal{O}$. For $z \in S$ the operator $F(z)$ has eigenvalue -1 with finite multiplicity.*

For the proof see Reed and Simon (1980: 201).

Problem 4.4 Show that the Riesz–Schauder theorem follows from the analytic Fredholm theorem. (Hint: $(T - z) = -z(I - T/z)$.)

Theorem 4.4 *Let $V \in L^2(\mathbb{R}^3)$ and let $H = H_0 + V$ on $\mathcal{H} = L^2(\mathbb{R}^3)$. Then $\sigma(H) \cap (-\infty, 0)$ is a bounded countable set $E_1 < E_2 < E_3 < \cdots < 0$ (possibly empty) with no limit points except possibly zero. Each E_j is an eigenvalue with finite multiplicity.*

Remark We characterize $\sigma(H) \cap [0, \infty)$ in section 4.5.

Proof H is self-adjoint by theorem 4.2. Consider $E < 0$. On $D(H) = D(H_0)$ we have the identity

$$(H - E) = (I + V(H_0 - E)^{-1})(H_0 - E) \tag{4.24}$$

Since E is in the resolvent set for H_0, we see that E is in the resolvent set for H iff $(I + V(H_0 - E)^{-1})^{-1}$ exists as a bounded operator. Furthermore $(H - E)\psi$ has a nonzero solution in $D(H)$ iff $(I + V(H_0 - E)^{-1})\phi = 0$ has a nonzero solution in $\mathcal{H}$, that is iff $V(H_0 - E)^{-1}$ has eigenvalue -1.

First we show that the spectrum is bounded below. In Kato's inequality $\|V\psi\| \leq a\|H_0\psi\| + b\|\psi\|$ insert $\psi = (H_0 - E)^{-1}\phi$ and obtain

$$\|V(H_0 - E)^{-1}\phi\| \leq \left(a + \frac{b}{|E|}\right)\|\phi\| \tag{4.25}$$

Thus if E is sufficiently negative, $\|V(H_0 - E)^{-1}\| < 1$ and so $(I + V(H_0 - E)^{-1})^{-1}$ exists and E is in the resolvent set for H.

We next note that $V(H_0 - E)^{-1}$ is an analytic function in $\mathbb{C} - [0, \infty)$. Indeed one can compute directly that the derivative is $V(H_0 - E)^{-2}$.

We claim $V(H_0 - E)^{-1}$ is compact for all such E in $\mathbb{C} - [0, \infty)$. It suffices to show the momentum space version $\mathcal{F}V(H_0 - E)^{-1}\mathcal{F}^{-1}$ is compact, and we show it is Hilbert–Schmidt. Since the Fourier transform of a product is the convolution of the transforms, we have

$$\begin{aligned}(\mathcal{F}V(H_0 - E)^{-1}\mathcal{F}^{-1}\psi)(p) &= (2\pi)^{-3/2}(\tilde{V} * (\mathcal{F}(H_0 - E)^{-1}\mathcal{F}^{-1})\psi)(p) \\ &= (2\pi)^{-3/2}\int \tilde{V}(p - q)\left(\frac{|q|^2}{2m} - E\right)^{-1}\psi(q)\,dq \\ &\equiv \int k(p, q)\psi(q)\,dq\end{aligned} \tag{4.26}$$

This is Hilbert–Schmidt since

$$\begin{aligned}\|k\|_2^2 &= (2\pi)^{-3}\int \left|\tilde{V}(p - q)\left(\frac{|q|^2}{2m} - E\right)^{-1}\right|^2 dp\,dq \\ &= (2\pi)^{-3}\|V\|_2^2 \int \left|\frac{|q|^2}{2m} - E\right|^{-2} dq \\ &< \infty\end{aligned} \tag{4.27}$$

Now we can apply the analytic Fredholm theorem for $V(H_0 - E)^{-1}$ in the region $\mathbb{C} - [0, \infty)$. The alternative that $(I + V(H_0 - E)^{-1})^{-1}$ does not exist anywhere is ruled out for E very negative by (4.25). Thus we conclude $V(H_0 - E)^{-1}$ has eigenvalue -1 at a bounded discrete set of points $E_1, E_2, \ldots$ necessarily real and with no accumulation point except possibly zero. Hence H has these eigenvalues. The multiplicity is finite for $V(H_0 - E_j)^{-1}$ and hence the E_j have finite multiplicity as eigenvalues of H. □

Note that the theorem holds in particular for the Yukawa potential, example 4.1.

Problem 4.5 Prove the same result if for every $\epsilon > 0$ there exists a split $V = V_1 + V_2$ such that $V_1 \in L^2$ and $V_2 \in L^\infty$ with $\|V_2\|_\infty < \epsilon$.

Example 4.3 Consider the Coulomb potential of example 4.2 with the Hamiltonian

$$H = \frac{-\Delta}{2m} - \frac{e^2}{|x|} \tag{4.28}$$

The potential $V(x) = -e^2/|x|$ satisfies the conditions of the previous problem as we see by letting B_R be the ball of radius R and writing $V = V\,1_{B_R} + V\,1_{B_R^c}$ with R large depending on ϵ. We conclude that the negative spectrum is purely discrete with finite multiplicity.

Actually for this problem the spectrum can be computed exactly. Solving the differential equation $(H - E)\psi = 0$ by separation of variables one finds that the eigenvalues are

$$E_n = \frac{-me^4}{2n^2} \qquad n = 1, 2, 3, \ldots \tag{4.29}$$

and the dimension of the eigenspace for E_n is n^2. For details see any textbook on quantum mechanics.

These are the energy levels for hydrogen. Differences of these energies determine the frequencies of light emitted by hydrogen, something which is experimentally observable. The success of these predictions was one of the original triumphs of quantum mechanics.

Problem 4.6 Let V be the rank-one operator $V\psi = \lambda\chi(\chi, \psi)$ where λ is real and $\|\chi\| = 1$. Define $H = H_0 + V$ as a self-adjoint operator. Find the negative spectrum.

4.4 The harmonic oscillator

The harmonic oscillator potential is

$$V(x) = k|x|^2/2 \tag{4.30}$$

The classical Hamiltonian $|p|^2/2m + k|x|^2/2$ describes a particle with equilibrium position $x = 0$ subject to a linear restoring force $-kx$. The quantum Hamiltonian is

$$H = \frac{-\Delta}{2m} + \frac{k}{2}|x|^2 \tag{4.31}$$

This potential is not a small perturbation of H_0 and is not covered by any of our analysis so far. However we can analyze it directly. For simplicity take $m = 1, k = 1$ and dimension $d = 1$.

Theorem 4.5 *The Hamiltonian*

$$H = \frac{1}{2}\left(-\frac{d^2}{dx^2} + x^2\right) \tag{4.32}$$

is essentially self-adjoint on $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$ and has spectrum $\frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$.

Proof Introduce the operators

$$\begin{aligned} a^* &= \frac{1}{\sqrt{2}}\left(x - \frac{d}{dx}\right) \\ a &= \frac{1}{\sqrt{2}}\left(x + \frac{d}{dx}\right) \end{aligned} \tag{4.33}$$

and then

$$H = a^* a + \frac{1}{2} \tag{4.34}$$

We can find one eigenfunction by solving $a\Omega_0 = 0$, for then $H\Omega_0 = \frac{1}{2}\Omega_0$. The normalized solution is

$$\Omega_0(x) = (\pi)^{-1/4} e^{-x^2/2} \tag{4.35}$$

From this we can generate more eigenfunctions defining

$$\Omega_n = (a^*)^n \Omega_0 \tag{4.36}$$

Since $[a, a^*] = 1$, we have $[H, a^*] = a^*$ and hence can compute

$$H\Omega_n = \left(n + \frac{1}{2}\right)\Omega_n \tag{4.37}$$

Thus the spectrum of H consists of at least the positive half integers.

In fact the Ω_n are the Hermite polynomial and form a complete orthogonal set. A proof of the completeness is sketched in the following problem. Since H has a complete set of eigenfunctions, it is essentially self-adjoint and the spectrum is exactly the eigenvalues $\frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$; see problem 1.13. □

Problem 4.7 Show that the eigenfunctions Ω_n form a complete set by the following steps:

1. Show that the subspace spanned by finite linear combinations is the space $\mathcal{P}$ of all functions of the form $P(x)\exp(-x^2/2)$ where $P(x)$ is a polynomial with complex coefficients.
2. Show that for any $k \in \mathbb{R}$ the function $e^{ikx}e^{-x^2/2}$ is in the closure $\overline{\mathcal{P}}$.
3. Show that $\overline{\mathcal{P}} = L^2(\mathbb{R})$ by showing that the orthogonal complement is $\{0\}$.

Theorem 4.5 readily generalizes to any number of dimensions. The features of this example will surface again in quantum field theory. A scalar quantum field can be thought of as an infinite collection of coupled harmonic oscillators, one for each point in space. The displacement is not in physical space, but in field strength.

Problem 4.8 Find the spectrum for the harmonic oscillator in $d = 3$.

Finally we note that there is an explicit formula for the kernel of the semi-group e^{-tH} known as Mehler's formula. It is

$$(e^{-tH}f)(x) = \int q_t(x,y)f(y)dy \tag{4.38}$$

where

$$q_t(x,y) = (2\pi \sinh t)^{-1/2} \exp\left(-\frac{1}{2}(\coth t)(x^2+y^2) + (\sinh t)^{-1}xy\right) \tag{4.39}$$

Problem 4.9 Check Mehler's formula by verifying the following:

1. For $t > 0$

$$\left(\frac{\partial}{\partial t} + H\right) q_t(x,y) = 0 \tag{4.40}$$

2. For $f \in \mathcal{S}(\mathbb{R})$

$$\lim_{t\to 0} \int q_t(x,y)f(y)dy = f(x) \tag{4.41}$$

4.5 Scattering

4.5.1 Wave operators

We continue to consider the single particle in a potential with Hamiltonian $H = H_0 + V$ and ask for the behavior of the state $e^{-iHt}\psi$ as $t \to \infty$. One possibility is that ψ is an eigenvector $H\psi = E\psi$ and in this case $e^{-iHt}\psi = e^{-iEt}\psi$. The state is localized and stationary and is called a *bound state*.

Another possibility is that the particle escapes the potential and behaves like a free particle. In this case there would be a free state $e^{-iH_0t}\phi$ such that

$$\lim_{t\to\infty} \|e^{-iHt}\psi - e^{-iH_0t}\phi\| = 0 \tag{4.42}$$

This is equivalent to

$$\lim_{t\to\infty} \|\psi - e^{iHt}e^{-iH_0t}\phi\| = 0 \tag{4.43}$$

This motivates the definition of *wave operators* $\Omega^{\pm}$ as the limits

$$\Omega^{\pm}\phi = \lim_{t\to\pm\infty} e^{iHt}e^{-iH_0t}\phi \tag{4.44}$$

when they exist for ϕ in $\mathcal{H}$. Given ϕ let $\psi = \Omega^{\pm}\phi$. Then $\|e^{-iHt}\psi - e^{-iH_0t}\phi\| \to 0$ as $t \to \pm\infty$. Thus we have found a state with given asymptotic behavior either in the past or the future. Note also that

$$\|\Omega^{\pm}\phi\| = \lim_{t\to\pm\infty} \|e^{iHt}e^{-iH_0t}\phi\| = \|\phi\| \tag{4.45}$$

Thus $\Omega^{\pm}$ is an isometry if it exists.

Theorem 4.6 *In $\mathcal{H} = L^2(\mathbb{R}^3)$, if $V \in L^2(\mathbb{R}^3)$, then $\Omega^{\pm}$ exists.*

Proof Let $\Omega_t = e^{iHt}e^{-iH_0t}$. We must show that $\Omega_t\phi$ has a limit as $t \to \pm\infty$. Since $\|\Omega_t\| = 1$, we can approximate $\Omega_t\phi$ by $\Omega_t\phi'$ with $\phi' \in \mathcal{S}(\mathbb{R}^3)$ uniformly in t. Thus it suffices to prove the result for $\phi \in \mathcal{S}(\mathbb{R}^3)$.

Such ϕ are in $D(H) = D(H_0)$ and so we may compute the derivative

$$\begin{aligned} \frac{d}{dt}\Omega_t\phi &= \frac{d}{dt}(e^{iHt}e^{-iH_0t}\phi) \\ &= e^{iHt}(iH - iH_0)e^{-iH_0t}\phi \\ &= e^{iHt}iVe^{-iH_0t}\phi \end{aligned} \tag{4.46}$$

Here we use the fact that e^{-iH_0t} leaves $D(H_0)$ invariant.

Now we write for $t' > t > 0$

$$(\Omega_{t'} - \Omega_t)\phi = \int_t^{t'} e^{iHs}iVe^{-iH_0s}\phi \, ds \tag{4.47}$$

Here the integral can be understood as a Hilbert space valued Riemann integral since the integrand is a continuous Hilbert space valued function. It follows that

$$\begin{aligned} \|(\Omega_{t'} - \Omega_t)\phi\| &\leq \int_t^{t'} \|Ve^{-iH_0s}\phi\| ds \\ &\leq \int_t^{t'} \|V\|_2 \|e^{-iH_0s}\phi\|_\infty ds \\ &\leq 2c\|V\|_2(t^{-1/2} - t'^{-1/2}) \end{aligned} \tag{4.48}$$

Here we have used the bound (4.7) which says

$$\|e^{-iH_0s}\phi\|_\infty \leq cs^{-3/2} \tag{4.49}$$

Thus $\Omega_{t_j}\phi$ is a Cauchy sequence for any $t_j \to \infty$ and hence the limits $\lim_{t\to\infty}\Omega_t$ exists. The limit $t \to -\infty$ is similar. □

Problem 4.10 Let V be a rank-one operator as in problem 4.6, now with $\chi \in L^1 \cap L^2$. Show that $\Omega_\pm$ exist.

4.5.2 Asymptotic completeness

Suppose the wave operators exist. We define

$$\mathcal{H}^{\pm} = \mathrm{Ran}(\Omega^{\pm}) \tag{4.50}$$

Since the range of an isometry is closed, this is a closed subspace of $\mathcal{H}$. These are states which become free as $t \to \pm\infty$. We also define a subspace of bound states

$$\mathcal{H}_{bd} = \text{closed subspace spanned by eigenvectors of } H \tag{4.51}$$

Lemma 4.2 *$\mathcal{H}^{\pm}$ and $\mathcal{H}_{bd}$ are orthogonal subspaces.*

Proof Let $\psi = \Omega^{+}\phi$ and let $H\chi = E\chi$. We show that $(\chi, \psi) = 0$. We have

$$(\chi, \psi) = \lim_{t\to\infty} (\chi, e^{iHt}e^{-iH_0t}\phi) = \lim_{t\to\infty} e^{iEt}(\chi, e^{-iH_0t}\phi) \tag{4.52}$$

Thus it suffices to show for $\chi, \phi \in L^2(\mathbb{R}^3)$ that $\lim_{t\to\infty}(\chi, e^{-iH_0t}\phi) = 0$. We can assume that $\chi, \phi \in \mathcal{S}(\mathbb{R}^3)$ since we can approximate the general case uniformly in t. Then the result follows from the bound (4.49) since

$$|(\chi, e^{-iH_0t}\phi)| \leq \|\chi\|_1 \|e^{-iH_0t}\phi\|_\infty \to 0 \tag{4.53}$$

□

It is possible that these subspaces exhaust the Hilbert space, that is

$$\mathcal{H} = \mathcal{H}^{\pm} \oplus \mathcal{H}_{bd} \tag{4.54}$$

If this is true, we say that the system exhibits *asymptotic completeness*. Note that this entails that $\mathcal{H}^{+} = \mathcal{H}^{-}$. Roughly it means that either a state is stationary or it goes out to infinity both in the distant past and in the distant future.

Asymptotic completeness is generally true. For example if $V \in L^1 \cap L^2$, it is true, although not especially easy to prove.

4.5.3 The scattering operator

Next we discuss actual scattering experiments, for example scattering a neutron off a nucleus, which we model as scattering the neutron off the potential created by the nucleus. We prepare the neutron in a certain state, which we model by the condition that it behaves like $e^{-iH_0t}\phi$ as $t \to -\infty$. Thus the prepared state (at time zero) is $\Omega^{-}\phi$. After the scattering has taken place we measure the state of the system to be something with asymptotic behavior $e^{-iH_0t}\chi$ as $t \to \infty$. Thus the measured state (at time zero) is $\Omega^{+}\chi$. By (3.2) the probability of this occurring is

$$|(\Omega^{+}\chi, \Omega^{-}\phi)|^2 = |(\chi, S\phi)|^2 \tag{4.55}$$

Here we have introduced the *scattering operator*

$$S \equiv (\Omega^{+})^{*}\Omega^{-} \tag{4.56}$$

One can then study the structure of the nucleus by hypothesizing a potential V, computing the scattering operator S for the potential, and then comparing the scattering probabilities $|(\chi, S\phi)|^2$ with the observed events. A more ambitious program is the so-called inverse scattering problem which asks to find V given S.

Lemma 4.3 *If $\mathcal{H}^{+} = \mathcal{H}^{-}$, then S is unitary.*

Proof Since $\Omega^{\pm}$ are isometries, we have

$$(\Omega^{\pm})^*\Omega^{\pm} = I \qquad \Omega^{\pm}(\Omega^{\pm})^* = P_{\mathcal{H}^{\pm}} \tag{4.57}$$

Then we compute

$$S^*S = (\Omega^-)^*\Omega^+(\Omega^+)^*\Omega^- = (\Omega^-)^*\Omega^- = I \tag{4.58}$$

Similarly $SS^* = I$. Hence S is unitary. □

4.5.4 Continuous spectrum

Finally we complete our characterization of the spectrum using the wave operators.

Theorem 4.7 *Suppose that $\Omega^{\pm}$ exist.*

1. $e^{iHt}\Omega^{\pm} = \Omega^{\pm}e^{iH_0t}$.
2. *H restricted to $\mathcal{H}^{\pm}$ has continuous spectrum* $[0, \infty)$.

Proof

1. This follows from the computation

$$\begin{aligned} e^{iHt}\Omega^{\pm}\phi &= \lim_{s\to\pm\infty} e^{iH(t+s)}e^{-iH_0s}\phi \\ &= \lim_{u\to\pm\infty} e^{iHu}e^{-iH_0(u-t)}\phi \\ &= \Omega^{\pm}e^{iH_0t}\phi \end{aligned} \tag{4.59}$$

2. $\Omega^{\pm} : \mathcal{H} \to \mathcal{H}^{\pm}$ is unitary. Then e^{iHt} restricted to $\mathcal{H}^{\pm}$ is unitarily equivalent to e^{iH_0t} on $\mathcal{H}$ by

$$e^{iHt} = \Omega^{\pm}e^{iH_0t}(\Omega^{\pm})^{-1} \tag{4.60}$$

It follows that the generators $H|\mathcal{H}^{\pm}$ and H_0 are unitarily equivalent. Since H_0 has continuous spectrum $[0, \infty)$ (problem 4.2), the same is true for $H|\mathcal{H}^{\pm}$.

□

4.6 Spin

4.6.1 Representations of the rotation group

As noted earlier the natural representation of the rotation group $SO(3)$ on our Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3)$ is $(U(R)\psi)(x) = \psi(R^{-1}x)$. There are however other possibilities for a single particle. Suppose that $R \to T(R)$ is a representation of $SO(3)$ by unitary $m \times m$ matrices so that $T(R_1)T(R_2) = T(R_1R_2)$. Then we could take the Hilbert space

to be $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^m)$, the square integrable $\mathbb{C}^m$ valued functions on $\mathbb{R}^3$, and define a unitary representation on $\mathcal{H}$ by

$$(U(R)\psi)(x) = T(R)\psi(R^{-1}x) \qquad R \in SO(3) \tag{4.61}$$

If we want our particle to be elementary (that is not composite), we would add the requirement that T is irreducible.

It turns out that the most common elementary particles (electrons, protons, neutrons) do behave nontrivially under rotations, but not exactly in the manner suggested above. Instead there is a representation of the universal covering group of $SO(3)$. This is $SU(2)$, the 2×2 complex matrices A satisfying $A^*A = 1$ and $\det A = 1$. As we explain below there is a two-to-one homomorphism $A \to R(A)$ from $SU(2)$ onto $SO(3)$ such that $R(-A) = R(A)$. If $T(A)$ is an m-dimensional unitary representation of $SU(2)$, then there is a unitary representation of $SU(2)$ on the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^m)$ defined by

$$(U(A)\psi)(x) = T(A)\psi(R(A)^{-1}x) \qquad A \in SU(2) \tag{4.62}$$

This also gives a representation of $SO(3)$ if we recall that the states are really rays. To find the action of $R \in SO(3)$ choose $\pm A$ so $R(\pm A) = R$ and define $(U(R)\psi)(x) = T(\pm A)\psi(R^{-1}x)$. For an irreducible representation, $T(-A) = \pm T(A)$. Thus the choice of $\pm A$ at worst changes the sign, and this has no effect on the ray.

It turns out there is an m-dimensional irreducible representation of $SU(2)$ for all positive integers m. The dimension is written $m = 2s + 1$ with $s = 0, 1/2, 1, 3/2, 2, \ldots$ and we say that the particle has *spin* s; more on this terminology later. For spin zero we have $T(A) = I$. This is the case we have been discussing and describes pions. For spin 1/2 we have $T(A) = A$; this is the case that describes electrons, protons, neutrons, etc. For spin one we have $T(A) = R(A)$ and we are back to a special case of (4.61). A modification of this describes photons; see section 9.4. Higher spins are also possible.

4.6.2 The covering group

Now we explain the covering map. First define *Pauli matrices* by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.63}$$

These form a basis for the real vector space of self-adjoint traceless 2×2 matrices. (The matrices $-i\sigma_1/2, -i\sigma_2/2, -i\sigma_3/2$ form a standard basis for the skew-adjoint traceless matrices, the Lie algebra of $SU(2)$.) With any $x \in \mathbb{R}^3$ associate the matrix

$$\sigma \cdot x = \sum_{i=1}^{3} \sigma_i x_i = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} \tag{4.64}$$

Note that

$$\det(\sigma \cdot x) = -|x|^2 \tag{4.65}$$

Now for $A \in SU(2)$ we have that $A(\sigma \cdot x)A^{-1}$ is again a self-adjoint traceless matrix and thus it has the form $\sigma \cdot y$ for a unique $y \in \mathbb{R}^3$. The map $x \to y$ is easily seen to be linear and so $y = R(A)x$ for some matrix $R(A)$. Thus $R(A)$ is defined by[2]

$$A(\sigma \cdot x)A^{-1} = \sigma \cdot (R(A)x) \tag{4.66}$$

Lemma 4.4 *The map $A \to R(A)$ is a two-to-one homomorphism from $SU(2)$ onto $SO(3)$.*

Proof First $R(A)$ is a homomorphism because

$$\begin{aligned} \sigma \cdot (R(AB)x) &= (AB)(\sigma \cdot x)(AB)^{-1} \\ &= A(\sigma \cdot R(B)x)A^{-1} \\ &= \sigma \cdot (R(A)R(B)x) \end{aligned} \tag{4.67}$$

which implies $R(AB) = R(A)R(B)$. Second $R(A)$ is orthogonal because

$$|R(A)x|^2 = -\det(\sigma \cdot (R(A)x)) = -\det(\sigma \cdot x) = |x|^2 \tag{4.68}$$

To see that $R(A)$ is a rotation we need $\det(R(A)) = 1$. This follows from the facts that $\det(R(I)) = \det I = 1$ and that $A \to \det(R(A))$ is continuous from $SU(2)$ to $\{\pm 1\}$ and that $SU(2)$ is connected (see problem 4.11).

Next note that

$$\exp\left(\frac{-i\sigma_3\theta}{2}\right) = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \tag{4.69}$$

is an element of $SU(2)$. By explicit computation we find that

$$\exp\left(\frac{-i\sigma_3\theta}{2}\right)(\sigma \cdot x)\exp\left(\frac{i\sigma_3\theta}{2}\right) = \sigma \cdot (R(e_3, \theta)x) \tag{4.70}$$

where $R(e_3, \theta)$ is the rotation by θ around the e_3 axis (2.38), and so

$$R\left(\exp\left(\frac{-i\sigma_3\theta}{2}\right)\right) = R(e_3, \theta) \tag{4.71}$$

Rotations about the other axes are generated similarly. Since rotations about the three axes generate $SO(3)$, the homomorphism is onto.

To show that the homomorphism is two-to-one it suffices to check that $R(A) = I$ implies $A = \pm I$ which we omit. □

Problem 4.11 Show that every element of $SU(2)$ can be written in the form

$$A = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{4.72}$$

[2] Essentially we are defining $R(A)$ as the adjoint representation of $SU(2)$.

for complex α, β satisfying $|\alpha|^2 + |\beta|^2 = 1$. Thus $SU(2)$ can be identified with the three-sphere S^3 and hence is simply connected.

4.6.3 Spin 1/2 particles

Now we discuss quantum mechanics for a spin 1/2 particle. The Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^2)$ and the representation of $SU(2)$ is

$$(U(A)\psi)(x) = A\psi(R(A)^{-1}x) \qquad A \in SU(2) \tag{4.73}$$

Our definition of angular momentum should now be modified. The third component is the generator of rotations around the third axis and is now given by

$$(J_3\psi)(x) = i\frac{d}{d\theta}\left[\exp\left(\frac{-i\sigma_3\theta}{2}\right)\psi(R(e_3,\theta)^{-1}x)\right]_{\theta=0} \tag{4.74}$$

This is computed as

$$J_3 = L_3 + \frac{\sigma_3}{2} \tag{4.75}$$

Here L_3 is the "orbital" angular momentum computed in (3.27). To this is added an intrinsic angular momentum $\sigma_3/2$ called spin. The spin operator has eigenvalues $\pm 1/2$ hence the term "spin 1/2." Other components are treated similarly. Spin has no classical analog.

A spin 1/2 particle has a modified Hamiltonian in the presence of electric and magnetic fields with potentials (Φ, A). This is the Pauli Hamiltonian

$$H = \frac{1}{2m}\left(-i\nabla - \frac{e}{c}A\right)^2 + e\Phi - \frac{e}{mc}\left(\frac{\sigma}{2}\cdot B\right) \tag{4.76}$$

where $B = \nabla \times A$ is the magnetic field. This arises naturally as an approximation to a relativistic Dirac Hamiltonian.

Problem 4.12 Suppose that the magnetic field B is constant with $A = (B \times x)/2$. Show that for e/c small

$$\frac{1}{2m}\left(-i\nabla - \frac{e}{c}A\right)^2 = \frac{-\Delta}{2m} - \frac{e}{2mc}(L\cdot B) + \mathcal{O}(e^2/c^2) \tag{4.77}$$

This problem shows that in the Pauli Hamiltonian the spin angular momentum $\sigma/2$ couples to a magnetic field in the same way as the orbital angular momentum L, except for a factor of 2 known as the "gyromagnetic ratio."

Notes on chapter 4: See Reed and Simon (1980), Reed and Simon (1975), Reed and Simon (1979), Reed and Simon (1978), and Cycon *et al.* (1987) for much

more about self-adjointness, spectra, and scattering. For spin and representations of $SU(2)$, see Miller (1972). Besides potentials that are functions one can also study delta function potentials, see Albeverio *et al.* (1988).

The expression $-i\nabla - ec^{-1}A$ in (4.23) or (4.76) can be interpreted as a covariant derivative on a complex line bundle. See section 7.4

5 Many particles

5.1 Two particles

5.1.1 A first look

Suppose we have two (spinless) particles. As explained in example 2.3 the classical Hamiltonian for the system might have the form

$$H(p_1, p_2, x_1, x_2) = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(x_1 - x_2) \tag{5.1}$$

Here $p_i \in \mathbb{R}^3, x_i \in \mathbb{R}^3$ are the momentum and position of the ith particle, m_i is the mass of the ith particle and V is a potential giving the interaction between them. Following our canonical quantization procedure we replace p_1, p_2, x_1, x_2 by operators $\hat{p}_1, \hat{p}_2, \hat{x}_1, \hat{x}_2$ satisfying the canonical commutation relations. We take $\hat{p}_i = -i\nabla_{x_i}$ and $\hat{x}_i = [x_i]$ acting in the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^3 \times \mathbb{R}^3) = L^2(\mathbb{R}^6)$. With this choice the quantum Hamiltonian becomes

$$H = \frac{-\Delta_1}{2m_1} + \frac{-\Delta_2}{2m_2} + V(x_1 - x_2) \tag{5.2}$$

where Δ_i is the Laplacian in x_i. If $V = 0$, then the Hamiltonian is

$$H_0 = \frac{-\Delta_1}{2m_1} + \frac{-\Delta_2}{2m_2} \tag{5.3}$$

This can be defined as a self-adjoint operator with the Fourier transform on $\mathbb{R}^6$ by

$$H_0 = \mathcal{F}^{-1}\left[\frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2}\right]\mathcal{F} \tag{5.4}$$

where the multiplication operator has the natural domain.

Problem 5.1 If $V \in L^2(\mathbb{R}^3)$, show that $H = H_0 + V$ is self-adjoint on $D(H_0)$.

5.1.2 Generalization

Now we give a more general treatment. Suppose we have two particles: the first with Hilbert space $\mathcal{H}_1$ and time evolution $U_1(t) = \exp(-iH_1 t)$ and the second with Hilbert space $\mathcal{H}_2$ and time evolution $U_2(t) = \exp(-iH_2 t)$. The Hilbert spaces may allow spin or other internal degrees of freedom like charge. To describe the two particle system, the prescription is that the Hilbert space is the tensor product

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{5.5}$$

(See appendix B for the definition of tensor product.) If the particles do not interact, then they should evolve in time just as they would by themselves, that is

$$U(t) = U_1(t) \otimes U_2(t) \tag{5.6}$$

This is a strongly continuous unitary group and so by Stone's theorem it has a self-adjoint generator $U(t) = \exp(-iHt)$. We compute $H\Psi = i\, d/dt[U(t)\Psi]_{t=0}$ when it exists. For example if $\psi \in D(H_1)$ and $\phi \in D(H_2)$

$$H(\psi \otimes \phi) = (H_1\psi \otimes \phi) + (\psi \otimes H_2\phi) \tag{5.7}$$

More generally let $D(H_1) \otimes D(H_2)$ be the algebraic tensor product of $D(H_1)$ and $D(H_2)$, that is finite linear combinations of $\psi \otimes \phi$ as above. On this dense domain we have

$$H = (H_1 \otimes I) + (I \otimes H_2) \tag{5.8}$$

Let us see how this reproduces the earlier treatment. Suppose we have two free spinless particles so that $\mathcal{H}_i = L^3(\mathbb{R}^3)$ and $H_i = -\Delta/2m_i$. In this case we have a natural identification

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3) \Longleftrightarrow L^2(\mathbb{R}^6) \tag{5.9}$$

The identification is given by a unitary operator which sends the vector $\psi \otimes \phi \in L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$ to the function $\psi \otimes \phi \in L^2(\mathbb{R}^6)$ defined by

$$(\psi \otimes \phi)(x_1, x_2) = \psi(x_1)\phi(x_2) \tag{5.10}$$

(See appendix B for details.) Under this identification we have

$$\begin{aligned} H_1 \otimes I &= \left(\frac{-\Delta}{2m_1} \otimes I\right) \Longleftrightarrow \frac{-\Delta_1}{2m_1} \\ I \otimes H_2 &= \left(I \otimes \frac{-\Delta}{2m_2}\right) \Longleftrightarrow \frac{-\Delta_2}{2m_2} \end{aligned} \tag{5.11}$$

Thus the Hamiltonian $H = H_1 \otimes I + I \otimes H_2$ is identified with the Hamiltonian $H_0 = -\Delta_1/2m_1 - \Delta_2/2m_2$ on $L^2(\mathbb{R}^6)$ as defined earlier.

5.1.3 Center of mass coordinates

How can we incorporate potentials into this tensor product structure? We give one answer now and another answer later in section 5.4.2. Working on $L^2(\mathbb{R}^6)$ we make a change of coordinates

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} \qquad x = x_1 - x_2 \tag{5.12}$$

Then X represents the center of mass of the system and x is the relative displacement of the particles. If $M = m_1 + m_2$ is the total mass, then the inverse is

$$x_1 = X + \frac{m_2}{M} x \qquad x_2 = X - \frac{m_1}{M} x \tag{5.13}$$

The coordinate change is implemented by the operator

$$(V\psi)(X, x) = \psi\left(X + \frac{m_2}{M} x,\ X - \frac{m_1}{M} x\right) \tag{5.14}$$

The operator V is unitary on $L^2(\mathbb{R}^6)$ since the Jacobian determinant for the transformation has absolute value one.

If H is the full two particle Hamiltonian (5.2), we find in the new coordinates

$$H' \equiv VHV^{-1} = H_{cm} + H_{rel} \tag{5.15}$$

where on $\Psi(X, x) \in L^2(\mathbb{R}^6)$

$$\begin{aligned} H_{cm} &= \frac{-\Delta_X}{2M} \\ H_{rel} &= \frac{-\Delta_x}{2\mu} + V(x) \end{aligned} \tag{5.16}$$

and where

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{5.17}$$

is called the reduced mass.

Now under the correspondence $L^2(\mathbb{R}^6) \leftrightarrow L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$ with the identification $\phi(X)\psi(x) \leftrightarrow \phi \otimes \psi$ we find as before

$$\begin{aligned} e^{-iH't} &= \exp(-iH_{cm}t) \otimes \exp(-iH_{rel}t) \\ H' &= (H_{cm} \otimes I) + (I \otimes H_{rel}) \end{aligned} \tag{5.18}$$

Thus the center of mass and a fictitious relative particle evolve in time independently of each other. The motion of the center of mass is free. The motion of the relative particle is the same as that of a single particle in a potential which we have studied

at length in the previous chapter. The only change is that we have the relative mass μ instead of the actual mass. If m_2 is much larger than m_1, then $\mu \approx m_1$. This is the case for the hydrogen atom where the mass of the proton is much larger than the mass of the electron.

Problem 5.2 Verify (5.15).

5.2 Identical particles

Until now we have been implicitly assuming that our two particles are distinguishable in the sense that they have a different mass or spin or charge. But suppose that this is not the case, for example suppose we have two electrons. Empirically there is no way to tell which is which. This statement is true in the strong sense that there is no way to label the particles and follow their individual evolution through the course of an experiment. This seems to be a deep fact and not just a reflection of our limited skill at experiments. Then describing the two particle system by a Hilbert space $\mathcal{H} \otimes \mathcal{H}$, which effectively labels the particles, would be a substantial over-description. Indeed nature does not choose this Hilbert space, but rather a subspace which is invariant under permutation of the labels.

We define a permutation operator P on $\mathcal{H} \otimes \mathcal{H}$ by

$$P(\psi_1 \otimes \psi_2) = \psi_2 \otimes \psi_1 \tag{5.19}$$

This satisfies $P^2 = I$ and $P^* = P$. We restrict to the subspace which is invariant under P. The orthogonal projection onto this subspace is

$$\Pi^+ = \frac{1}{2}(I + P) \tag{5.20}$$

since this is a projection and $P\psi = \psi$ iff $\psi \in \text{Ran } \Pi^+$. Thus our Hilbert space is

$$\mathcal{H}_2^+ = \Pi^+(\mathcal{H} \otimes \mathcal{H}) \tag{5.21}$$

and is called the *symmetric tensor product*. An example of an element of this space is

$$\Pi^+(f \otimes g) = \frac{1}{2}(f \otimes g + g \otimes f) \tag{5.22}$$

Particles for which this is the correct Hilbert space are called *bosons*. Examples are pions and photons.

Remarkably this is not the only interesting possibility. Another possibility is that the state changes sign under the permutation operator P. A change of sign still gives

the same ray so our description is still invariant under P. The projection onto the subspace which changes signs under P is

$$\Pi^- = \frac{1}{2}(I - P) \tag{5.23}$$

Now the Hilbert space is

$$\mathcal{H}_2^- = \Pi^-(\mathcal{H} \otimes \mathcal{H}) \tag{5.24}$$

and is called the *anti-symmetric tensor product*. An example of an element of this space is

$$\Pi^-(f \otimes g) = \frac{1}{2}(f \otimes g - g \otimes f) \tag{5.25}$$

Particles for which this is the correct Hilbert space are called *fermions*. Examples are electrons, protons, and neutrons.

The choice of the symmetric or anti-symmetric tensor product is called the *statistics* of the particle. It turns out that particles with integer spin are always bosons, and particles with half-integer spin are always fermions. This fact has an explanation in quantum field theory, but for us it is just an empirical fact. In spite of this spin-statistics connection we will sometimes find it useful to consider spinless fermions.

5.3 *n*-particles

Starting with a Hilbert space for a single particle we want to construct a Hilbert space for n identical particles. Consider the n-fold tensor product

$$\mathcal{H}_n = \mathcal{H} \otimes \cdots \otimes \mathcal{H} \tag{5.26}$$

We define operators on $\mathcal{H}_n$ by

$$\begin{aligned} \Pi^+(f_1 \otimes \cdots \otimes f_n) &= \frac{1}{n!}\sum_\pi f_{\pi(1)} \otimes \cdots \otimes f_{\pi(n)} \\ \Pi^-(f_1 \otimes \cdots \otimes f_n) &= \frac{1}{n!}\sum_\pi sgn(\pi) f_{\pi(1)} \otimes \cdots \otimes f_{\pi(n)} \end{aligned} \tag{5.27}$$

where the sum is over permutations π of $(1, \ldots, n)$ and $sgn(\pi)$ is the sign of the permutation.[1] One can check that this defines an operator on the dense domain of finite combinations of vectors $f_1 \otimes \cdots \otimes f_n$. One also checks that $(\Pi^\pm)^2 = \Pi^\pm$ and $(\Pi^\pm)^* = \Pi^\pm$. It follows that $\|\Pi^\pm \psi\| \leq \|\psi\|$ and hence $\Pi^\pm$ extends to a

[1] A permutation π is a bijection on $(1, \ldots, n)$. The sign $sgn(\pi)$ depends on the number of elementary exchanges to return $\pi(1), \ldots, \pi(n)$ to its original order. It is ± 1 depending on whether the number is even or odd. We have $sgn(\pi \circ \sigma) = sgn(\pi) sgn(\sigma)$.

bounded operator on $\mathcal{H}_n$ (theorem 1.3). The extensions are projection operators since the identities hold for the extensions.

The Hilbert space for n identical particles is taken as

$$\mathcal{H}_n^{\pm} = \Pi^{\pm}\mathcal{H}_n = \Pi^{\pm}(\mathcal{H} \otimes \cdots \otimes \mathcal{H}) \tag{5.28}$$

with the plus sign for bosons and the minus sign for fermions. Exchanging two entries in this space has no effect for bosons, and changes the sign for fermions

$$\begin{aligned} &\Pi^{\pm}(f_1 \otimes \cdots \otimes f_i \otimes \cdots \otimes f_j \otimes \cdots \otimes f_n) \\ &= \pm\Pi^{\pm}(f_1 \otimes \cdots \otimes f_j \otimes \cdots \otimes f_i \otimes \cdots \otimes f_n) \end{aligned} \tag{5.29}$$

For fermions this means that if $f_i = f_j$ for some $i \neq j$, then

$$\Pi^{-}(f_1 \otimes \cdots \otimes f_n) = 0 \tag{5.30}$$

This is the *Pauli exclusion principle*: two identical fermions cannot be in the same state.

Next we consider a simple dynamics on $\mathcal{H}_n^{\pm}$. In general if U is a unitary operator on $\mathcal{H}$, then

$$\Gamma_n(U) \equiv U \otimes \cdots \otimes U \tag{5.31}$$

defines a unitary operator on $\mathcal{H}_n$ which preserves the subspaces $\mathcal{H}_n^{\pm}$. In particular if e^{-iHt} is a time evolution on the single particle space $\mathcal{H}$, and if the particles do not interact with each other, then

$$\Gamma_n(e^{-iHt}) \equiv e^{-iHt} \otimes \cdots \otimes e^{-iHt} \tag{5.32}$$

is the time evolution on $\mathcal{H}_n^{\pm}$. This is a one-parameter unitary group and so has a self-adjoint generator H_n such that

$$e^{-iH_n t} = \Gamma_n(e^{-iHt}) \tag{5.33}$$

On $D(H) \otimes \cdots \otimes D(H)$ we compute by taking derivatives

$$H_n = d\Gamma_n(H) \equiv H \otimes \cdots \otimes I + \cdots + I \otimes \cdots \otimes H \tag{5.34}$$

Example 5.1 **Atoms** An atom with atomic number N is described by N-electrons each moving under the influence of a central potential created by an atomic nucleus of charge N

$$V(x) = \frac{-e^2 N}{|x|} \tag{5.35}$$

The single particle Hamiltonian is $H = (-\Delta/2m) + V$ on $\mathcal{H} = L^2(\mathbb{R}^3)$ as in example 4.3. (Or we could take the Pauli Hamiltonian (4.76) on $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^2)$.)

Since electrons are fermions, the multiparticle Hamiltonian is then $H_N = d\Gamma_N(H)$ on $\mathcal{H}_N^-$ as above.

Now H has the spectrum of the hydrogen atom but with e^2 replaced by e^2N. Thus it has eigenfunctions $\{\phi_k\}$ with eigenvalues $\{e_k\}$ labeled so that $e_1 \leq e_2 \leq e_3 \leq \dots$. The multiparticle Hamiltonian H_N has eigenfunctions

$$\phi_{\alpha_1,\dots,\alpha_N} = \Pi^-(\phi_{\alpha_1} \otimes \cdots \otimes \phi_{\alpha_n}) \tag{5.36}$$

with eigenvalues

$$e_{\alpha_1,\dots,\alpha_N} = \sum_{i=1}^{N} e_{\alpha_i} \tag{5.37}$$

Here $\alpha_1, \dots, \alpha_N$ is a sequence of positive integers. However because of the Pauli exclusion principle they must be distinct integers. Thus the lowest energy states, physically the stable states, will be states like $\phi_{1,2,\dots,N}$. The electrons fill the lowest energy levels ("shells").

This is the starting point for chemistry. Our model of an atom still needs refinement since we have neglected a number of additional effects, starting with the interaction between the electrons themselves.

Example 5.2 **Interacting particles** Consider the case of N spinless bosons or fermions interacting only with each other. The single particle Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^3)$ with Hamiltonian $H_0 = -\Delta/2m$. The N-particle Hilbert space is $\mathcal{H}_N^{\pm}$ and the free Hamiltonian is $H_{0,N} = d\Gamma_N(H_0)$. However $\mathcal{H}_N^{\pm}$ is isomorphic to $L^2_{\pm}(\mathbb{R}^{3N})$ the symmetric or anti-symmetric subspace of $L^2(\mathbb{R}^{3N})$. Under this isomorphism we have

$$H_{0,N} = \sum_{i=1}^{N} \frac{-\Delta_i}{2m} \tag{5.38}$$

If v is the potential between two particles with $v(x) = v(-x)$, then the total potential is V_N defined by

$$V_N(x_1, \dots, x_N) = \sum_{1 \leq i < j \leq N} v(x_i - x_j) \tag{5.39}$$

This acts on $L^2_{\pm}(\mathbb{R}^{3N})$ and the total Hamiltonian on this space is

$$H_N = H_{0,N} + V_N \tag{5.40}$$

As in problem 5.1 we can show that this is self-adjoint on $D(H_{0,N})$.

Problem 5.3 In the previous example define representations of the translation and rotation groups. Find the total momentum and angular momentum. Are they conserved? (This is an extension of example 3.2.)

Problem 5.4 From the definition (5.27) check that $\Pi^{\pm}$ is well-defined, that $(\Pi^{\pm})^2 = \Pi^{\pm}$ and $(\Pi^{\pm})^* = \Pi^{\pm}$, and deduce that $\|\Pi^{\pm}\psi\| \leq \|\psi\|$.

5.4 Fock space

5.4.1 Definitions

In non-relativistic quantum mechanics, for a closed system, the number of particles is fixed. Nevertheless it is convenient to introduce a formalism in which there are an indefinite number of particles. There are several reasons for this:

1. The formalism for an indefinite number of particles has some elegant features which can be used even when the particle number is fixed.
2. The number of particles may undergo statistical fluctuations due to contact with an external system (more about this later).
3. It makes contact with relativistic quantum systems where the number of particles in a closed system can actually change.

Start with a single particle Hilbert space $\mathcal{H}$, and define the n-particle Hilbert space $\mathcal{H}_n^{\pm}$ to be the symmetric or anti-symmetric n-fold tensor product as in the last section. Then the (boson/fermion) *Fock space* over $\mathcal{H}$ is the infinite direct sum of the $\mathcal{H}_n^{\pm}$. It is

$$\mathcal{F}^{\pm}(\mathcal{H}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n^{\pm} \tag{5.41}$$

where $\mathcal{H}_0^{\pm} \equiv \mathbb{C}$ corresponds to no particles. The elements are sequences $\psi = (\psi_0, \psi_1, \psi_2, \dots)$ with $\psi_n \in \mathcal{H}_n^{\pm}$ such that

$$\|\psi\|^2 = \sum_{n=0}^{\infty} \|\psi_n\|^2 < \infty \tag{5.42}$$

This is a Hilbert space with inner product

$$(\psi, \chi) = \sum_{n=0}^{\infty} (\psi_n, \chi_n) \tag{5.43}$$

The no-particle state

$$\Omega_0 = (1, 0, 0, \dots) \tag{5.44}$$

is also called the vacuum.

We describe some operators on this space:

(1) The number operator N is defined by

$$N(\psi_0, \psi_1, \psi_2, \psi_3, \dots) = (0, \psi_1, 2\psi_2, 3\psi_3, \dots) \tag{5.45}$$

or equivalently

$$(N\psi)_n = n\psi_n \tag{5.46}$$

It is self-adjoint on

$$D(N) = \{\psi : \sum_n n^2 \|\psi_n\|^2 < \infty\} \tag{5.47}$$

It describes the number of particles in the state. For $\|\psi\| = 1$ the quantity $\|\psi_n\|^2$ is the probability of finding n-particles in the state.

(2) In general if U is a unitary operator on $\mathcal{H}$, then $\Gamma(U)$ is the unitary operator on $\mathcal{F}^{\pm}(\mathcal{H})$ defined by

$$\Gamma(U) = \bigoplus_{n=0}^{\infty} \Gamma_n(U) \tag{5.48}$$

where Γ_n is defined in (5.31). In particular if $e^{-iH_1 t}$ is a time evolution on $\mathcal{H}$, then we define a time evolution on $\mathcal{F}^{\pm}(\mathcal{H})$ as

$$\Gamma(e^{-iH_1 t}) = \bigoplus_{n=0}^{\infty} \Gamma_n(e^{-iH_1 t}) \tag{5.49}$$

This is a strongly continuous one-parameter unitary group and so has a self-adjoint generator H such that $e^{-iHt} = \Gamma(e^{-iH_1 t})$. On a domain with a finite number of particles and with wave functions in the algebraic tensor product $D(H) \otimes \cdots \otimes D(H)$ we find that

$$H = d\Gamma(H_1) \equiv \bigoplus_{n=0}^{\infty} H_n \tag{5.50}$$

where $H_n = d\Gamma_n(H_1)$ is defined in (5.34). This turns out to be a domain of essential self-adjointness and so determines the operator completely. Note also that $N = d\Gamma(1)$

(3) We introduce creation and annihilation operators, at first without statistics on $\mathcal{H}_n = \mathcal{H} \otimes \cdots \otimes \mathcal{H}$. For $h \in \mathcal{H}$ define $\alpha^*(h) : \mathcal{H}_n \to \mathcal{H}_{n+1}$ and $\alpha(h) : \mathcal{H}_n \to \mathcal{H}_{n-1}$ by

$$\begin{aligned} \alpha^*(h)(f_1 \otimes \cdots \otimes f_n) &= \sqrt{n+1}\, h \otimes f_1 \otimes \cdots \otimes f_n \\ \alpha(h)(f_1 \otimes \cdots \otimes f_n) &= \sqrt{n}\, (h, f_1)\, f_2 \otimes \cdots \otimes f_n \end{aligned} \tag{5.51}$$

One checks that these formulas do indeed define operators. Furthermore the operators are bounded with

$$\|\alpha^*(h)\| \leq \sqrt{n+1}\|h\| \qquad \|\alpha(h)\| \leq \sqrt{n}\|h\| \tag{5.52}$$

They are adjoint to each other because for $\psi \in \mathcal{H}_n$ and $\chi \in \mathcal{H}_{n+1}$ we have $(\alpha^*(h)\psi, \chi) = (\psi, \alpha(h)\chi)$. Note also that $\alpha^*(h)$ is linear in h while $\alpha(h)$ is anti-linear in h.

Next we define $a^*(h) : \mathcal{H}_n^{\pm} \to \mathcal{H}_{n+1}^{\pm}$ and $a(h) : \mathcal{H}_n^{\pm} \to \mathcal{H}_{n-1}^{\pm}$ by

$$\begin{aligned} a^*(h) &= \Pi^{\pm}\alpha^*(h) \\ a(h) &= \Pi^{\pm}\alpha(h) \end{aligned} \tag{5.53}$$

These are still bounded operators with the same bound and they are still adjoint to each other. We need a more explicit expression for these operators

Lemma 5.1

$$\begin{aligned} &a^*(h)\Pi_n^{\pm}(f_1 \otimes \cdots \otimes f_n) \\ &\qquad = \sqrt{n+1}\ \Pi_{n+1}^{\pm}(h \otimes f_1 \otimes \cdots \otimes f_n) \\ &a(h)\Pi_n^{\pm}(f_1 \otimes \cdots \otimes f_n) \\ &\qquad = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (\pm 1)^{j+1} (h, f_j) \Pi_{n-1}^{\pm}(f_1 \otimes \cdots \otimes \hat{f}_j \otimes \cdots \otimes f_n) \end{aligned} \tag{5.54}$$

where the "hat" on $\hat{f}_j$ means omit this entry.

Proof We give the proof for fermions; bosons are easier. With $h = f_0$ we have

$$\begin{aligned} &\alpha^*(f_0)\Pi_n^{-}(f_1 \otimes \cdots \otimes f_n) \\ &= \frac{\sqrt{n+1}}{n!} \sum_{\pi} sgn(\pi) f_0 \otimes f_{\pi(1)} \otimes \cdots \otimes f_{\pi(n)} \\ &= \frac{\sqrt{n+1}}{n!} \sum_{\pi' : \pi'(0)=0} sgn(\pi') f_{\pi'(0)} \otimes f_{\pi'(1)} \otimes \cdots \otimes f_{\pi'(n)} \end{aligned} \tag{5.55}$$

In the second expression we have replaced the sum over permutations π on $(1, \ldots, n)$ with a sum over permutations π' on $(0, 1, \ldots, n)$ that leave 0 fixed. Now to get $a^*(f_0)\Pi_n^{-}(f_1 \otimes \cdots \otimes f_n)$ we apply Π_{n+1}^{-}. However

$$\Pi_{n+1}^{-}(sgn(\pi') f_{\pi'(0)} \otimes f_{\pi'(1)} \otimes \cdots \otimes f_{\pi'(n)}) = \Pi_{n+1}^{-}(f_0 \otimes f_1 \otimes \cdots \otimes f_n) \tag{5.56}$$

and

$$\frac{1}{n!} \sum_{\pi' : \pi'(0)=0} 1 = 1 \tag{5.57}$$

so we have the first result.

For the second result we compute

$$\begin{aligned}
&\alpha(f_0)\Pi_n^-(f_1\otimes\cdots\otimes f_n)\\
&=\frac{\sqrt{n}}{n!}\sum_\pi sgn(\pi)(f_0,f_{\pi(1)})f_{\pi(2)}\otimes\cdots\otimes f_{\pi(n)}\\
&=\frac{\sqrt{n}}{n!}\sum_j(-1)^{j+1}(f_0,f_j)\sum_\sigma sgn(\sigma)f_{\sigma(2)}\otimes\cdots\otimes f_{\sigma(n)}\\
&=\frac{1}{\sqrt{n}}\sum_j(-1)^{j+1}(f_0,f_j)\Pi_{n-1}^-(f_2\otimes\cdots\hat{f}_j\cdots\otimes f_n)
\end{aligned}\tag{5.58}$$

Here in the second step we have replaced the sum over permutations π of $(1,\dots,n)$ by a sum over $j=\pi(1)$ and a sum over bijections σ from $(2,\dots,n)$ to $(1,\dots,\hat{j},\dots,n)$. We have also used $sgn(\pi)=(-1)^{j+1}sgn(\sigma)$ where $sgn(\sigma)$ is the number of elementary exchanges to return $\sigma(2),\dots,\sigma(n)$ to its natural order. (It takes $j-1$ exchanges to move $\pi(1)=j$ back to the first position, and $(-1)^{j-1}=(-1)^{j+1}$.)

Thus $\alpha(h)$ already maps into the symmetrized subspace, hence $a(h)=\alpha(h)$, and hence the result. □

The operators $a^*(h),a(h)$ on $\mathcal{H}_n^\pm$ induce operators on the boson/fermion Fock spaces $\mathcal{F}^\pm(\mathcal{H})$ by letting them act on each component. We make the convention that $a(h)=0$ on $\mathcal{H}_0$. As a dense domain for these operators we take the finite particle vectors

$$\mathcal{D}_0=\{\psi\in\mathcal{F}^\pm(\mathcal{H}):\exists N \text{ so } \psi_n=0 \text{ for } n\geq N\}\tag{5.59}$$

The operators preserve this domain. From (5.54) we have

$$\Pi_n^\pm(f_1\otimes\cdots\otimes f_n)=\frac{1}{\sqrt{n!}}\,a^*(f_1)\cdots a^*(f_n)\Omega_0\tag{5.60}$$

Thus we can create a dense set of states by acting on the vacuum with creation operators and taking the linear span.

Lemma 5.2 *With $[A,B]_\pm = AB\pm BA$ the following commutation or anti-commutation relations hold on $\mathcal{F}^\pm(\mathcal{H})$*

$$\begin{aligned}
[a(g),a(h)]_\mp&=0\\
[a^*(g),a^*(h)]_\mp&=0\\
[a(g),a^*(h)]_\mp&=(g,h)
\end{aligned}\tag{5.61}$$

Proof To check the last we compute using (5.54)

$$\begin{aligned}
&a(g)a^*(h)\Pi_n^\pm(f_1\otimes\cdots\otimes f_n)\\
&=(g,h)\Pi_n^\pm(f_1\otimes\cdots\otimes f_n)\\
&+\sum_{j=1}^n(\pm1)^j(g,f_j)\Pi_n^\pm(h\otimes f_1\otimes\cdots\otimes\hat{f}_j\otimes\cdots\otimes f_n)
\end{aligned}\tag{5.62}$$

On the other hand

$$\begin{aligned} &a^*(h)a(g)\Pi_n^{\pm}(f_1 \otimes \cdots \otimes f_n) \\ &= \sum_{j=1}^{n} (\pm 1)^{j+1} (g, f_j) \Pi_n^{\pm}(h \otimes f_1 \otimes \cdots \otimes \hat{f}_j \otimes \cdots \otimes f_n) \end{aligned} \tag{5.63}$$

Comparing these gives the result $[a(g), a^*(h)]_{\mp} = (g, h)$. □

We still have that $a^*(h)$ is linear in h, that $a(h)$ is anti-linear in h, and that for $\psi, \chi \in \mathcal{D}_0$

$$(a^*(h)\psi, \chi) = (\psi, a(h)\chi) \tag{5.64}$$

This says that $(a(h))^* \supset a^*(h)$ and $(a^*(h))^* \supset a(h)$. Since the adjoints are densely defined, the operators $a(h), a^*(h)$ have closures which we denote by the same symbol. To characterize the closure we have the following:

Lemma 5.3

1. *For bosons $D(\sqrt{N}) \subset D(a^*(h))$, also for $a(h)$, and for $\psi \in D(\sqrt{N})$*

$$\|a^*(h)\psi\| \leq \|h\| \|\sqrt{N+1}\psi\| \qquad \|a(h)\psi\| \leq \|h\| \|\sqrt{N}\psi\| \tag{5.65}$$

2. *For fermions the closed operators are bounded and satisfy*

$$\|a^*(h)\psi\| \leq \|h\| \|\psi\| \qquad \|a(h)\psi\| \leq \|h\| \|\psi\| \tag{5.66}$$

Proof

1. (Bosons) The inequality holds for $\psi \in \mathcal{D}_0$ by (5.52). But $\mathcal{D}_0$ is a core for $\sqrt{N}$, which means that given $\psi \in D(\sqrt{N})$ there are sequences $\psi_k \in \mathcal{D}_0$ such that $\psi_k \to \psi, \sqrt{N}\psi_k \to \sqrt{N}\psi$ as $k \to \infty$. (For example let ψ_k be the truncation of ψ at the kth entry.) Then we have

$$\|a^*(h)(\psi_j - \psi_k)\| \leq \|h\| \|\sqrt{N+1}(\psi_j - \psi_k)\| \to 0 \tag{5.67}$$

   as $j, k \to \infty$. Thus $\lim_{k\to\infty} a^*(h)\psi_k$ exists and since $a^*(h)$ is closed it follows that $\psi \in D(a^*(h))$ and that $a^*(h)\psi = \lim_{k\to\infty} a^*(h)\psi_k$. Hence $D(\sqrt{N}) \subset D(a^*(h))$. Taking the limit in $\|a^*(h)\psi_k\| \leq \|h\| \|\sqrt{N+1}\psi_k\|$ we get the inequality for ψ.
2. (Fermions) The anti-commutation relation $a^*(h)a(h) + a(h)a^*(h) = \|h\|^2$ implies for $\psi \in \mathcal{D}_0$

$$\|a(h)\psi\|^2 + \|a^*(h)\psi\|^2 = \|h\|^2 \|\psi\|^2 \tag{5.68}$$

   Hence the inequalities hold for $\psi \in \mathcal{D}_0$ and it follows that closures are bounded operators satisfying the same bound.

□

Problem 5.5 Verify the claims made for $\alpha^*(h), \alpha(h)$.

Problem 5.6 If U is unitary on $\mathcal{H}$, show that

$$\begin{aligned}\Gamma(U)a(h)\Gamma(U^{-1}) &= a(Uh)\\ \Gamma(U)a^*(h)\Gamma(U^{-1}) &= a^*(Uh)\end{aligned} \tag{5.69}$$

Problem 5.7 Let T be a contraction on $\mathcal{H}$, that is $\|T\| \leq 1$. Show that $\Gamma(T)$ defines a contraction on $\mathcal{F}^{\pm}(\mathcal{H})$. Show that $\Gamma(T)$ may not be bounded if $\|T\| > 1$.

5.4.2 Fock space over L^2

We now specialize to the case where the one-particle Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^3)$. In this case $\mathcal{H}_n^{\pm}$ is identified with the symmetric or anti-symmetric subspace $L^2_{\pm}(\mathbb{R}^{3n})$ of $L^2(\mathbb{R}^{3n})$.

Lemma 5.4 *If $\psi \in \mathcal{H}_n^{\pm}$ is identified with $\psi \in L^2_{\pm}(\mathbb{R}^{3n})$, then $a(h)\psi \in \mathcal{H}_{n-1}^{\pm}$ is identified with $a(h)\psi \in L^2_{\pm}(\mathbb{R}^{3(n-1)})$ given by*

$$(a(h)\psi)(x_1,\ldots,x_{n-1}) = \sqrt{n}\int \overline{h(x)}\psi(x,x_1,\ldots,x_{n-1})\,dx \tag{5.70}$$

Proof We give the proof for fermions. Consider $\psi = \Pi^-(f_1\otimes\cdots\otimes f_n)$ which is identified with

$$\psi(x_1,\ldots,x_n) = \frac{1}{n!}\sum_{\pi} sgn(\pi) f_{\pi(1)}(x_1)\cdots f_{\pi(n)}(x_n) \tag{5.71}$$

Then $a(h)\psi$ is identified with

$$\begin{aligned}&(a(h)\psi)(x_2,\ldots,x_n)\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(-1)^{j+1}(h,f_j)[\Pi^-_{n-1}(f_1\otimes\cdots\hat{f}_j\cdots\otimes f_n)](x_2,\ldots,x_n)\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(-1)^{j+1}(h,f_j)\frac{1}{(n-1)!}\sum_{\sigma} sgn(\sigma) f_{\sigma(2)}(x_2)\cdots f_{\sigma(n)}(x_n)\\ &= \sqrt{n}\int \overline{h(x_1)}\left[\frac{1}{n!}\sum_{\pi} sgn(\pi) f_{\pi(1)}(x_1)\cdots f_{\pi(n)}(x_n)\right]dx_1\\ &= \sqrt{n}\int \overline{h(x_1)}\psi(x_1,\ldots,x_n)dx_1\end{aligned} \tag{5.72}$$

The fourth line follows as in (5.58). □

If $\psi \in L^2_{\pm}(\mathbb{R}^{3n})$ is a continuous function, then in $a(h)\psi$ we can take h to be a δ-function and define an operator $a(x)$ by

$$(a(x)\psi)(x_1, \ldots, x_{n-1}) = \sqrt{n}\psi(x, x_1, \ldots, x_{n-1}) \tag{5.73}$$

We recover the original operator by $a(h)\psi = \int \overline{h(x)}(a(x)\psi)dx$. We also get an operator $a(x)$ on the Fock space by $(a(x)\psi)_n = a(x)\psi_{n+1}$. Then

$$(a(x)\psi)_n(x_1, \ldots, x_n) = \sqrt{n+1}\ \psi_{n+1}(x, x_1, \ldots, x_n)dx \tag{5.74}$$

As a suitable dense domain we might take finite particle vectors with wave functions in Schwartz space

$$\mathcal{D}_{\mathcal{S}} = \{\psi \in \mathcal{D}_0 : \psi_n \in \mathcal{S}(\mathbb{R}^{3n})\} \tag{5.75}$$

The operator $a(x)$ has no adjoint; the formal adjoint creates delta functions which are not in L^2. Nevertheless we can define $a^*(x)$ as the bilinear form on $\mathcal{D}_{\mathcal{S}} \times \mathcal{D}_{\mathcal{S}}$, anti-linear in the first factor, which sends ψ, ϕ to

$$(\psi, a^*(x)\phi) \equiv (a(x)\psi, \phi) \tag{5.76}$$

With this interpretation we have an elegant representation of some of our basic operators:

Lemma 5.5 *In $\mathcal{F}^{\pm}(\mathcal{H})$, $\mathcal{H} = L^2(\mathbb{R}^3)$, as bilinear forms on $\mathcal{D}_{\mathcal{S}} \times \mathcal{D}_{\mathcal{S}}$:*

1. *The number operator satisfies*

$$N = \int a^*(x)a(x)dx \tag{5.77}$$

2. *The free Hamiltonian $H_0 = d\Gamma(-\Delta/2m)$ satisfies*

$$H_0 = \int a^*(x)\left(\frac{-\Delta}{2m}\right)a(x)dx \tag{5.78}$$

3. *Let v be an interparticle potential, for $n \geq 2$ let V_n be the associated n-particle potential defined in (5.39), and let $V = \oplus_n V_n$ be the Fock space potential. Then*

$$V = \frac{1}{2}\int a^*(x)a^*(y)v(x-y)a(x)a(y)dxdx \tag{5.79}$$

Proof Let $\psi, \phi \in \mathcal{D}_{\mathcal{S}}$. Then we have

$$\begin{aligned}
&(\psi, (\int a^*(x)a(x)dx)\phi) \equiv \int (a(x)\psi, a(x)\phi)dx \\
&= \sum_{n=0}^{\infty}\int ((a(x)\psi)_n, (a(x)\phi)_n)dx \\
&= \sum_{n=0}^{\infty}(n+1)\int \overline{\psi_{n+1}(x, x_1, \ldots, x_n)}\phi_{n+1}(x, x_1, \ldots, x_n)dxdx_1 \ldots dx_n \\
&= (\psi, N\phi)
\end{aligned} \tag{5.80}$$

This proves the identity for N, and the identity for H_0 is similar. For the last we have

$$\begin{aligned}
&\int \frac{1}{2}\,(a(x)a(y)\psi, a(x)a(y)\phi)\; v(x-y)dxdy \\
&= \sum_{n=0}^{\infty} \frac{(n+1)(n+2)}{2} \int \overline{\psi_{n+2}(x,y,x_1,\ldots,x_n)}\phi_{n+2}(x,y,x_1,\ldots,x_n) \\
&\qquad v(x-y)\; dxdy\; dx_1 \ldots dx_n \\
&= \sum_{n=2}^{\infty} \frac{n(n-1)}{2} \int \overline{\psi_n(x_1,\ldots,x_n)}\phi_n(x_1,\ldots,x_n)v(x_1-x_2)\; dx_1 \ldots dx_n \qquad (5.81) \\
&= \sum_{n=2}^{\infty} \sum_{1\le i<j\le n} \int \overline{\psi_n(x_1,\ldots,x_n)}\phi_n(x_1,\ldots,x_n)v(x_i-x_j)\; dx_1 \ldots dx_n \\
&= \sum_{n=2}^{\infty} (\psi_n, V_n\phi_n) \\
&= (\psi, V\phi)
\end{aligned}$$

□

Notes on chapter 5: For more on multiparticle quantum mechanics see Reed and Simon (1979), Reed and Simon (1978), or Gustafson and Sigal (2003).

6 Statistical mechanics

6.1 Mixed states

Until now the states of a physical system have been described by unit vectors (actually rays) in a Hilbert space. These are states which are prepared so that we have as much knowledge about them as possible. They are also called *pure states*. But we also want to consider states whose preparation is incomplete. We only know the probability that it is in any of various pure states. These are known as *mixed states*.

The mathematical definition is that a mixed state is a positive trace class operator Q on the Hilbert space $\mathcal{H}$ with $Tr(Q) = 1$. The operator Q is called a *density operator*. If $\{\phi_n\}$ is an orthonormal basis of eigenvectors for Q with eigenvalues $\mu_n \geq 0$, then $Q\psi = \sum_{n=1}^{\infty} \mu_n \phi_n (\phi_n, \psi)$ which we write as

$$Q = \sum_{n=1}^{\infty} \mu_n \phi_n (\phi_n, \cdot) \tag{6.1}$$

The condition $Tr(Q) = 1$ means that

$$\sum_n \mu_n = 1 \tag{6.2}$$

If a property of a physical measurement is described by a projection operator P, then the probability of a positive result in state Q is taken to be

$$Tr(PQ) = \sum_{n=1}^{\infty} \mu_n (\phi_n, P\phi_n) \tag{6.3}$$

Taking $P = P_\psi = \psi(\psi, \cdot)$, the projection onto ψ, the probability of finding the system in the pure state ψ is

$$Tr(P_\psi Q) = \sum_{n=1}^{\infty} \mu_n |(\psi, \phi_n)|^2 \tag{6.4}$$

In particular μ_n is the probability of finding the system in the state ϕ_n.

Observable quantities are still described by self-adjoint operators A. If $E(B)$ are the spectral projections for A, then the probability that a measurement of A in the state Q gives a value in B is

$$Tr(E(B)Q) = \sum_{n=1}^{\infty} \mu_n(\phi_n, E(B)\phi_n) \tag{6.5}$$

The expected value of repeated measurements of A in the state Q is

$$Tr(AQ) = \sum_{n=1}^{\infty} \mu_n(\phi_n, A\phi_n) \tag{6.6}$$

if it exists. It exists if A is bounded, and it may or may not exist for unbounded operators.

Note that we can identify pure states with mixed states of rank one via the map $\psi \to Q_\psi = \psi(\psi, \cdot)$. This is defined on rays: ψ and $e^{i\theta}\psi$ give the same mixed state. The probabilities $Tr(PQ_\psi) = (\psi, P\psi)$ are the same as before.

6.2 Equilibrium states

We describe some mixed states appropriate for describing large numbers of particles called *equilibrium states*. In this section we give a general discussion of both the classical and quantum versions, but in subsequent sections we only consider the quantum version in detail. In the classical versions the states are probability measures on phase space – the measure of a set is the probability of finding the system in that set. In the quantum versions the states are density operators as described above – we assign a probability to finding the system in various pure states.

For this discussion we suppose that we are in a bounded open region $\Lambda \subset \mathbb{R}^3$ or possibly the torus $\Lambda = \mathbb{R}^3/L\mathbb{Z}^3$ of width L. The important point is that Λ have finite volume.

6.2.1 Microcanonical ensemble

The first case is a system which is isolated from its surroundings and has a fixed energy E and a fixed number of particles N. With no further knowledge of the system an appropriate state is one which assigns equal weight to all states with this energy and particle number. States which enter into such a description with fixed E and N are said to constitute a *microcanonical ensemble*. One can create such states in either a classical or a quantum version. However we do not go into details.

6.2.2 Canonical ensemble

In the second case the system still has a fixed number of particles N, but now the energy E is not fixed due to interactions with its surroundings. The surroundings

are taken to be a heat bath at a temperature T. The temperature is a measure of the average kinetic energy of particles in the bath. We do not attempt a mathematical description of the heat bath or of its interaction with the system, but we do make a hypothesis about the probabilities for states of the system in this circumstance. The fundamental hypothesis is that the probability of finding the system in a state with energy E is proportional to $e^{-E/kT}$. Here $k = 1.38 \times 10^{-16}$ ergs/$^\circ K$ is a constant which sets the temperature scale and is known as Boltzmann's constant. We usually let $\beta = 1/kT$ and write $e^{-E/kT} = e^{-\beta E}$. With probabilities assigned in this fashion the states are called *Gibbs states* and are said to constitute a *canonical ensemble*. We now spell out the construction in more detail.

In the classical case the phase space is $\mathcal{P} = \Lambda^N \times \mathbb{R}^{3N}$ and we have a Hamiltonian $H_N(x, p)$ on this space. We define a probability measure on $\mathcal{P}$ by

$$d\mu_\beta(x, p) = \frac{1}{Z(\beta)} e^{-\beta H_N(x,p)} dxdp \tag{6.7}$$

The normalizing factor is known as the *partition function* and it is given by

$$Z(\beta) = \int e^{-\beta H_N(x,p)} dxdp \tag{6.8}$$

This integral is required to converge. A classical observable A is a function on phase space, hence a random variable, and its expected value is

$$< A >_\beta = \int A(x, p) d\mu_\beta(x, p) = \frac{\int A(x,p) e^{-\beta H_N(x,p)} dxdp}{\int e^{-\beta H_N(x,p)} dxdp} \tag{6.9}$$

In the quantum case there is a single particle space $\mathcal{H}$, for example $\mathcal{H} = L^2(\Lambda)$, and an N-particle space $\mathcal{H}_N^\pm = \Pi^\pm(\mathcal{H} \otimes \cdots \otimes \mathcal{H})$. The Hamiltonian H_N is required to be a self-adjoint operator on this space. The basic hypothesis is that the state at inverse temperature β has the density operator

$$Q_\beta = Z(\beta)^{-1} e^{-\beta H_N} \tag{6.10}$$

where the partition function is now

$$Z(\beta) = Tr(e^{-\beta H_N}) \tag{6.11}$$

For this to make sense we need $e^{-\beta H_N}$ to be trace class. An observable A is a self-adjoint operator on $\mathcal{H}_N$ and the expected value is

$$< A >_\beta \equiv Tr(AQ_\beta) = \frac{Tr(Ae^{-\beta H_N})}{Tr(e^{-\beta H_N})} \tag{6.12}$$

The quantum state is invariant under time evolution in the following sense. If $A_t = e^{iH_N t} A e^{-iH_N t}$ is the time evolution of A in the Heisenberg picture, then

$$< A_t >_\beta = < A >_\beta \tag{6.13}$$

This follows from the cyclicity of the trace. We say that the state is a *stationary state*.

6.2.3 Grand canonical ensemble

In the third case neither E or N is fixed. Instead the system is supposed to be in contact with a heat bath and a particle bath. The basic hypothesis is that the probability of finding the system in a state with energy E and particle number N is proportional to $\exp(-\beta(E-\mu N))$. Here β is again the inverse temperature, and μ is a parameter called the *chemical potential*. A particle has energy $-\mu$ just by its presence in the system. States weighted in this fashion constitute a *grand canonical ensemble*.

In the classical case the phase space is the disjoint union of the n-particle phase spaces $\Omega = \cup_{n=0}^{\infty}\Omega_n$. A measure $\mu_{\beta,\mu}$ is defined on Ω by stipulating that its restriction to Ω_n is

$$d\mu_{\beta,\mu} \mid \Omega_n = Z(\beta,\mu)^{-1}\exp(-\beta(H(x,p)-\mu n))dxdp \tag{6.14}$$

The partition function which gives the overall normalization factor is

$$Z(\beta,\mu) = \sum_{n=0}^{\infty}\int_{\Omega_n}\exp(-\beta(H(x,p)-\mu n))dxdp \tag{6.15}$$

One can also define expectations of classical observables.

In the quantum case the Hilbert space is taken to be the Fock space $\mathcal{F}^{\pm}(\mathcal{H}) = \oplus_n \mathcal{H}_n^{\pm}$ and the Hamiltonian is $H = \oplus_n H_n$ where H_n is the n-particle Hamiltonian. The density operator has the form

$$Q_{\beta,\mu} = Z(\beta,\mu)^{-1}e^{-\beta(H-\mu N)} \tag{6.16}$$

where now N is the number operator. The partition function is

$$Z(\beta,\mu) = Tr(e^{-\beta(H-\mu N)}) \tag{6.17}$$

An observable is a self-adjoint operator A on the Fock space and has the expectation

$$< A >_{\beta,\mu} \equiv Tr(AQ_{\beta,\mu}) = \frac{Tr(Ae^{-\beta(H-\mu N)})}{Tr(e^{-\beta(H-\mu N)})} \tag{6.18}$$

This is also a stationary state.

6.2.4 General problems

One set of problems is concerned with motivating the above discussion. A great deal of effort has been expended over the years attempting to derive the various ensembles from more basic hypotheses, and to study the relation between them. There are many interesting developments here but it would take us too far afield to explore them. We just take the ensembles as defined as our starting point.

Another class of problems is concerned with picking specific models or classes of models and studying detailed properties of the states. We do this for a few simple models working in the grand canonical ensemble. This will just give a taste of what is a very large subject.

Let us first mention some items of interest for the quantum grand canonical ensemble. We define a *free energy* in terms of the partition function by

$$F(\beta, \mu) = -\beta^{-1} \log Z(\beta, \mu) \tag{6.19}$$

It turns out this can be interpreted as the amount of energy available to do work. The *pressure* is defined as minus the free energy per unit volume and is given by[1]

$$p(\beta, \mu) = -|\Lambda|^{-1} F(\beta, \mu) = \beta^{-1} |\Lambda|^{-1} \log Z(\beta, \mu) \tag{6.20}$$

where $|\Lambda|$ is the volume of Λ. The expected number of particles is

$$< N >_{\beta,\mu} = Z(\beta, \mu)^{-1} Tr(N\, e^{-\beta(H-\mu N)}) \tag{6.21}$$

It can be computed from the partition function by

$$< N >_{\beta,\mu} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log Z(\beta, \mu) \tag{6.22}$$

The *density* is the expected number of particles divided by the volume

$$\rho(\beta, \mu) = |\Lambda|^{-1} < N >_{\beta,\mu} \tag{6.23}$$

Recall that the number operator can be expressed as $N = \int a^*(x) a(x) dx$. We could also consider the number of particles in a region $B \subset \mathbb{R}^3$ defined by $N_B = \int_B a^*(x) a(x) dx$ or the kinetic energy in B which would be defined by $\int_B a^*(x)(-\Delta/2m) a(x) dx$. Thus if we knew the expectations $< a^*(x) a(y) >_{\beta,\mu}$, we could compute expectations of many interesting observables. More generally we would like to compute *correlation functions* defined by

$$< a^*(x_1) \cdots a^*(x_n) a(y_1) \cdots a(y_n) >_{\beta,\mu} \tag{6.24}$$

These may or may not be well-defined. The situation is improved if we replace $a(x)$ by the more regular $a(f) = \int f(x) a(x) dx, f \in \mathcal{S}(\mathbb{R}^3)$. Then the correlation functions are

$$< a^*(f_1) \cdots a^*(f_n) a(g_1) \cdots a(g_n) >_{\beta,\mu} \tag{6.25}$$

If it exists, this gives (6.24) a meaning as a distribution. Indeed it is a multilinear functional on $\mathcal{S}(\mathbb{R}^3)$ and hence by the kernel theorem determines a distribution in $\mathcal{S}'(\mathbb{R}^{3n})$. (See appendix C for the basic facts about distributions.)

[1] The definition says that pressure has dimensions of energy/volume. But since energy has dimensions of force × distance, pressure has dimensions of force/area as expected.

All these quantities depend on the volume Λ, and generally the volume must be finite in order that they be well-defined. At first this seems to correlate well with actual physical situations where the volume really is finite. However volumes are typically very large on an atomic scale, and so it is a good idealization to treat the system as infinite. If it can be done mathematically, it is worthwhile because then one avoids uninteresting boundary effects. Furthermore collective phenomena like phase transitions generally have more dramatic manifestations at infinite volume. In short it is not enough to study the quantities $p(\beta,\mu)$, $\rho(\beta,\mu)$ and correlation functions in a finite volume Λ, one should also take the limit $\Lambda \to \mathbb{R}^3$, known as the *thermodynamic limit*.

6.3 Free boson gas

We consider the case of free bosons in detail. We take Λ to be the torus $\Lambda = \mathbb{R}^3/L\mathbb{Z}^3$. Making this choice is the same as taking Λ to be the cube $[-L/2, L/2]^3$ and imposing periodic boundary conditions. This might not seem like a good starting point for an actual gas. Perhaps a better model would be the cube with some local boundary conditions. We make the choice anyway with the idea that the infinite volume limit should be independent of the boundary conditions.

For a single particle the Hilbert space is $\mathcal{H} = L^2(\Lambda)$ and specializing to $m = 1/2$ the Hamiltonian is $-\Delta$ as for the $\mathbb{R}^3$ construction. The trigonometric polynomials are functions in $L^2(\Lambda)$ of the form

$$\phi_k(x) = \frac{e^{ikx}}{L^{3/2}} \qquad k \in \frac{2\pi}{L}\mathbb{Z}^3 \tag{6.26}$$

These form a complete orthonormal set in $L^2(\Lambda)$. This statement is equivalent to the L^2- convergence of Fourier series. Furthermore the ϕ_k are eigenfunctions of $-\Delta$

$$-\Delta\phi_k = |k|^2\phi_k \tag{6.27}$$

Thus $-\Delta$ is naturally a positive self-adjoint operator by problem 1.13. As in the $\mathbb{R}^3$ construction, the k are called momenta.

Now consider n such particles. The Hilbert space for bosons is now $\mathcal{H}_n^+ = \Pi^+(\mathcal{H}\otimes\cdots\otimes\mathcal{H})$ and the Hamiltonian is $H_n = d\Gamma_n(-\Delta)$. States $\phi_{k_1}\otimes\cdots\otimes\phi_{k_n}$ form a basis for $\mathcal{H}\otimes\cdots\otimes\mathcal{H}$ and so states

$$\Phi_{k_1,\ldots,k_n} = \Pi^+(\phi_{k_1}\otimes\cdots\otimes\phi_{k_n}) \tag{6.28}$$

span $\mathcal{H}_n^+$. Note that $\Phi_{k_1,\ldots,k_n}$ depends only on the collection $\{k_1,\ldots,k_n\}$, not on the ordering. Different collections give orthogonal states. They are eigenfunctions of H_n and satisfy

$$H_n \Phi_{k_1,\dots,k_n} = \left(\sum_{i=1}^{n} |k_i|^2 \right) \Phi_{k_1,\dots,k_n} \tag{6.29}$$

We reformulate as follows. First identify $\mathcal{H}_n^+$ as a subspace of the Fock space $\mathcal{F}^+(\mathcal{H})$ and write $\Phi_{k_1,\dots,k_n}$ as a constant times $a^*(\phi_{k_1}) \dots a^*(\phi_{k_n})\Omega_0$ (see (5.60)). Then label these basis vectors by the number of times a particular momentum occurs. Thus let $\{n_k\}$ be a collection of nonnegative integers indexed by $k \in (2\pi/L)\mathbb{Z}^3$ such that $\sum_k n_k = n$. For each such collection define

$$\Phi(\{n_k\}) = \left(\prod_k \frac{1}{\sqrt{n_k!}} \right) \prod_k a^*(\phi_k)^{n_k} \Omega_0 \tag{6.30}$$

With this choice of normalization the $\Phi(\{n_k\})$ form an orthonormal basis for $\mathcal{H}_n^+$ (problem 6.1). We have

$$H_n \Phi(\{n_k\}) = \left(\sum_k n_k |k|^2 \right) \Phi(\{n_k\}) \tag{6.31}$$

Finally consider the full Fock space. We drop the restriction $\sum_k n_k = n$ and instead take infinite sequences $\{n_k\}$ with the condition that $n_k = 0$ except for a finite number of k. Then the $\Phi(\{n_k\})$ form an orthonormal basis for the entire Fock space. Furthermore $\Phi(\{n_k\})$ is an eigenvector for the full Hamiltonian H with eigenvalue $\sum_k n_k |k|^2$. Note also that $\Phi(\{n_k\})$ is an eigenvector for N with eigenvalue $\sum_k n_k$.

Now we are ready to calculate the partition function in the grand canonical ensemble. We have for $\mu < 0$

$$\begin{aligned} Z(\beta,\mu) &= Tr(e^{-\beta(H-\mu N)}) \\ &= \sum_{\{n_k\}} (\Phi(\{n_k\}), e^{-\beta(H-\mu N)} \Phi(\{n_k\})) \\ &= \sum_{\{n_k\}} \exp\left(-\beta \sum_k n_k (|k|^2 - \mu) \right) \\ &= \sum_{\{n_k\}} \prod_k e^{-\beta n_k (|k|^2 - \mu)} \\ &= \prod_k \sum_{n=0}^{\infty} e^{-\beta n (|k|^2 - \mu)} \\ &= \prod_k \frac{1}{1 - e^{-\beta(|k|^2 - \mu)}} \end{aligned} \tag{6.32}$$

The infinite product converges since

$$\sum_k \left(\frac{1}{1 - e^{-\beta(|k|^2-\mu)}} - 1 \right) = \sum_k \frac{e^{-\beta(|k|^2-\mu)}}{1 - e^{-\beta(|k|^2-\mu)}} < \infty \tag{6.33}$$

From the partition function we compute the pressure (6.20) and the density (6.23) and find

$$\begin{aligned} p(\beta,\mu) &= |\Lambda|^{-1}\beta^{-1}\sum_k \log\left(\frac{1}{1-e^{-\beta(|k|^2-\mu)}}\right) \\ \rho(\beta,\mu) &= |\Lambda|^{-1}\sum_k \frac{e^{-\beta(|k|^2-\mu)}}{1-e^{-\beta(|k|^2-\mu)}} \end{aligned} \tag{6.34}$$

These expressions also have a nice infinite volume limit. As $L\to\infty$ the sum over $k\in(2\pi/L)\mathbb{Z}^3$ becomes an integral over $k\in\mathbb{R}^3$ and we find

$$\begin{aligned} p(\beta,\mu) &= (2\pi)^{-3}\beta^{-1}\int \log\left(\frac{1}{1-e^{-\beta(|k|^2-\mu)}}\right)dk \\ \rho(\beta,\mu) &= (2\pi)^{-3}\int \frac{e^{-\beta(|k|^2-\mu)}}{1-e^{-\beta(|k|^2-\mu)}}\,dk \end{aligned} \tag{6.35}$$

Returning to finite volume, the correlation functions can also be computed. We illustrate with the two-point function.

Lemma 6.1 *Let $f,g\in\mathcal{C}^\infty(\Lambda)$ and let $h=-\Delta$. Then for $\mu<0$*

$$< a^*(f)a(g) >_{\beta,\mu} = \left(g,\left(\frac{e^{-\beta(h-\mu)}}{1-e^{-\beta(h-\mu)}}\right)f\right) \tag{6.36}$$

Remark Explicitly

$$< a^*(f)a(g) >_{\beta,\mu} = \sum_k \bar{g}_k\left(\frac{e^{-\beta(|k|^2-\mu)}}{1-e^{-\beta(|k|^2-\mu)}}\right)f_k \tag{6.37}$$

where $f_k=(\phi_k,f)$ are the Fourier coefficients for f. The expression (6.36) also holds in the infinite volume limit, but now defined with the Fourier transform instead of Fourier series.

Proof Since $e^{-\beta(H-\mu N)}=\Gamma(e^{-\beta(h-\mu)})$, we can compute (cf. problem 5.6)

$$e^{-\beta(H-\mu N)}a^*(f) = a^*(e^{-\beta(h-\mu)}f)e^{-\beta(H-\mu N)} \tag{6.38}$$

Using the cyclicity of the trace and then the commutation relations for a,a^*, (5.61) yields

$$\begin{aligned} &Tr(a^*(f)a(g)e^{-\beta(H-\mu N)}) = Tr(a(g)e^{-\beta(H-\mu N)}a^*(f)) \\ &= Tr(a(g)a^*(e^{-\beta(h-\mu)}f)e^{-\beta(H-\mu N)}) \\ &= Tr(a^*(e^{-\beta(h-\mu)}f)a(g)e^{-\beta(H-\mu N)}) + Z(\beta,\mu)(g,e^{-\beta(h-\mu)}f) \end{aligned} \tag{6.39}$$

Dividing by $Z(\beta,\mu)$ this can be written

$$< a^*((1-e^{-\beta(h-\mu)})f)a(g) >_{\beta,\mu} = (g, e^{-\beta(h-\mu)}f) \tag{6.40}$$

Now replacing f by $(1-e^{-\beta(h-\mu)})^{-1}f$ gives the result. □

Problem 6.1 Check that $\|\Phi(\{n_k\})\| = 1$.

Problem 6.2 Prove that pressure and density in (6.35) really are the infinite volume limits of the finite volume expressions (6.34) as claimed.

Problem 6.3 Take $\Lambda = [0,L]^N$. On $\mathcal{H} = L^2(\Lambda)$ define $-\Delta$ as a self-adjoint operator by taking eigenfunctions with Dirichlet boundary conditions, that is vanishing on the boundary. Compute the grand canonical partition function in this case.

Problem 6.4 Compute the n-point correlation functions (6.25) for the free boson gas by establishing first that

$$\begin{aligned} &< a^*(f_1)\cdots a^*(f_n)a(g_1)\cdots a(g_n) >_{\beta,\mu} \\ &= \sum_{j=1}^{n} < a^*(f_1)a(g_j) >_{\beta,\mu} < a^*(f_2)\cdots a^*(f_n)a(g_1)\cdots \widehat{a(g_j)}\cdots a(g_n) >_{\beta,\mu} \end{aligned} \tag{6.41}$$

where the factor $a(g_j)$ is omitted in the last expectation. Then show that

$$< a^*(f_1)\cdots a^*(f_n)a(g_1)\cdots a(g_n) >_{\beta,\mu} = \sum_{\pi}\prod_{i=1}^{n} < a^*(f_i)a(g_{\pi(i)}) >_{\beta,\mu} \tag{6.42}$$

where the sum is over permutations π of $(1,\ldots,n)$.

6.4 Free fermion gas

We consider noninteracting fermions. For simplicity we neglect spin and take the one particle space to be $\mathcal{H} = L^2(\Lambda)$. The treatment is the same as for bosons except that the full Hilbert space is now the anti-symmetric Fock space $\mathcal{F}^-(\mathcal{H})$. We still have the basis vectors $\Phi(\{n_k\})$ given by (6.30) but now the Pauli exclusion principle means that each n_k can only take the values zero and one.

The partition function in the grand canonical ensemble is now computed just as in (6.32)

$$\begin{aligned} Z(\beta,\mu) &= Tr(e^{-\beta(H-\mu N)}) \\ &= \sum_{\{n_k\}} \prod_k e^{-\beta n_k(|k|^2-\mu)} \\ &= \prod_k \sum_{n=0,1} e^{-\beta n(|k|^2-\mu)} \\ &= \prod_k \left(1 + e^{-\beta(|k|^2-\mu)}\right) \end{aligned} \tag{6.43}$$

This converges for all $\beta > 0$ and all $\mu \in \mathbb{R}$ since

$$\sum_k e^{-\beta(|k|^2-\mu)} < \infty \tag{6.44}$$

For the pressure and the density we compute

$$\begin{aligned} p(\beta,\mu) &= |\Lambda|^{-1}\beta^{-1} \sum_k \log(1 + e^{-\beta(|k|^2-\mu)}) \\ \rho(\beta,\mu) &= |\Lambda|^{-1} \sum_k \frac{e^{-\beta(|k|^2-\mu)}}{1 + e^{-\beta(|k|^2-\mu)}} \end{aligned} \tag{6.45}$$

One can also compute the two-point function as for bosons and we find with $h = -\Delta$

$$< a^*(f)a(g) >_{\beta,\mu} = \left(g, \left(\frac{e^{-\beta(h-\mu)}}{1 + e^{-\beta(h-\mu)}}\right) f\right) \tag{6.46}$$

Note that the only difference from bosons is the plus sign in the denominator.

Next we investigate the zero temperature limit. Let N_k be the number of particles with momentum k defined by $N_k\Phi(\{n_k\}) = n_k\Phi(\{n_k\})$. Then $N_k = a^*(\phi_k)a(\phi_k)$ and from (6.46) the expectation is

$$< N_k >_{\beta,\mu} = \frac{e^{-\beta(|k|^2-\mu)}}{1 + e^{-\beta(|k|^2-\mu)}} \tag{6.47}$$

For $\mu > 0$ we have the zero temperature limit

$$\lim_{\beta\to\infty} < N_k >_{\beta,\mu} = \begin{cases} 1 & |k|^2 < \mu \\ 0 & |k|^2 > \mu \end{cases} \tag{6.48}$$

This is saying that at zero temperature all states with energy $|k|^2 < \mu$ are occupied while states with energy $|k|^2 > \mu$ are empty. This corresponds to the lowest energy state and is referred to as the *Fermi sea*.

The situation for bosons is quite different as there is no exclusion principle. A more extensive analysis shows that at fixed density below a certain critical temperature a substantial fraction of particles occupy the lowest energy state $k = 0$, and at zero temperature they all occupy it. This phenomenon is known as *Bose–Einstein condensation*.

6.5 Interacting bosons

Now we consider the statistical mechanics of interacting bosons on the three-dimensional torus $\Lambda = \mathbb{R}^3/L\mathbb{Z}^3$; there is a similar treatment for fermions. The n-particle Hilbert space is now $L^2_+(\Lambda^n)$, the symmetric subspace of $L^2(\Lambda^n)$. The n-particle Hamiltonian is

$$H_n = H_{0,n} + V_n = \sum_{i=1}^{n} \frac{-\Delta_i}{2m} + \sum_{1\le i<j\le n} v(x_i - x_j) \tag{6.49}$$

where the interparticle potential v is taken to be in $L^2(\Lambda)$. Then one can show that H_n is self-adjoint on the domain of H_0. (See example 5.2 and problem 5.1 for the $\mathbb{R}^3$ result.)

We assume that the following *stability condition* is satisfied: there is a constant B such that for all n and all points $x_i \in \Lambda$

$$\sum_{1\le i<j\le n} v(x_i - x_j) \ge -Bn \tag{6.50}$$

This is a fairly restrictive condition. It is trivially satisfied if $v \ge 0$ and obviously false if $v(0) < 0$ (take all $\{x_i\}$ coincident and n large). Physically it means that particles have a hard core repelling other particles.

Problem 6.5 v is said to be of positive type if the Fourier series has nonnegative coefficients. Show that the stability condition is satisfied if v is of positive type

For a stable interaction we have[2]

$$0 \le H_{0,n} - (B+\mu)n \le H_n - \mu n \tag{6.51}$$

where the first inequality holds if $\mu < -B$. Under the same assumption the full Hamiltonian on the Fock space satisfies

$$0 \le H_0 - (B+\mu)N \le H - \mu N \tag{6.52}$$

Our modest goal is to prove the existence of the grand canonical ensemble on the torus, that is we want to show that $\exp(-\beta(H-\mu N))$ is trace class. First we have:

Lemma 6.2 *For $\mu < -B$ the Hamiltonian $H - \mu N$ has pure point spectrum with finite multiplicity and no accumulation points.*

Proof Let $S = H - \mu N + 1$ and $T = H_0 - (\mu + B)N + 1$ so $1 \le T \le S$. We have seen in our treatment of the free boson gas that T has point spectrum with finite multiplicity and no accumulation points. Hence $T^{-1/2}$ is compact (see problem 1.17).

[2] $T \le S$ means $D(S) \subset D(T)$ and $(\psi, T\psi) \le (\psi, S\psi)$ for $\psi \in D(S)$.

Now $D(T) = D(S)$ is contained in both $D(T^{1/2})$ and $D(S^{1/2})$ by the spectral theorem. The inequality $T \leq S$ implies that for $\psi \in D(S)$

$$\|T^{1/2}\psi\| \leq \|S^{1/2}\psi\| \tag{6.53}$$

But $D(S)$ is a core for $D(S^{1/2})$ again by the spectral theorem. It follows that $D(S^{1/2}) \subset D(T^{1/2})$ and that the inequality (6.53) holds for $\psi \in D(S^{1/2})$. (We made a similar argument in lemma 5.3.) Replacing ψ by $S^{-1/2}\chi$ for any χ we see that $T^{1/2}S^{-1/2}$ is a bounded operator with $\|T^{1/2}S^{-1/2}\| \leq 1$.

Now write

$$S^{-1/2} = T^{-1/2}(T^{1/2}S^{-1/2}) \tag{6.54}$$

This exhibits $S^{-1/2}$ as the product of a compact operator and a bounded operator. Hence it is compact and has point spectrum with finite multiplicity and no accumulation points except zero. The result for $H - \mu N = (S^{-1/2})^{-2} - 1$ follows. □

We also need:

Theorem 6.1 *(min–max theorem) Let T be a self-adjoint operator on a Hilbert space with pure point spectrum*

$$\lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots \tag{6.55}$$

repeated by multiplicity. Then

$$\lambda_n = \sup_{\xi_1,\dots,\xi_{n-1}} \left[\inf_{\psi \in [\xi_1,\dots,\xi_{n-1}]^\perp, \|\psi\|=1} (\psi, T\psi) \right] \tag{6.56}$$

where the infimum is over $\psi \in D(T)$.

The theorem is useful because it gives control over the eigenvalues without knowing the eigenvectors. We do not give the full proof.[3] But observe that for $n = 0$ it says

$$\lambda_0 = \inf_{\|\psi\|=1} (\psi, T\psi) \tag{6.57}$$

This is true since if ϕ_n are a basis of eigenvectors with $T\phi_n = \lambda_n\phi_n$, then for any $\psi \in D(T)$ with $\|\psi\| = 1$

$$(\psi, T\psi) = \sum_n \lambda_n |(\psi, \phi_n)|^2 \geq \sum_n \lambda_0 |(\psi, \phi_n)|^2 = \lambda_0 \|\psi\|^2 = \lambda_0 \tag{6.58}$$

On the other hand the lower bound λ_0 is actually attained at $\psi = \phi_0$.

Corollary 6.1 *Let T, S satisfy the hypotheses of the theorem and suppose $T \leq S$. Then the eigenvalues $\lambda_n(T), \lambda_n(S)$ satisfy $\lambda_n(T) \leq \lambda_n(S)$.*

[3] For the proof see Reed and Simon (1978: 76).

Proof We have

$$\inf_{\psi\in[\xi_1,\ldots,\xi_{n-1}]^\perp,\|\psi\|=1}(\psi,T\psi)\leq\inf_{\psi\in[\xi_1,\ldots,\xi_{n-1}]^\perp,\|\psi\|=1}(\psi,S\psi) \tag{6.59}$$

Now take the supremum over $\xi_1,\ldots,\xi_{n-1}$. □

Now we have the main result:

Theorem 6.2 *If the two particle potential is stable with constant B and if the chemical potential satisfies $\mu < -B$, then the grand canonical partition function $Z(\beta,\mu) = Tr(\exp(-\beta(H-\mu N)))$ exists.*

Proof Let $\lambda_n(H-\mu N)$ be the eigenvalues of $H-\mu N$ repeated by multiplicity. By (6.52) and the corollary we have

$$\lambda_n(H-\mu N)\geq\lambda_n(H_0-(\mu+B)N) \tag{6.60}$$

Then

$$\begin{aligned} Tr\left(e^{-\beta(H-\mu N)}\right) &= \sum_n e^{-\beta\lambda_n(H-\mu N)} \\ &\leq \sum_n e^{-\beta\lambda_n(H_0-(\mu+B)N)} \\ &= Tr\left(e^{-\beta(H_0-(\mu+B)N)}\right) \end{aligned} \tag{6.61}$$

The last expression is finite by our results on the free boson gas. □

6.6 Further developments

Suppose we are working in the grand canonical ensemble as defined in (6.18). Let $K = H-\mu N$ be the Hamiltonian, let A, B be observables, and let

$$\alpha_t(B) = e^{iKt}Be^{-iKt} \tag{6.62}$$

be the time evolution of B. Then we have

$$<A\ \alpha_t(B)>_{\beta,\mu} = \frac{Tr(Ae^{iKt}Be^{-(\beta+it)K})}{Tr(e^{-\beta K})} \tag{6.63}$$

Assuming that $K\geq 0$ the function e^{iKt} is the boundary value on the real axis of an analytic function from the upper half plane to the bounded operators on Fock space. Then (6.63) is the boundary value on the real axis of a function analytic in the strip $0 < \mathrm{Im}\, t < \beta$. But by the cyclicity of the trace we also have

$$Tr(Ae^{iKt}Be^{-(\beta+it)K}) = Tr(Be^{-(\beta+it)K}Ae^{iKt}) \tag{6.64}$$

Taking the boundary value at $t = i\beta$ we obtain the identity

$$< A\ \alpha_t(B) >_{\beta,\mu} |_{t=i\beta} =< BA >_{\beta,\mu} \tag{6.65}$$

This is known as the *KMS condition*.[4] It turns out that the KMS condition completely characterizes the expectation $< \cdots >_{\beta,\mu}$, and so gives an alternate definition of equilibrium states. An advantage is that it also makes sense in infinite volume.

Let us discuss further the general problem of infinite volume. We would like to have a full quantum theory in infinite volume, rather than just certain special limits of finite volume quantities. This turns out to be possible and we sketch the idea. One must modify the basic quantum structure founded on a Hilbert space. Instead the basic object is taken to be a C^* algebra[5] $\mathcal{A}$ whose self-adjoint elements correspond to observables. For example it might be the C^* algebra generated by creation and annihilation operators on Fock space. A state is now specified by giving the expectations of all the elements of $\mathcal{A}$. More precisely a state ω is a continuous positive[6] linear functional on $\mathcal{A}$ with norm one. Finite volume Gibbs states are states in this sense and so are infinite volume limits of Gibbs states. Time evolution is given by a family α_t of automorphisms of $\mathcal{A}$ such that $\alpha_0 = id$ and $\alpha_t\alpha_s = \alpha_{t+s}$. For example α_t could be generated by a Hamiltonian as in (6.62). Finally ω is defined to be an equilibrium state at inverse temperature β if the β-KMS condition is satisfied, that is if for $A, B \in \mathcal{A}$ the function $t \to \omega(A\ \alpha_t(B))$ is the boundary value of a function analytic in $0 < \mathrm{Im}\, t < \beta$ such that

$$\omega(A\ \alpha_t(B))|_{t=i\beta} = \omega(BA) \tag{6.66}$$

These ideas can be carried a great deal further with beneficial results.

Notes on chapter 6: For the basic structure of statistical mechanics see Ruelle (1969). For quantum statistical mechanics as we have presented it see Bratteli and Robinson (1981); for the algebraic approach consult Bratteli and Robinson (1981) and Haag (1992); for stability bounds with realistic potentials see Lieb and Seiringer (2010).

For physics books on statistical mechanics, try Landau and Lifschitz (1969) or Huang (2009).

There are important topics in statistical mechanics which have not been discussed at all. These include phase transitions, critical phenomena, and nonequilibrium statistical mechanics.

[4] KMS stands for Kubo–Martin–Schwinger.

[5] A C^* algebra is a Banach $*$-algebra with $\|A^*A\| = \|A\|^2$. Think of a closed subalgebra of the algebra of bounded operators on a Hilbert space.

[6] Positive means $\omega(A^*A) \geq 0$ for all $A \in \mathcal{A}$.

Part II

Relativistic

7 Relativity

7.1 Principles of relativity

Einstein arrived at the principles of relativity by thinking about the relationship between measurements made by observers in uniform motion relative to each other. We depart from this historical path and instead start with Maxwell's equations (0.3).

In the absence of charges and currents ($\rho = 0, j = 0$) one can deduce from Maxwell's equations that any component u of the electric or magnetic field obeys the wave equation

$$\left(-\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \Delta\right) u = 0 \tag{7.1}$$

Disturbances propagate with the velocity c which is the speed of light. Hence light is explained as a wave in the electric and magnetic fields.

Suppose we think of space and time as a single entity called *spacetime* and modeled by $\mathbb{R}^4$. A point is labeled $x = (x^0, x^1, x^2, x^3)$ with $x^0 = ct$ a scaled time with units of distance. Then the wave equation can be written

$$\sum_{\mu\nu} \eta^{\mu\nu} \frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} u = 0 \tag{7.2}$$

where $\eta^{\mu\nu}$ is the diagonal matrix with entries $(-1, 1, 1, 1)$. We take these coefficients in the wave equation above as a clue to the structure of spacetime. The coefficients can be interpreted as coming from a metric on $\mathbb{R}^4$ given by

$$\eta = \sum_{\mu\nu} \eta_{\mu\nu} dx^\mu dx^\nu = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \tag{7.3}$$

This is not a positive definite metric, but a Lorentzian metric, that is a metric which has signature $-+++$. The basic postulate of *special relativity* is that spacetime is to be modeled by the pair $(\mathbb{R}^4, \eta)$. The idea is that field equations are to be built out of this metric as above. Furthermore free particles should respect the metric in the sense that their trajectories should be geodesics for the metric. All the structure of special relativity follows from these assumptions, some of which we develop as we go on.

The scope of special relativity is limited as the name suggests. It is only supposed to be valid in situations where there is no gravitational field, or more precisely where

the effect of gravitational fields is negligible. The effect of the gravitational field is to distort the metric and indeed possibly even distort the topology of spacetime. The basic postulate of *general relativity* is that spacetime is to be modeled by a Lorentzian manifold $(\mathcal{M}, g)$ consisting of a four-dimensional manifold $\mathcal{M}$ and a Lorentzian metric g which in local coordinates takes the form

$$g = \sum_{\mu\nu} g_{\mu\nu} dx^\mu dx^\nu \tag{7.4}$$

Field equations are to be constructed out of the metric. For example the wave equation would have the form

$$\sum_{\mu\nu} |\det g|^{-1/2} \frac{\partial}{\partial x^\mu} \left(|\det g|^{1/2} g^{\mu\nu} \frac{\partial u}{\partial x^\nu} \right) = 0 \tag{7.5}$$

where $\{g^{\mu\nu}\}$ is the inverse matrix to $\{g_{\mu\nu}\}$. Again free particles should travel on geodesics. The metric itself is determined by the distribution of energy and matter by an equation known as Einstein's equation.

A manifold is anyway the best model for spacetime with or without gravity. Manifolds are defined to treat all coordinate systems equally and this is matched by the phenomenon that nature has no distinguished coordinate systems. In constructing the basic equations to describe nature one should take care that the construction is not tied to any particular coordinate system. This naturalness condition is sometimes known as the *principle of general covariance*. Einstein's theory of gravity outlined above respects it. Quantum mechanics has a more difficult time with it, primarily because it has a more rigid notion of time.

In the next three chapters we concentrate on combining quantum mechanics with special relativity and ignore gravitational effects. It is appropriate in any case because on microscopic scales gravity is much weaker than electromagnetic or nuclear forces and thus is usually negligible for elementary particle physics.

7.2 Minkowski space

7.2.1 Definitions

We start by developing some general features of $(\mathbb{R}^4, \eta)$ called *Minkowksi space*. The tangent space to any point $\mathbb{R}^4$ can be taken to be $\mathbb{R}^4$ itself. The metric determines an indefinite inner product on the tangent space by defining for tangent vectors $v = (v^0, \ldots, v^3)$ and $w = (w^0, \ldots, w^3)$

$$v \cdot w = \sum_{\mu\nu} \eta_{\mu\nu} v^\mu w^\nu = -v^0 w^0 + v^1 w^1 + v^2 w^2 + v^3 w^3 \tag{7.6}$$

We employ the *summation convention* that repeated indices are summed over, so this is written $v \cdot w = \eta_{\mu\nu} v^\mu w^\nu$. We also write $v = (v^0, \mathbf{v})$ and $w = (w^0, \mathbf{w})$ and then

$$v \cdot w = -v^0 w^0 + \mathbf{v} \cdot \mathbf{w} \tag{7.7}$$

where the second dot product $\mathbf{v} \cdot \mathbf{w}$ is the usual scalar product in $\mathbb{R}^3$.

There are also linear functions on tangent vectors called cotangent vectors. These are also identified with $\mathbb{R}^4$ and written $\theta = (\theta_0, \ldots, \theta_3)$. The cotangent vector θ sends tangent vector v to $\theta_\mu v^\mu$. The metric enables us to identify tangent vectors and cotangent vectors by $v_\mu = \eta_{\mu\nu} v^\nu$ or $\theta^\mu = \eta^{\mu\nu} \theta_\nu$. Then for tangent vectors v, w we have $v \cdot w = \eta_{\mu\nu} v^\mu w^\nu = v^\mu w_\mu$.

Tangent vectors are defined to be *spacelike*, *lightlike*, or *timelike* according to whether $v \cdot v = -(v^0)^2 + \mathbf{v} \cdot \mathbf{v}$ is positive, zero, or negative. The timelike vectors form a double cone in $\mathbb{R}^4$. The component with $v^0 > 0$ is called *future directed* and the component with $v^0 < 0$ is called *past directed*.

We consider parametrized curves $x : [a, b] \to \mathbb{R}^4$. If the tangent vector dx/dt is always spacelike, then the curve is spacelike and we define the length of the curve to be

$$L(x) = \int_a^b \sqrt{\frac{dx}{d\tau} \cdot \frac{dx}{d\tau}} d\tau \tag{7.8}$$

If the tangent vector $dx/d\tau$ is always timelike, then the curve is timelike and we define the elapsed *proper time* to be

$$T(x) = \int_a^b \sqrt{-\frac{dx}{d\tau} \cdot \frac{dx}{d\tau}} d\tau \tag{7.9}$$

Such curves are past or future directed according to whether $dx/d\tau$ is past or future directed. Future directed timelike curves are possible trajectories of massive particles. The curve is called the *worldline* of the particle.

Finally if the tangent vector $dx/d\tau$ is always lightlike, then the curve is lightlike. Future directed lightlike curves are possible trajectories of light rays and massless particles.

7.2.2 Free particles

In the absence of external forces massive particles travel on the timelike *geodesics* of the metric. These geodesics are the forward directed timelike curves between two given points (x, y) in spacetime which maximize the proper time among all such curves. To find the geodesics note that the proper time does not depend on the parametrization of the curve. Thus it suffices to consider curves for which $-dx/d\tau \cdot dx/d\tau$ is a positive constant. Such a curve is parametrized proportional to proper time.

Theorem 7.1 *Let $x : [a,b] \to \mathbb{R}^4$ be a future directed timelike geodesic that is parametrized proportional to proper time with $x(a) = x, x(b) = y$. Then $d^2x/d\tau^2 = 0$ and hence x is a straight line. The elapsed proper time is*

$$\sqrt{-(x-y)\cdot(x-y)} \tag{7.10}$$

Proof Let $\eta : [a,b] \to \mathbb{R}^4$ be any smooth function with $\eta(a) = \eta(b) = 0$. Then for s sufficiently small

$$x_s(\tau) = x(\tau) + s\eta(\tau) \tag{7.11}$$

is also a forward timelike curve from x to y. Then we can compute the proper time $T(x_s)$. This has a maximum at $s = 0$ and so $d/ds[T(x_s)]_{s=0} = 0$. We compute

$$\frac{d}{ds}T(x_s) = \frac{1}{2}\int_a^b \left(-\frac{dx_s}{d\tau}\cdot\frac{dx_s}{d\tau}\right)^{-1/2} \frac{d}{ds}\left(-\frac{dx_s}{d\tau}\cdot\frac{dx_s}{d\tau}\right) d\tau \tag{7.12}$$

At $s = 0$ this becomes

$$0 = -\int_a^b \frac{d\eta}{d\tau}\cdot\frac{dx}{d\tau}d\tau = \int_a^b \eta\cdot\frac{d^2x}{d\tau^2}d\tau \tag{7.13}$$

Since η is arbitrary, it follows that $dx^2/d\tau^2 = 0$. Then the solution is

$$x(\tau) = x + \left(\frac{\tau - a}{b-a}\right)(y-x) \tag{7.14}$$

For this curve, $T(x)$ has the value (7.10). □

We still have some freedom in the parametrization of our geodesics. We use this to make a choice of the constant $-dx/d\tau \cdot dx/d\tau$. For particles of mass $m > 0$ a convenient choice is to set

$$-\frac{dx}{d\tau}\cdot\frac{dx}{d\tau} = m^2c^2 \tag{7.15}$$

which we now assume.

We write the dynamical equation $d^2x/d\tau^2 = 0$ as the first-order system

$$\frac{dx^\mu}{d\tau} = p^\mu \qquad \frac{dp_\mu}{d\tau} = 0 \tag{7.16}$$

where $p^\mu = \eta^{\mu\nu}p_\nu$. This is a Hamiltonian system with variables x^μ, p_μ and Hamiltonian $p\cdot p/2 = \eta^{\mu\nu}p_\mu p_\nu/2 = \eta_{\mu\nu}p^\mu p^\nu/2$, not now the energy. The constant $p = (p_0, p_1, p_2, p_3) = (p_0, \mathbf{p})$ is called the *four-momentum*. For a forward directed solution of mass m we have

$$-p\cdot p = m^2c^2 \qquad p^0 > 0 \tag{7.17}$$

Once we restrict to solutions of mass m we can reduce the number of variables by eliminating x^0 and $p_0 = -p^0$. By (7.17) we have $p^0 = \omega(\mathbf{p})$ where for $\mathbf{p} \in \mathbb{R}^3$

$$\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2c^2} \tag{7.18}$$

The x^0 equation is then $dx^0/d\tau = \omega(\mathbf{p})$, which we use to replace the parameter τ by the time coordinate x^0. We have then $d\mathbf{x}/dx^0 = (d\mathbf{x}/d\tau)(d\tau/dx^0) = \mathbf{p}/\omega(\mathbf{p})$. Finally with $x^0 = ct$ the remaining equations become

$$\frac{d\mathbf{x}}{dt} = \frac{c\,\mathbf{p}}{\omega(\mathbf{p})} \qquad \frac{d\mathbf{p}}{dt} = 0 \tag{7.19}$$

This is a Hamiltonian system with Hamiltonian $c\,\omega(\mathbf{p})$. It is interpreted as describing particle of mass m, position $\mathbf{x}$, momentum $\mathbf{p}$, and energy

$$E = c\,\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2c^2 + m^2c^4} \tag{7.20}$$

Note that the four-momentum $(p^0, \mathbf{p}) = (E/c, \mathbf{p})$ is an energy–momentum vector.

Next we make the connection with the corresponding non-relativistic concepts. For $|\mathbf{p}|/mc$ small we have the expansion

$$E = mc^2 + \frac{|\mathbf{p}|^2}{2m} + \dots \tag{7.21}$$

Thus for small momenta the energy is the same as the non-relativistic energy $|\mathbf{p}|^2/2m$, but shifted by the constant amount mc^2. The latter is the energy of the particle at rest and comes only from the mass.

Also note that according to (7.19) the velocity $\mathbf{v} = d\mathbf{x}/dt$ of a particle of momentum $\mathbf{p}$ is

$$\mathbf{v} = \frac{\mathbf{p}c^2}{E} = \frac{\mathbf{p}c}{\sqrt{|\mathbf{p}|^2 + m^2c^2}} \tag{7.22}$$

Note that $|\mathbf{v}|/c < 1$, that is velocities of massive particles are always less than the speed of light. Eliminating $\mathbf{p}$ in favor of $\mathbf{v}$ we find instead of (7.20), (7.22)

$$E = \frac{mc^2}{\sqrt{1 - |\mathbf{v}|^2/c^2}} \qquad \mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - |\mathbf{v}|^2/c^2}} \tag{7.23}$$

These also reduce to the non-relativistic definitions of energy and momentum for $|\mathbf{v}|/c$ small.

Finally consider free particles of mass $m = 0$ modeled by solutions of $d^2x/d\tau^2 = 0$ satisfying $-dx/d\tau \cdot dx/d\tau = 0$. This still can be written as the Hamiltonian system (7.16) but now restricted to solutions with $p \cdot p = 0$. Energy and momentum are related by $E = |\mathbf{p}|c$. The velocity is still $\mathbf{v} = \mathbf{p}c^2/E$, but now $|\mathbf{v}| = c$ so massless particles travel at the speed of light. There is no non-relativistic approximation for these particles. An example of a massless particle is a photon which is a particle of light. These are quantum mechanical entities, but for some purposes can be treated as classical particles as we have done here.

The value of c depends on which system of units we are using. Hereafter we choose units so that $c = 1$ and the parameter c disappears from our equations.

7.2.3 Forces

There is no fully consistent theory of interacting classical relativistic particles. Hence in this framework we do not attempt to use symmetries of a system to identify conserved quantities such as total momentum, as we did in the non-relativistic case (section 2.4). Nevertheless one can assign a total four-momentum to a system of relativistic particles by adding the individual four-momenta. It is an empirical fact that this total four-momentum is conserved in collisions of elementary particles, even when particles are created or destroyed. This feature emerges naturally in quantum field theory.

What we can do now is consider relativistic particles acted on by external forces. Suppose a massive charged particle is acted on by external electric and magnetic fields $\mathbf{E}, \mathbf{B}$. We combine these fields into a single entity, the electromagnetic field. It is a matrix of functions $F_{\mu\nu} : \mathbb{R}^4 \to \mathbb{R}$ given by

$$\{F_{\mu\nu}\} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{7.24}$$

Note that this is anti-symmetric $F_{\mu\nu} = -F_{\nu\mu}$.

The electromagnetic force on a particle with world line $x^\mu(\tau)$ and charge e is given by the *Lorentz force* $eF^\mu_{\ \nu}\, dx^\nu/d\tau$ where $F^\mu_{\ \nu} = \eta^{\mu\rho}F_{\rho\nu}$. With a parametrization satisfying $-dx/d\tau \cdot dx/d\tau = m^2$, the equation of motion of the particle is

$$\frac{d^2x^\mu}{d\tau^2} = eF^\mu_{\ \nu}\,\frac{dx^\nu}{d\tau} \tag{7.25}$$

Note that the value of $dx/d\tau \cdot dx/d\tau$ is preserved by the time evolution, since if $x(\tau)$ is a solution, then

$$\frac{d}{d\tau}\left(\frac{dx}{d\tau}\cdot\frac{dx}{d\tau}\right) = 2\frac{d^2x}{d\tau^2}\cdot\frac{dx}{d\tau} = 2eF_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0 \tag{7.26}$$

by the anti-symmetry of $F_{\mu\nu}$. The equation (7.25) is the relativistic generalization of the Lorentz equation (0.2).[1]

Now suppose that $F_{\mu\nu}$ is derived from a potential A_μ by[2]

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{7.27}$$

[1] If we reparametrize by $\sigma = m\tau$, then $dx/d\sigma \cdot dx/d\sigma = -1$ so the worldline is parametrized by proper time. In this case the equation takes the form

$$m\frac{d^2x}{d\sigma^2} = eF^\mu_{\ \nu}\,\frac{dx^\nu}{d\sigma}$$

which shows the mass dependence.

[2] This relation is naturally expressed in terms of differential forms. If we consider the two-form $F = F_{\mu\nu}dx^\mu dx^\nu$ and the one-form $A = A_\mu dx^\mu$, then (7.27) says that $F = dA$, that is F is the exterior derivative of A.

where $\partial_\mu = \partial/\partial x^\mu$. If $A = (\Phi, \mathbf{A})$, then this says

$$\mathbf{E} = \frac{\partial \mathbf{A}}{\partial x^0} - \nabla\Phi \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{7.28}$$

so we can identify Φ as the electrostatic potential and $\mathbf{A}$ as the magnetic potential. In addition a time-dependent $\mathbf{A}$ induces an electric field $\partial\mathbf{A}/\partial x^0$.

Line integrals of A over curves C are naturally defined by

$$\int_C A_\mu dx^\mu = \int_a^b A_\mu(x(\tau))\frac{dx^\mu}{d\tau}d\tau \tag{7.29}$$

whenever $x : [a, b] \to \mathbb{R}^d$ is a parametrization of C. The integral is independent of parametrization. This makes it natural to consider integrals such as

$$I(x) = \int_a^b \left[m\sqrt{-\frac{dx}{d\tau}\cdot\frac{dx}{d\tau}} - eA_\mu(x(\tau))\frac{dx^\mu}{d\tau}\right] d\tau \tag{7.30}$$

Then the next problem gives an indication of why the Lorentz equation is natural.

Problem 7.1 Let $x : [a, b] \to \mathbb{R}^4$ be forward timelike parametrized so that $-dx/d\tau \cdot dx/d\tau = m^2$. Show that if x maximizes the integral $I(x)$ among all forward timelike curves between the same endpoints, then x satisfies the Lorentz equation (7.25) with $F_{\mu\nu}$ given by (7.27).

7.2.4 Lorentz transformations

Now consider the isometries (symmetries) of the spacetime $(\mathbb{R}^4, \eta)$. These are maps $y = \kappa(x)$ which preserve the metric or equivalently preserve proper time intervals and distances. Thus they satisfy

$$(\kappa(x_1) - \kappa(x_2)) \cdot (\kappa(x_1) - \kappa(x_2)) = (x_1 - x_2) \cdot (x_1 - x_2) \tag{7.31}$$

for all $x_1, x_2 \in \mathbb{R}^4$. Translations $y = x + a$ are isometries and linear transformations $y = \Lambda x$, also written $y^\mu = \Lambda^\mu{}_\nu x^\nu$, are isometries if $\Lambda x \cdot \Lambda x = x \cdot x$. This is equivalent to

$$\Lambda^T \eta \Lambda = \eta \tag{7.32}$$

also written $\eta_{\mu\nu}\Lambda^\mu{}_{\mu'}\Lambda^\nu{}_{\nu'} = \eta_{\mu',\nu'}$. These are called Lorentz transformations. They form a group known as the *Lorentz group*.

It turns out these are all the isometries, that is the general isometry has the form

$$\{a, \Lambda\}x = \Lambda x + a \tag{7.33}$$

The group of all such transformations is called the *Poincaré group* denoted $\mathcal{P}$. The group law is

$$\{a, \Lambda\}\{a', \Lambda'\} = \{\Lambda a' + a, \Lambda\Lambda'\} \tag{7.34}$$

It is the semi-direct product of the Lorentz group and the translation group.

The Lorentz group is denoted $\mathcal{L}$ or $O(1,3)$. It inherits a topology as a subset of $\mathbb{R}^{16}$ and is in fact a Lie group. It follows from the defining relation (7.32) that elements of the group satisfy $\det \Lambda = \pm 1$ and this condition divides the group into disjoint sets denoted $\mathcal{L}_\pm$. An example of $\Lambda \in \mathcal{L}_-$ is a reflection through a hyperplane. The set $\mathcal{L}_+$ contains the identity and is a subgroup. Furthermore (7.32) implies that $(\Lambda^0{}_0)^2 = 1 + \sum_{k=1}^3 (\Lambda^k{}_0)^2$ so $\mathcal{L}_+$ is divided into disjoint sets with $\pm\Lambda^0{}_0 > 1$ and denoted respectively $\mathcal{L}_+^\uparrow$ and $\mathcal{L}_+^\downarrow$. Elements of $\mathcal{L}_+^\downarrow$ involve time reversal. The set $\mathcal{L}_+^\uparrow$ contains the identity and is a subgroup known as the *proper Lorentz group*. Correspondingly there is a *proper Poincaré group* $\mathcal{P}_+^\uparrow$.

An example of a proper Lorentz transformation is a rotation of the form

$$\Lambda_R = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \tag{7.35}$$

where R is a rotation on $\mathbb{R}^3$. Another proper Lorentz transformation is a boost along the first axis of the form

$$\Lambda_\beta = \begin{pmatrix} \cosh\beta & \sinh\beta & 0 & 0 \\ \sinh\beta & \cosh\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.36}$$

There are also boosts in any spacelike direction. It turns out that any element of $\Lambda \in \mathcal{L}_+^\uparrow$ can be written in the form

$$\Lambda = \Lambda_{R_1} \Lambda_\beta \Lambda_{R_2} \tag{7.37}$$

for some R_1, R_2, β. Since each of these special transformations can be continuously connected to the identity, it follows that any element of $\mathcal{L}_+^\uparrow$ can be continuously connected to the identity and hence the group is connected.

Physical laws are differential equations built from the metric η and will have the proper Poincaré transformations as a symmetry. This means that the proper Poincaré group acts on the space of solutions. More precisely this is true when there are no forces external to the system we are describing. For a simple example suppose $x(\tau) = p\tau + b$ is a solution of $dx^2/d\tau^2 = 0$ describing a free particle of energy–momentum p with mass m so $-p\cdot p = m^2$ and $p^0 > 0$. If $\{a, \Lambda\}$ is a proper Poincaré transformation, then the transformed world line $x'(\tau) \equiv \Lambda x(\tau) + a = (\Lambda p)\tau + \Lambda b + a$ is also a solution, now with energy–momentum $p' = \Lambda p$ and still satisfying $-p' \cdot p' = m^2$ and $(p')^0 > 0$.

The physical interpretation of Poincaré transformations can be made in either an active sense or a passive sense. In the active sense they carry physical configurations to different physical configurations. In the passive sense they describe the same

physical system in new coordinates. For boosts the new coordinates are associated with an observer in relative motion with respect to the original coordinates, hence the term "relativity."

Problem 7.2 Show that if $\Lambda^T \eta \Lambda = \eta$, then the same holds for Λ^{-1}, Λ^T.

Problem 7.3 Show that if $q^\mu = \Lambda^\mu_{\ \nu} p^\nu$, then $q_\mu = p_\nu (\Lambda^{-1})^\nu_{\ \mu}$.

Problem 7.4 Show that the boost Λ_β takes a particle at rest to a particle with velocity $\mathbf{v} = (\tanh \beta, 0, 0)$.

7.3 Classical free fields

We now begin a discussion of various fields on spacetime. Mathematically these are functions on spacetime. Physically they model some sort of local disturbance, often not directly observable. Although we introduce them in a classical context the complete physical interpretation involves quantum mechanics.

7.3.1 Scalar fields

A scalar field ϕ is a function $\phi : \mathbb{R}^4 \to \mathbb{R}$, which is a solution of the Klein–Gordon equation

$$(-\Box + m^2)\phi = 0 \tag{7.38}$$

Here $\Box = \partial \cdot \partial$ is the Laplacian for the Minkowski metric, called the *d'Alembertian*. Written out with $\partial_\mu = \partial/\partial x^\mu$ it is

$$\Box = \eta^{\mu\nu}\partial_\mu \partial_\nu = -\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2 \tag{7.39}$$

The parameter m is called the mass. This is related to the definition of mass for particles (think $p_\mu \leftrightarrow -i\partial_\mu$) but we do not make the connection precise until we quantize.

There is a basic existence and uniqueness theorem which says that given smooth functions f, g on $\mathbb{R}^3$ there is a unique smooth function $u(x) = u(x^0, \mathbf{x})$ on $\mathbb{R}^4$ such that $(-\Box + m^2)u = 0$ with intial values $u(0, \mathbf{x}) = f(\mathbf{x})$ and $(\partial_0 u)(0, \mathbf{x}) = g(\mathbf{x})$. In addition influence propagates with unit speed. We explore some variations of these facts in the following.

We also note Green's identity for the d'Alembertian, which says that for suitable functions u, v

$$\begin{aligned}&\int_{[a,b]\times\mathbb{R}^3} u(-\Box+m^2)v - v(-\Box+m^2)u \\ &= \int_{x^0=b} u\partial_0 v - v\partial_0 u \;\; - \int_{x^0=a} u\partial_0 v - v\partial_0 u\end{aligned} \tag{7.40}$$

This is proved by integration by parts.

Problem 7.5 Show that if ϕ satisfies the Klein–Gordon equation, the Poincaré transformed function $\phi_{a,\Lambda}(x) \equiv \phi(\Lambda^{-1}(x-a))$ is also a solution.

7.3.2 Charged scalar fields

The next simplest possibility is a charged scalar field ϕ of mass m, which is a function from $\mathbb{R}^4$ to $\mathbb{R}^2$. We write $\phi = (\phi_1, \phi_2)$ and require each component to solve the Klein–Gordon equation: $(-\Box + m^2)\phi_i = 0$.

By Green's identity if $\phi = (\phi_1, \phi_2)$ is a solution, then

$$Q = \int_{x^0=t} \phi_1 \partial^0 \phi_2 - \phi_2 \partial^0 \phi_1 \tag{7.41}$$

is independent of t and is called a *charge*. There is an associated charge-current density $j = (j^0, \mathbf{j}) = (j^0, j^1, j^2, j^3)$ defined by

$$j^\mu(\phi) = \phi_1 \partial^\mu \phi_2 - \phi_2 \partial^\mu \phi_1 \tag{7.42}$$

This satisfies the conservation law $\partial_\mu j^\mu = \partial_0 j^0 + \nabla \cdot \mathbf{j} = 0$ and the conservation of $Q = \int_{x^0=t} j^0$ can also be understood from

$$\frac{dQ}{dt} = \int_{x^0=t} \partial_0 j^0 = -\int_{x^0=t} \nabla \cdot \mathbf{j} = 0 \tag{7.43}$$

Multiples of $(j^0, \mathbf{j})$ will be identified with the actual electric charge density and current density when the field ϕ is coupled to the electromagnetic field.

It will also be useful to write the field as a single complex valued function $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$. In this case the current is given by

$$j^\mu = -i(\bar{\phi}\partial^\mu \phi - \phi \partial^\mu \bar{\phi}) \tag{7.44}$$

These considerations generalize to fields $\phi : \mathbb{R}^4 \to \mathbb{R}^n$ with each component satisfying $(-\Box + m^2)\phi_\alpha = 0$. In this case there are $n(n-1)/2$ conserved charges

$$Q_{\alpha\beta} = \int_{x^0=t} \phi_\alpha \partial^0 \phi_\beta - \phi_\beta \partial^0 \phi_\alpha \tag{7.45}$$

7.3.3 Dirac fields

Next consider fields which solve the Dirac equation. The equation is based on a first-order linear differential operator whose square is the d'Alembertian.

The starting point is the Clifford algebra for Minkowski space. The algebra is generated by complex matrices γ^μ with $\mu = 0, 1, 2, 3$ satisfying the anti-commutation relations[3]

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} I \tag{7.46}$$

For $a \in \mathbb{R}^4$ we define

$$\gamma \cdot a = \gamma^\mu a_\mu = \gamma^0 a_0 + \gamma^1 a_1 + \gamma^2 a_2 + \gamma^3 a_3 \tag{7.47}$$

These satisfy

$$\{\gamma \cdot a,\ \gamma \cdot b\} = 2\eta^{\mu\nu} a_\mu b_\nu = 2a \cdot b \tag{7.48}$$

In particular $(\gamma \cdot a)^2 = a \cdot a$ and thus $\gamma \cdot a$ provides a linear square root of the Minkowski inner product. Note also that $(\gamma^0)^2 = -I$ and $(\gamma^k)^2 = I$ for $k = 1, 2, 3$.

There are various possibilities for the gamma matrices. One possible choice is the 4×4 matrices

$$\gamma^0 = -i \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \qquad \gamma^k = -i \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \tag{7.49}$$

where σ_k are the Pauli matrices (4.63). These satisfy $(\gamma^0)^* = -\gamma^0$ and $(\gamma^k)^* = \gamma^k$ and we usually assume we have a representation which has this property. If $\tilde{\gamma}^\mu$ is another choice of 4×4 gamma matrices, then there is a nonsingular matrix M such that[4]

$$\tilde{\gamma}^\mu = M \gamma^\mu M^{-1} \tag{7.50}$$

The operator $\gamma \cdot \partial = \gamma^\mu \partial_\mu$ has the desired property

$$(\gamma \cdot \partial)^2 = \partial \cdot \partial = \Box \tag{7.51}$$

The *Dirac equation* is now defined by

$$(\gamma \cdot \partial + m)\psi = 0 \tag{7.52}$$

for some function $\psi : \mathbb{R}^4 \to \mathbb{C}^4$ called a *spinor* field. Applying the operator $(-\gamma \cdot \partial + m)$ we see that any solution of the Dirac equation also satisfies the Klein–Gordon equation $(-\Box + m^2)\psi = 0$. Hence m is again a mass.

[3] The anticommutator is $\{A, B\} = [A, B]_+ = AB + BA$.

[4] See for example Miller (1972: 363).

Next let $\beta = i\gamma^0$ and define an indefinite inner product on $\mathbb{C}^4$ by $u, v \to u^\dagger \beta v$. Here if u is a column vector, then $u^\dagger$ is the conjugate transpose row vector so

$$u^\dagger \beta v = \sum_{ab} \bar{u}_a \beta_{ab} v_b \tag{7.53}$$

The γ matrices are skew-adjoint with respect to this inner product

$$(\gamma^\mu u)^\dagger \beta v = -u^\dagger \beta (\gamma^\mu v) \tag{7.54}$$

Green's identity for the Dirac operator says

$$\begin{aligned} &\int_{[a,b]\times\mathbb{R}^3} \chi^\dagger \beta(\gamma\cdot\partial + m)\psi - ((\gamma\cdot\partial + m)\chi)^\dagger \beta\psi \\ &= \int_{x^0=b} \chi^\dagger \beta(\gamma^0\psi) - \int_{x^0=a} \chi^\dagger \beta(\gamma^0\psi) \end{aligned} \tag{7.55}$$

Then if ψ is a solution of the Dirac equation, the quantity

$$Q = \int_{x^0=t} \psi^\dagger \beta[i\gamma^0]\psi = \int_{x^0=t} |\psi|^2 \tag{7.56}$$

is positive definite and independent of t. A multiple will be interpreted as the electric charge. The associated conserved current

$$j^\mu = \psi^\dagger \beta[i\gamma^\mu]\psi \tag{7.57}$$

satisfies $Q = \int_{x^0=t} j^0$ and $\partial_\mu j^\mu = 0$.

Problem 7.6

1. Show that $Tr(\gamma^\mu\gamma^\nu) = \eta^{\mu\nu} Tr(I)$.
2. Show that the γ^μ are linearly independent.

7.3.4 The electromagnetic field

We have already noted that the electromagnetic field can be described by a matrix of functions $F_{\mu\nu}$. If the total charge density ρ and the total current density $\mathbf{j}$ are combined into a spacetime vector field $j = (j^0, j^1, j^2, j^3) = (\rho, \mathbf{j})$, then Maxwell's equations can be written in the form[5]

$$\begin{aligned} \partial_\sigma F_{\mu\nu} + \partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} &= 0 \\ \partial_\nu F^{\mu\nu} &= j^\mu \end{aligned} \tag{7.58}$$

In this form the necessity of the charge conservation law $\partial_\mu j^\mu = 0$ is transparent.

[5] In terms of the two-form $F = F_{\mu\nu} dx^\mu dx^\nu$, the one-form $j = j_\mu dx^\mu$, the exterior derivative d, and its adjoint δ these say $dF = 0$ and $\delta F = j$. ■

In a simply connected region the first equation says that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{7.59}$$

for some functions A_μ.[6] We identify $A = (A_0, \mathbf{A})$ with the electric potential and the magnetic potential as in section 7.2.3. The potential can be assumed to satisfy $\partial_\mu A^\mu = 0$ (see problem 7.7). Then Maxwell's equations are replaced by the pair of equations

$$\Box A_\mu = -j_\mu \qquad \partial^\mu A_\mu = 0 \tag{7.60}$$

It is just the wave equation with a source and a constraint. This is the equation we will eventually quantize (for $j = 0$).

Problem 7.7 In $\mathbb{R}^4$ let $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

1. Show that if $A'_\mu = A_\mu + \partial_\mu \chi$ for any smooth function χ, then $F_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu$.
2. Show that one can choose χ so that $\partial^\mu A'_\mu = 0$.

Problem 7.8 Show that if A_μ solves (7.60) with $j = 0$, then so does $A'_\mu(x) = (\Lambda^{-1})^\nu{}_\mu A_\nu(\Lambda^{-1}(x - a))$.

7.4 Interacting classical fields

7.4.1 The gauge principle

The equations (7.58) or (7.60) show how the charge-current density acts as the source of the electromagnetic field. But the electromagnetic field also affects charges and in particular charged fields. This occurs in a geometrically natural way as we now explain.

Let us start with the charged scalar field whose configuration is a function $u : \mathbb{R}^4 \to \mathbb{R}^n$. Just as the vector space $\mathbb{R}^4$ is not the best model for spacetime, yielding that honor to a manifold, so the vector space $\mathbb{R}^n$ is not the best model for charge space. We continue to treat it as an inner product space, but now do not single out any special oriented orthonormal basis. Instead we consider all possible oriented orthonormal bases denoted $\{e_\alpha\}$, and instead of a fixed vector in $\mathbb{R}^n$ we consider its expression $v = \sum_\alpha v_\alpha e_\alpha$ in each of these bases. The bases are related to each other by an element of the rotation group $SO(n)$ in $\mathbb{R}^n$ and hence so are the components

[6] Every closed form $dF = 0$ is exact $F = dA$.

relative to each choice of basis. Thus what we want is an assignment to each oriented orthonormal basis $\{e_\alpha\}$ a vector $v \in \mathbb{R}^n$ such that if

$$e_\beta = \sum_\alpha R_{\alpha\beta} e'_\alpha \tag{7.61}$$

for some $R \in SO(n)$, then

$$v'_\alpha = \sum_\beta R_{\alpha\beta} v_\beta \tag{7.62}$$

For then we have

$$\sum_\beta v_\beta e_\beta = \sum_\alpha v'_\alpha e'_\alpha \tag{7.63}$$

Next we want to allow the possibility of choosing a different basis at each point in spacetime. Thus we consider functions from $x \in \mathbb{R}^4$ to orthonormal bases $\{e_\alpha(x)\}$. Two such functions are related by a function $R : \mathbb{R}^4 \to SO(n)$ such that

$$e_\beta(x) = \sum_\alpha R_{\alpha\beta}(x) e'_\alpha(x) \tag{7.64}$$

The function $R(x)$ is known as a (local) *gauge transformation* and is assumed to be smooth. Now if our charged scalar field is expressed as $\sum_\beta u_\beta(x) e_\beta(x)$ and also as $\sum_\alpha u'_\alpha(x) e'_\alpha(x)$, then

$$u'_\alpha(x) = \sum_\beta R_{\alpha\beta}(x) u_\beta(x) \tag{7.65}$$

This is the structure of a trivial $SO(n)$ vector bundle. If we allow $\mathbb{R}^4$ to be a manifold and allow different choices of bases in $\mathbb{R}^n$ in different open subsets of the manifold, we would have the general definition of an $SO(n)$ vector bundle. In the terminology of vector bundles the functions $u(x)$ are sections of the vector bundle in a particular trivialization and the gauge transformations $R(x)$ are the transition functions for a change of trivialization.

Let us specialize now to the case $n = 2$ in which case $\phi : \mathbb{R}^4 \to \mathbb{R}^2$ is the charged scalar field and the gauge group is $SO(2)$. We regard $\mathbb{R}^2$ as the complex numbers $\mathbb{C}$. Then $SO(2)$ is identified with the group $U(1)$, the complex numbers of modulus one, under the identification

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \leftrightarrow e^{i\theta} \tag{7.66}$$

Thus we consider families of functions $\phi : \mathbb{R}^4 \to \mathbb{C}$ assigned to oriented orthonormal bases in $\mathbb{C}$ (that is choices of the real axis) such that if two bases are related by the gauge transformation $e^{i\theta(x)}$, then the functions are related by

$$\phi'(x) = e^{i\theta(x)} \phi(x) \tag{7.67}$$

A family of functions with this transformation law is called a section of a complex line bundle.

Next we want to write a differential equation for a section of a complex line bundle. The equation will be expressed in terms of a particular function representing the section, but should be natural in the sense that it should not matter which function we take.

To build such an equation we need the notion of a connection. A connection is given by a family of functions A_μ each associated with choice of basis. Under a change of basis given by $e^{iq\lambda(x)}$ the connection changes (by definition) by

$$A'_\mu(x) = A_\mu(x) + \partial_\mu \lambda(x) \tag{7.68}$$

Here we have inserted a constant q in the transformation $e^{iq\lambda(x)}$ to allow a different weighting for fields and connections.

The connection determines a covariant derivative $\partial_\mu - iqA_\mu$. If $\phi' = e^{iq\lambda}\phi$, then

$$\begin{aligned}(\partial_\mu - iqA'_\mu)\phi' &= (\partial_\mu - iqA_\mu - iq\partial_\mu\lambda)e^{iq\lambda}\phi \\ &= e^{iq\lambda}(\partial_\mu - iqA_\mu)\phi\end{aligned} \tag{7.69}$$

Thus the functions $(\partial_\mu - iqA_\mu)\phi$ give a new section of the complex line bundle.

Now we can define a differential equation for the charged field $\phi : \mathbb{R}^4 \to \mathbb{C}$ by treating it as a section of a complex line bundle. Take the Klein–Gordon equation and replace the derivatives by covariant derivatives. Then we have

$$(-(\partial - iqA)\cdot(\partial - iqA) + m^2)\phi = 0 \tag{7.70}$$

The interpretation is that the connection A_μ is the electromagnetic potential, that the equation is describing the time evolution of the charged field ϕ in the presence of A, and that q is the charge measuring the strength of the coupling.

Recall that potentials A, A' related by a gauge transformation (7.68) have the same field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Thus it is possible to regard the different versions of (7.70) as describing the same physical situation. The field strength $F_{\mu\nu}$ also has a natural interpretation in the vector bundle language. Namely it is the curvature of the connection defined as the commutator of the covariant derivative. We have

$$-iqF_{\mu\nu} = [(\partial_\mu - iqA_\mu), (\partial_\nu - iqA_\nu)] \tag{7.71}$$

This all generalizes to $SO(n)$ and indeed any Lie group G. Suppose G is represented by $n \times n$ real or complex matrices. A section of a G-bundle is a family of functions $\phi : \mathbb{R}^4 \to \mathbb{R}^n$ or $\mathbb{C}^n$ connected by gauge transformations $g : \mathbb{R}^4 \to G$ by

$$\phi'(x) = g(x)\phi(x) \tag{7.72}$$

Let $\mathcal{G}$ be Lie algebra of G; for example if $G = SO(n)$, then $\mathcal{G}$ is the skew-symmetric $n \times n$ matrices. A connection or *gauge potential* is given by a family of functions $A_\mu : \mathbb{R}^4 \to \mathcal{G}$ which are related by[7]

$$A'_\mu(x) = g(x)A_\mu(x)g^{-1}(x) - (\partial_\mu g)(x)g^{-1}(x) \tag{7.73}$$

The covariant derivative $\partial_\mu + A_\mu$ again maps sections to sections in the sense that

$$(\partial_\mu + A'_\mu)\phi' = g(\partial_\mu + A_\mu)\phi \tag{7.74}$$

One can form a dynamical equation by

$$(-(\partial + A) \cdot (\partial + A) + m^2)\phi = 0 \tag{7.75}$$

This describes a field with a generalized charge. The curvature is now

$$\begin{aligned} F_{\mu\nu} &= [(\partial_\mu + A_\mu), (\partial_\nu + A_\nu)] \\ &\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \end{aligned} \tag{7.76}$$

and it is interpreted as the field strength of the potential. Equivalent connections A, A' do not now give the same field strength. We have instead $F'_{\mu\nu} = gF_{\mu\nu}g^{-1}$.

7.4.2 Systems

Now we can put together some of our equations to form complete systems in which each field has a dynamics and influences the other fields.

To start, consider a system consisting of a charged Dirac field $\psi : \mathbb{R}^4 \to \mathbb{C}^4$ interacting with an electromagnetic field $A : \mathbb{R}^4 \to \mathbb{R}$. In this case the charge-current density $j^\mu = q\psi^\dagger\beta[i\gamma^\mu]\psi$ is the source for the electromagnetic field in Maxwell's equations (7.58). The influence of A on ψ is given by replacing the ordinary derivative by the covariant derivative in the Dirac equation. Thus we have

$$\begin{aligned} (\gamma \cdot (\partial - iqA) + m)\psi &= 0 \\ \partial_\nu F^{\mu\nu} &= q\psi^\dagger\beta[i\gamma^\mu]\psi \end{aligned} \tag{7.77}$$

Note that the same charge q is used in both equations. Because of this the system of equations can be derived from a simple variational principle. Define an action by

$$S(\psi, A) = \int_{t_0}^{t_1} \int_{\mathbb{R}^3} \left(\psi^\dagger\beta(\gamma \cdot (\partial - iqA) + m)\psi + \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \right) dx \tag{7.78}$$

The least action principle says that dynamical fields must minimize the action and these turn out to be exactly the fields satisfying (7.77).

[7] A' is again in the Lie algebra $\mathcal{G}$. To see that $v^\mu(\partial_\mu g)(x)g^{-1}(x)$ is in $\mathcal{G}$ for any $x, v \in \mathbb{R}^4$ let $x(t)$ be a curve in $\mathbb{R}^4$ with $x(0) = x, x'(0) = v$. Then $\gamma(t) = g(x(t))g^{-1}(x)$ is a curve in G with $\gamma(0) = I, \gamma'(0) = v^\mu(\partial_\mu g)(x)g^{-1}(x)$ and hence the latter is in $\mathcal{G}$.

Next consider a charged scalar field $\phi : \mathbb{R}^4 \to \mathbb{C}$ interacting with an electromagnetic potential A. The influence of A on the ϕ field is given by taking covariant derivatives in the Klein–Gordon equation as in (7.75). We complete the system by taking $j^\mu = -iq(\bar{\phi}\partial^\mu\phi - \phi\partial^\mu\bar{\phi})$ as the source for the electromagnetic field. Thus we have the system:

$$\begin{aligned}(-(\partial - iqA)\cdot(\partial - iqA) + m^2)\phi &= 0\\ \partial_\nu F^{\mu\nu} &= -iq(\bar{\phi}\partial^\mu\phi - \phi\partial^\mu\bar{\phi})\end{aligned} \tag{7.79}$$

Usually one also adds a term $-2q^2A^\mu|\phi|^2$ to the right side of the second equation. Then the equations can be derived from a least action principle with action

$$S(\phi, A) = \int_{t_0}^{t_1}\int_{\mathbb{R}^3}\left(\frac{1}{2}\overline{(\partial - iqA)\phi}\cdot(\partial - iqA)\phi + m^2|\phi|^2 + \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right)dx \tag{7.80}$$

The equations were constructed in accordance with the gauge principle for gauge group $SO(2) = U(1)$. For a general gauge group G we would replace ϕ or ψ by a vector valued function and A by a Lie algebra valued function. The actions (7.80) or (7.78) still hold with $F_{\mu\nu}$ now given by (7.76) and $\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ replaced by the gauge invariant $\frac{1}{4}\mathrm{tr}(F^{\mu\nu}F_{\mu\nu})$. Variation of the action leads to a system of equations similar to (7.77) or (7.79), but even more nonlinear. These are known as *nonabelian gauge theories*. With the gauge group $SU(3)\times SU(2)\times U(1)$ a quantized version provides a model for the strong, weak, and electromagnetic interactions of elementary particles, known as the *standard model*.

These are difficult systems of nonlinear equations. A simpler case is a single scalar field which interacts with itself according to the equation

$$(-\Box + m^2)\phi + 4\lambda\phi^3 = 0 \tag{7.81}$$

with $\lambda > 0$. This can be derived from the action

$$S(\phi) = \int_{t_0}^{t_1}\int_{\mathbb{R}^3}\left(\frac{1}{2}(\partial\phi\cdot\partial\phi + m^2\phi^2) + \lambda\phi^4\right) \tag{7.82}$$

Problem 7.9 Consider smooth functions $\phi(t, \mathbf{x})$ such that for $t_0 \le t \le t_1$ the function $\phi(t, \cdot)$ has compact support and so $\phi(t_0, \mathbf{x}) = f_0(\mathbf{x}), \phi(t_1, \mathbf{x}) = f_1(\mathbf{x})$ Show that if $\phi(t, \mathbf{x})$ minimizes the action (7.82) among all such functions, then it satisfies the equation (7.81).

7.5 Fundamental solutions

In this last section we depart from general considerations to treat a specific problem. We define and study certain fundamental solutions of the Klein–Gordon equation.

For any $x \in \mathbb{R}^4$ define

$$J^{\pm}(x) = \{y \in \mathbb{R}^4 : (x-y)\cdot(x-y) \leq 0,\ \pm(y^0 - x^0) \geq 0\} \tag{7.83}$$

This is the causal future or past of the point x, that is the point which can be reached by a future or past directed timelike curve. For any subset $A \subset \mathbb{R}^4$ we define

$$J^{\pm}A = \cup_{x\in A} J^{\pm}x \tag{7.84}$$

Theorem 7.2 *Then there exist linear operators* $E^{\pm} : \mathcal{C}_0^{\infty}(\mathbb{R}^4) \to \mathcal{C}^{\infty}(\mathbb{R}^4)$ *such that*

$$\begin{aligned} (-\Box + m^2)(E^{\pm}f) &= f \\ \text{supp}\,(E^{\pm}f) &\subset J^{\pm}(\text{supp}\, f) \end{aligned} \tag{7.85}$$

Proof We first define a distribution $E_0^{\pm}$ by

$$E_0^{\pm}(x) = (2\pi)^{-4} \int_{\Gamma_{\pm}\times\mathbb{R}^3} \frac{e^{ip\cdot x}}{p\cdot p + m^2}\, dp \tag{7.86}$$

where $\Gamma_{\pm}$ is the contour $\mathbb{R} \pm i\alpha$ with $\alpha > 0$, and we use the Lorentz inner product in the exponential. The Fourier transform is in the sense of distributions so for $f \in \mathcal{C}_0^{\infty}(\mathbb{R}^4)$

$$< E_0^{\pm}, f >= (2\pi)^{-2} \int_{\Gamma_{\pm}\times\mathbb{R}^3} \frac{\hat{f}(-p)}{p\cdot p + m^2} dp \tag{7.87}$$

where the Fourier transform is with the Lorentz inner product

$$\hat{f}(p) = (2\pi)^{-2} \int_{\mathbb{R}^4} e^{-ip\cdot x} f(x) dx \tag{7.88}$$

Note that $p\cdot p + m^2 = -(p^0)^2 + |\mathbf{p}|^2 + m^2$ has zeros at $p^0 = \pm\omega(\mathbf{p})$, which we avoid by the choice of the contour $\Gamma_{\pm}$. The function $\hat{f}$ is an entire function which is rapidly decreasing in real directions.. Thus $< E_0^{\pm}, f >$ is well-defined and independent of the choice of α.

The operator is defined by the convolution $E^{\pm}f = E_0^{\pm} * f$, which means $(E^{\pm}f)(x) =< E_0^{\pm}, f(x-\cdot) >$. Replacing f by $f(x-\cdot)$ means replacing $\hat{f}(p)$ by $e^{-ip\cdot x}\hat{f}(-p)$ and so we obtain

$$(E^{\pm}f)(x) = (2\pi)^{-2} \int_{\Gamma_{\pm}\times\mathbb{R}^3} \frac{e^{ip\cdot x}\hat{f}(p)}{p\cdot p + m^2} dp \tag{7.89}$$

It satisfies $(-\Box + m^2)(E^{\pm}f)(x) = f(x)$ since after taking the derivatives we get $(2\pi)^{-2}\int_{\Gamma_{\pm}\times\mathbb{R}^3} e^{ip\cdot x}\hat{f}(p)dp$ in which we can take $\alpha = 0$ to identify $f(x)$.

The statement about supports is equivalent to supp $E_0^{\pm} \subset J^{\pm}(0)$. We consider E_0^{+} and first show that supp E_0^{+} is contained in the set $x^0 \geq 0$. Thus we want to show that if supp$f \subset \{x^0 < 0\}$, then $< E_0^{+}, f >= 0$. If $x^0 < 0$, then $|\exp(-ip\cdot x)| = \exp(-\text{Im}\ p^0 x^0)$ is bounded in Im $p^0 < 0$ and so $\hat{f}(p)$ is bounded in Im $p^0 < 0$. Then $\hat{f}(-p)$ is bounded in Im $p^0 > 0$, and in the formula (7.87) we can complete the contour Γ_+ in the upper half plane and get zero.

Finally we use the fact that E_0^+ is Lorentz invariant, $E_0^+(\Lambda x) = E_0^+(x)$. It follows that if $E_0^+(x)$ vanishes on a neighborhood of a point x, then it vanishes on a neighborhood of any point Λx in the orbit of x. Hence we can enlarge the region where it vanishes from $\{x^0 < 0\}$ to $\Lambda\{x^0 < 0\}$. The latter is the complement of $J^+(0)$, hence the support of E_0^+ is in $J^+(0)$. The argument for $E_0^-(x)$ is similar. □

From these fundamental solutions we construct the *propagator*[8]

$$E = E^+ - E^- \tag{7.90}$$

If $f \in \mathcal{C}_0^\infty(\mathbb{R}^4)$, then $u = Ef$ is a $\mathcal{C}^\infty$ solution of the Klein–Gordon equation with the property that for any t the function $u(t, \cdot)$ has compact support. Such a solution will be called a *regular* solution.

It will be useful to have an alternate expression for $(Ef)(x)$. It is given by the expression (7.89) but now integrated over $(\Gamma^+ - \Gamma^-) \times \mathbb{R}^3$. The p^0 contour $(\Gamma^+ - \Gamma^-)$ can be shrunk down to circles around $p^0 = \pm\omega(\mathbf{p})$ and evaluated. After a change of variables we have

$$\begin{aligned}(Ef)(x) = &\frac{i}{2\pi}\int_{\mathbb{R}^3} e^{i(-\omega(\mathbf{p})x^0 + \mathbf{p}\cdot\mathbf{x})}\hat{f}(\omega(\mathbf{p}), \mathbf{p})\frac{d\mathbf{p}}{2\omega(\mathbf{p})} \\ &- \frac{i}{2\pi}\int_{\mathbb{R}^3} e^{i(\omega(\mathbf{p})x^0 - \mathbf{p}\cdot\mathbf{x})}\hat{f}(-\omega(\mathbf{p}), -\mathbf{p})\frac{d\mathbf{p}}{2\omega(\mathbf{p})}\end{aligned} \tag{7.91}$$

Problem 7.10 Show that every regular solution of the Klein–Gordon equation has the form $u = Ef$ for some $f \in \mathcal{C}_0^\infty(\mathbb{R}^4)$

Notes on chapter 7: Special relativity is best understood as general relativity with a special metric. References are Misner *et al.* (1973) and Sachs and Wu (1977).

For gauge theories see Drechsler and Mayer (1977) or Frankel (2004).

It is also common to use the metric $\eta' = -\eta$ instead of η. Then the Klein–Gordon equation $(-\Box_\eta + m^2)\phi = 0$ becomes $(\Box_{\eta'} + m^2)\phi = 0$. The Dirac matrices γ^μ are replaced by $(\gamma')^\mu = i\gamma^\mu$. The Dirac equation $(\gamma^\mu\partial_\mu + m)u = 0$ becomes $(-i(\gamma')^\mu\partial_\mu + m)u = 0$.

[8] Not to be confused with the Feynman propagator.

8 Scalar particles and fields

In this chapter we develop a quantum mechanical description of both scalar particles and scalar fields, in each case without interaction. In fact there is a deep connection between particles and fields. Mathematically it is reflected in the fact that field operators naturally act on a multiparticle Hilbert space. Physically it is the fact that fields and particles are complementary manifestations of the same underlying reality. The prime example is electromagnetism which exists as electric and magnetic fields or as massless spin 1 particles called photons. This is taken up in the next chapter, after we treat the simpler case of scalars.

As we develop this and other relativistic field theories we will encounter states with arbitrarily negative energies. If present, these would lead to serious instabilities when our system is coupled to other systems. Our system could serve as an infinite energy source by dropping to lower and lower levels. This is completely unphysical so one of our tasks is to find a consistent way to discard or reinterpret the negative energy states.

8.1 Scalar particles

8.1.1 Canonical quantization

We consider a single free relativistic particle of mass m in the spacetime $(\mathbb{R}^4, \eta)$ and start with a canonical quantization procedure. Recall from section 7.2.2 that in reduced form the phase space is $\mathbb{R}^3 \times \mathbb{R}^3$ and the Hamiltonian for a particle of mass m is $H(\mathbf{x}, \mathbf{p}) = \omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}$. Correspondingly we take the Hilbert space $L^2(\mathbb{R}^3, d\mathbf{x})$ and the representation of the CCR

$$\hat{x}^i = [x^i] \qquad \hat{p}_i = -i\partial/\partial x^i \tag{8.1}$$

just as in (3.11). The quantum Hamiltonian is then

$$H = \sqrt{|\hat{p}|^2 + m^2} = \sqrt{-\Delta + m^2} \tag{8.2}$$

This is defined in terms of the Fourier transform by $H = \mathcal{F}^{-1}[\omega(\mathbf{p})]\mathcal{F}$ and time evolution is

$$e^{-iHt} = \mathcal{F}^{-1}[e^{-i\omega(\mathbf{p})t}]\mathcal{F} \tag{8.3}$$

Alternatively we can work directly in momentum space. Then the Hilbert space is $L^2(\mathbb{R}^3, d\mathbf{p})$, the position and momentum operators are

$$\hat{x}^i = i\partial/\partial p_i \qquad \hat{p}_i = [p_i] \tag{8.4}$$

The Hamiltonian is

$$H = \sqrt{|\hat{p}|^2 + m^2} = [\omega(\mathbf{p})] \tag{8.5}$$

and time evolution is $e^{-iHt} = [e^{-i\omega(\mathbf{p})t}]$. The two constructions are of course unitarily equivalent via the Fourier transform.

Problem 8.1 (Non-relativistic limit) If we reinstate the speed of light c as a parameter, then $\omega(\mathbf{p})$ is replaced by $\omega_c(\mathbf{p}) = \sqrt{|\mathbf{p}|^2c^2 + m^2c^4}$. Show that for $\psi \in L^2(\mathbb{R}^3, d\mathbf{p})$

$$L^2 - \lim_{c\to\infty} e^{imc^2t}e^{-i\omega_c(\mathbf{p})t}\psi(\mathbf{p}) = e^{-i|\mathbf{p}|^2t/2m}\psi(\mathbf{p}) \tag{8.6}$$

Thus when adjusted by the phase factor e^{imc^2t} the relativistic dynamics converges to the non-relativistic dynamics.

8.1.2 Quantization from Klein-Gordon

Another way to approach quantization is to start with the Klein–Gordon equation $(-\Box + m^2)\phi = 0$ and try to interpret solutions as wave functions. To find solutions take the partial Fourier transform

$$\tilde{\phi}(t, \mathbf{p}) = (2\pi)^{-3/2} \int e^{-i\mathbf{p}\cdot\mathbf{x}}\phi(t, \mathbf{x})d\mathbf{x} \tag{8.7}$$

This satisfies

$$\left(\frac{\partial^2}{\partial t^2} + |\mathbf{p}|^2 + m^2\right)\tilde{\phi} = 0 \tag{8.8}$$

and has solutions

$$\tilde{\phi}(t, \mathbf{p}) = e^{-i\omega(\mathbf{p})t}\phi_+(\mathbf{p}) + e^{i\omega(\mathbf{p})t}\phi_-(\mathbf{p}) \tag{8.9}$$

for any $\phi_\pm(\mathbf{p})$. Hence a general complex valued solution of the Klein–Gordon equation is

$$\phi(t, \mathbf{x}) = (2\pi)^{-3/2} \int e^{i\mathbf{p}\cdot\mathbf{x}} \left(e^{-i\omega(\mathbf{p})t}\phi_+(\mathbf{p}) + e^{i\omega(\mathbf{p})t}\phi_-(\mathbf{p})\right) d\mathbf{p} \tag{8.10}$$

with $\phi_\pm(\mathbf{p})$ determined by a pair of initial conditions. Real solutions have $\overline{\phi_-(\mathbf{p})} = \phi_+(-\mathbf{p})$.

For the quantum interpretation we specialize to positive energy solutions which have $\phi_-(\mathbf{p}) = 0$. Then the solution is determined by a single complex valued initial condition. The solution (8.9) with initial condition $\psi(\mathbf{p}) \in L^2(\mathbb{R}^3, d\mathbf{p})$ is $\psi(t,\mathbf{p}) = e^{-i\omega(\mathbf{p})t}\psi(\mathbf{p})$ which agrees with our momentum space quantization. The solution (8.10) with initial condition $\psi(\mathbf{x}) \in L^2(\mathbb{R}^3, d\mathbf{x})$ is given by

$$\psi(t,\mathbf{x}) = (2\pi)^{-3/2} \int e^{i\mathbf{p}\cdot\mathbf{x}} e^{-i\omega(\mathbf{p})t} \tilde{\psi}(\mathbf{p}) d\mathbf{p} \tag{8.11}$$

which agrees with our configuration space quantization (8.3).

Now consider the effect of spacetime translations by $a = (a^0, \mathbf{a}) \in \mathbb{R}^4$ and rotations $R \in SO(3)$. The transformation $\{a, R\}$ on $\mathbb{R}^4$ forms a subgroup of the proper Poincaré group; we exclude boosts for the time being. Define a representation of this subgroup on $L^2(\mathbb{R}^3, d\mathbf{p})$ by

$$(u_0(a,R)\psi)(\mathbf{p}) = e^{i(\omega(\mathbf{p})a^0 - \mathbf{p}\cdot\mathbf{a})} \psi(R^{-1}\mathbf{p}) \tag{8.12}$$

These operators are unitary since the Lebesgue measure is invariant under rotations and one checks that $u_0(a,R)u_0(a',R') = u_0(a + Ra', RR')$.

On $L^2(\mathbb{R}^3, d\mathbf{x})$ we define a representation

$$\hat{u}_0(a,R) = \mathcal{F}^{-1}\, u_0(a,R)\, \mathcal{F} \tag{8.13}$$

Then $\hat{u}_0(a,R)$ implements the action of these transformations on our wave functions. If $\psi(t,\mathbf{x})$ is a complete wave function given by (8.11), then

$$\Big(\hat{u}_0(a,R)\psi(t,\cdot)\Big)(\mathbf{x}) = \psi(t - a^0, R^{-1}(\mathbf{x} - \mathbf{a})) \tag{8.14}$$

Note that the representation of the rotation group corresponds to spin zero.

8.1.3 Covariant quantization

There is a third way to arrive at the quantization of a single relativistic particle. This begins with the idea that associated with any isolated system there should be a unitary representation of the proper Poincaré group describing the effect of Poincaré transformations. Since time translation is included the representation contains the complete dynamics of the system. Elementary particles are supposed to correspond to irreducible representations. These are labeled by two parameters interpreted as mass and spin. Here we want to find the irreducible representation for particle of mass m and spin zero.

Working in momentum space $\mathbb{R}^4$ note that the *mass shell*

$$V_m^+ = \{p = (p^0, p^1, p^2, p^3) \in \mathbb{R}^4 : p \cdot p = -m^2, p^0 > 0\} \tag{8.15}$$

is invariant under the proper Lorentz group $\mathcal{L}_+^\uparrow$. (In fact V_m^+ is the orbit of the point $(m,0,0,0)$.) Then we can define a representation of the proper Poincaré group $\mathcal{P}_+^\uparrow$ on functions on V_m^+ by

$$(u(a,\Lambda)\psi)(p) = e^{-ip\cdot a}\psi(\Lambda^{-1}p) \tag{8.16}$$

where $p\cdot a = -p^0a^0 + \mathbf{p}\cdot\mathbf{a}$. One checks that

$$u(a,\Lambda)u(a',\Lambda') = u(a+\Lambda a', \Lambda\Lambda') \tag{8.17}$$

There is an essentially unique measure on V_m^+ which is invariant under the Lorentz group which we denote μ_m. Then $u(a,\Lambda)$ is a unitary representation of the proper Poincaré group on $L^2(V_m^+,\mu_m)$ and it turns out to be irreducible.

We can be more specific about the measure μ_m. First note that the map

$$\phi(\mathbf{p}) = (\omega(\mathbf{p}),\mathbf{p}) \tag{8.18}$$

from $\mathbb{R}^3$ to V_m^+ provides a global set of coordinates for V_m^+. For measurable $B\subset\mathbb{R}^3$ we define

$$\mu_m(\phi(B)) = \int_B \frac{d\mathbf{p}}{2\omega(\mathbf{p})} \tag{8.19}$$

Then $\psi \to \psi\circ\phi$ is a unitary map from $L^2(V_m^+,\mu_m)$ to $L^2(\mathbb{R}^3, d\mathbf{p}/2\omega(\mathbf{p}))$ and the representation (8.16) becomes the representation on $L^2(\mathbb{R}^3, d\mathbf{p}/2\omega(\mathbf{p}))$ given by

$$(u(a,\Lambda)\psi)(\mathbf{p}) = e^{-i\phi(\mathbf{p})\cdot a}\psi(\phi^{-1}\Lambda^{-1}\phi(\mathbf{p})) \tag{8.20}$$

Problem 8.2 Check that the measure $d\mathbf{p}/2\omega(\mathbf{p})$ on $\mathbb{R}^3$ is invariant under Lorentz transformations $\mathbf{p}\to\phi^{-1}\Lambda^{-1}\phi(\mathbf{p})$. (Hint: Use the decomposition (7.37).)

Problem 8.3 Check that $u(a,\Lambda)$ as defined in (8.16) is a unitary representation of the proper Poincaré group on $L^2(V_m^+,\mu_m)$.

8.1.4 Comparison

We now compare the different approaches. In the first two the Hilbert space can be taken as $\mathcal{H}_0 = L^2(\mathbb{R}^3, d\mathbf{p})$ the same as for non-relativistic problems. In this case there are position operators $\hat{x}^i = i\partial/\partial p_i$ and a representation $u_0(a,R)$ of spacetime translations and rotations. In the third approach the Hilbert space is $\mathcal{H} = L^2(\mathbb{R}^3, d\mathbf{p}/2\omega(\mathbf{p}))$ and there is a representation of the proper Poincaré group $u(a,\Lambda)$.

These two possibilities are equivalent via the unitary operator $V:\mathcal{H}_0\to\mathcal{H}$ defined by

$$(V\psi)(\mathbf{p}) = \sqrt{2\omega(\mathbf{p})}\psi(\mathbf{p}) \tag{8.21}$$

This connects the representations since

$$V^{-1}u(a,R)\,V = u_0(a,R) \tag{8.22}$$

If we want a representation of the full Poincaré group on $\mathcal{H}_0$, we can extend the representation on the subgroup by defining

$$u_0(a,\Lambda) \equiv V^{-1}u(a,\Lambda)\,V \tag{8.23}$$

If we want a position operator on $\mathcal{H}$, we can take the Newton–Wigner operator

$$X^i = V\,\hat{x}^i\,V^{-1} \tag{8.24}$$

Note that if $\psi \in \mathcal{H}_0$ is strictly localized (that is if $\mathcal{F}^{-1}\psi$ has compact support), it is not true in general that $u_0(0,\Lambda)\psi$ is strictly localized. Strict localization of particles is not a Lorentz invariant concept. We will do better with fields.

Problem 8.4 Find explicit expressions for X^i and $u_0(a,\Lambda)$.

8.1.5 Many particles

Now suppose we want to describe many free bosons of mass m and spin zero. The Hilbert space would be the Fock space $\mathcal{F}^+(\mathcal{H}_0)$ or $\mathcal{F}^+(\mathcal{H})$. The time evolution would be given by the one-parameter unitary group

$$e^{-iHt} = \Gamma([e^{-i\omega(\mathbf{p})t}]) \tag{8.25}$$

where Γ is defined in (5.48). The Hamiltonian is the generator

$$H = d\Gamma([\omega(\mathbf{p})]) = \int \omega(\mathbf{p})a^*(\mathbf{p})a(\mathbf{p})d\mathbf{p} \tag{8.26}$$

Here the second expression is a bilinear form, $a(\mathbf{p})$ is defined as in (5.74) (now in momentum space), and the identity can be verified as in lemma 5.5. There is also a representation of the Poincaré group which on the relativistic Fock space $\mathcal{F}^+(\mathcal{H})$ is

$$U(a,\Lambda) = \Gamma(u(a,\Lambda)) \tag{8.27}$$

8.2 Scalar fields

8.2.1 Hamiltonian formulation

In section 7.3.1 we introduced the classical scalar field $\phi(t,\mathbf{x})$ on $\mathbb{R}^4$ satisfying the Klein–Gordon equation

$$(-\Box + m^2)\phi = 0 \tag{8.28}$$

This can also be written as a pair of first-order equations for functions $\phi(\mathbf{x}), \pi(\mathbf{x})$

$$\begin{aligned} \frac{d\phi}{dt} &= \pi \\ \frac{d\pi}{dt} &= -(-\Delta + m^2)\phi \end{aligned} \tag{8.29}$$

This is an infinite-dimensional Hamiltonian system with the Hamiltonian

$$H(\phi, \pi) = \frac{1}{2}\int_{\mathbb{R}^3} (\pi^2(\mathbf{x}) + (\nabla\phi(\mathbf{x}))^2 + m^2\phi^2(\mathbf{x}))d\mathbf{x} \tag{8.30}$$

That is we have the functional derivatives $\partial H/\partial\pi(\mathbf{x}) = \pi(\mathbf{x})$ and $\partial H/\partial\phi(\mathbf{x}) = ((-\Delta + m^2)\phi)(\mathbf{x})$. We are computing these derivatives formally, although they could be given a rigorous meaning without too much trouble.

In terms of the pair $\Phi = (\phi, \pi)$ the Hamiltonian can be written

$$H(\Phi) = \frac{1}{2}(\Phi, \hat{H}\Phi) \tag{8.31}$$

where the inner product is in (real) $L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$ and where

$$\hat{H} = \begin{pmatrix} (-\Delta + m^2) & 0 \\ 0 & I \end{pmatrix} \tag{8.32}$$

The equations become with $\nabla_\Phi = (\partial/\partial\phi, \partial/\partial\pi)$

$$\frac{d\Phi}{dt} = J\nabla_\Phi H = J\hat{H}\Phi \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{8.33}$$

Let Φ_t be the solution of this equation with $\Phi_0 = \Phi \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$. Time evolution preserves this space and Φ_t is linear in Φ so we can write

$$\Phi_t = T(t)\Phi \tag{8.34}$$

where $T(t)$ is a linear operator on $\mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$. It satisfies $T(t)T(s) = T(t+s)$ and $T(0) = I$. An explicit expression for $T(t)$ is given below.

Since we have a linear Hamiltonian system there is a symplectic form invariant under time evolution, see the remarks at the end of section 2.3. It is given by

$$\sigma(\Phi_1, \Phi_2) = (\Phi_1, J\Phi_2) = (\phi_1, \pi_2) - (\pi_1, \phi_2) \tag{8.35}$$

The invariance $\sigma(\Phi_{1,t}, \Phi_{2,t}) = \sigma(\Phi_1, \Phi_2)$ follows from Green's identity (7.40). Alternatively note that $J\hat{H}$ is skew-symplectic

$$\sigma(J\hat{H}\Phi_1, \Phi_2) = (\hat{H}\Phi_1, \Phi_2) = (\Phi_1, \hat{H}\Phi_2) = -\sigma(\Phi_1, J\hat{H}\Phi_2) \tag{8.36}$$

Hence

$$\frac{d}{dt}\,\sigma(\Phi_{1,t}, \Phi_{2,t}) = \sigma(J\hat{H}\Phi_{1,t}, \Phi_{2,t}) + \sigma(\Phi_{1,t}, J\hat{H}\Phi_{2,t}) = 0 \tag{8.37}$$

which gives the invariance.

Problem 8.5 With $\hat{\omega} = \sqrt{-\Delta + m^2}$ show that

$$T(t) = \begin{pmatrix} \cos(\hat{\omega}t) & \hat{\omega}^{-1}\sin(\hat{\omega}t) \\ -\hat{\omega}\sin(\hat{\omega}t) & \cos(\hat{\omega}t) \end{pmatrix} \tag{8.38}$$

Problem 8.6 Show that $T(t)$ commutes with $J\hat{H}$.

8.2.2 Canonical quantization

To quantize this Hamiltonian system by our canonical procedure we would begin with a representation of the CCR which we interpret as

$$\begin{aligned} [\phi(\mathbf{x}), \phi(\mathbf{y})] &= [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0 \\ [\phi(\mathbf{x}), \pi(\mathbf{y})] &= i\delta(\mathbf{x} - \mathbf{y}) \end{aligned} \tag{8.39}$$

Then put these operators in the classical Hamiltonian (8.30) to get a quantum Hamiltonian H. Then define field operators in the Heisenberg picture at time t by

$$\phi(t, \mathbf{x}) = e^{iHt}\phi(\mathbf{x})e^{-iHt} \qquad \pi(t, \mathbf{x}) = e^{iHt}\pi(\mathbf{x})e^{-iHt} \tag{8.40}$$

However H constructed in this manner is rather ill-defined.

For an alternate strategy recall that the operators $\phi(t, \mathbf{x}), \pi(t, \mathbf{x})$ should satisfy the classical dynamical equations (Ehrenfest's theorem). Thus we look for operator valued solutions of (8.29) which satisfy the CCR (8.39) at time zero.

Furthermore we treat the fields as distributions and consider averaged fields

$$\phi(h) = \int \phi(\mathbf{x})h(\mathbf{x})d\mathbf{x} \qquad \pi(h) = \int \pi(\mathbf{x})h(\mathbf{x})d\mathbf{x} \tag{8.41}$$

where the test functions h are in the (real) Schwartz space $\mathcal{S}(\mathbb{R}^3)$. The interpretation of the CCR is now that $\phi(h), \pi(h)$ should be symmetric operators on a dense invariant domain in some Hilbert space satisfying

$$\begin{aligned} [\phi(h), \phi(g)] &= [\pi(h), \pi(g)] = 0 \\ [\phi(h), \pi(g)] &= i(h, g) \end{aligned} \tag{8.42}$$

The dynamical equations are also interpreted in the sense of distributions.

We also consider the pair $\Phi = (\phi, \pi)$. As a test function for the pair we take a pair $F = (f, g)$ of elements in $\mathcal{S}(\mathbb{R}^3)$. Then we form the particular combination

$$\sigma(\Phi, F) = \phi(g) - \pi(f) \tag{8.43}$$

corresponding to the symplectic form. Then the CCR (8.42) imply

$$[\sigma(\Phi, F_1), \sigma(\Phi, F_2)] = i\sigma(F_1, F_2) \tag{8.44}$$

Conversely if we have a representation $\sigma(\Phi, F)$ of (8.44), then the operators $\phi(g) = \sigma(\Phi, (0, g))$ and $\pi(f) = -\sigma(\Phi, (f, 0))$ satisfy (8.42). The problem is now to find a distribution solution of $d\Phi/dt = J\hat{H}\Phi$ with initial conditions satisfying (8.44).

If we have a representation of the CCR, then a formal solution is $\Phi_t = T(t)\Phi$ and we take this as the definition of Φ_t interpreted in the sense of distributions. Thus taking into account the invariance of the symplectic form on functions we define

$$\sigma(\Phi_t, F) = \sigma(\Phi, T(-t)F) \tag{8.45}$$

Taking into account that it is skew-symplectic, we define $J\hat{H}\Phi$ as a distribution by

$$\sigma(J\hat{H}\Phi, F) = -\sigma(\Phi, J\hat{H}F) \tag{8.46}$$

Then we have

$$\begin{aligned} \frac{d}{dt}\sigma(\Phi_t, F) &= \frac{d}{dt}\sigma(\Phi, T(-t)F) = -\sigma(\Phi, J\hat{H}T(-t)F) \\ &= -\sigma(\Phi, T(-t)J\hat{H}F) = \sigma(J\hat{H}\Phi_t, F) \end{aligned} \tag{8.47}$$

which is the result we want.

There are many quantizations corresponding to different representations of the CCR. In the finite-dimensional case it does not matter which we take since (as it turns out) they are all unitarily equivalent and so physically equivalent. In the present infinite-dimensional case this is not true and different choices correspond to different physics. Which should we choose?

We have two criteria. The first is that we want a representation which differs only locally from empty space. This is the appropriate choice to describe a finite number of free elementary particles. For statistical mechanics or for interacting fields we would want to make a different choice. The second criterion is that the energy (that is the spectrum of the Hamiltonian) should be positive.

Our choice is based on a Fock space for a free relativistic particle, thus realizing the field-particle duality. As in section 8.1.5 we take the symmetric Fock space $\mathcal{F}^+(\mathcal{H}_0)$ where $\mathcal{H}_0 = L^2(\mathbb{R}^3, d\mathbf{p})$. If a^*, a are the creation and annihilation operators on this Fock space, we define on the domain $\mathcal{D}_0$ of finite particle vectors (see (5.59))

$$\sigma(\Phi, F) = i(a(K_0F) - a^*(K_0F)) \tag{8.48}$$

where with $\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}$

$$K_0(f, g) = \frac{1}{\sqrt{2}}\left(\omega^{1/2}\tilde{f} + i\omega^{-1/2}\tilde{g}\right) \tag{8.49}$$

The operator $K_0 : \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3) \to L^2(\mathbb{R}^3, d\mathbf{p})$ is selected to satisfy two criteria. The first is that it is symplectic from (real) $\mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$ with symplectic form σ to (complex) $L^2(\mathbb{R}^3)$ with symplectic form which is twice the imaginary part of the inner product, that is

$$2\,\mathrm{Im}(K_0F_1, K_0F_2) = \sigma(F_1, F_2) \tag{8.50}$$

This will give the CCR. Second it maps a real solution $F_t = T(t)F$ of the Klein–Gordon equation to a complex positive energy solution (in momentum space), that is

$$K_0 F_t = e^{-i\omega t} K_0 F \tag{8.51}$$

This will give the positive energy. Both of these are easily checked.

Let us see that this works:

Theorem 8.1 *The operators $\sigma(\Phi, F)$ defined by (8.48) (8.49) satisfy the CCR, and time evolution is implemented by the free particle Hamiltonian $H = d\Gamma(\omega) \geq 0$ in the sense that*

$$\sigma(\Phi_t, F) = e^{iHt}\sigma(\Phi, F)e^{-iHt} \tag{8.52}$$

Proof It gives a representation of the CCR since

$$\begin{aligned}[\sigma(\Phi, F_1), \sigma(\Phi, F_2)] &= (K_0F_1, K_0F_2) - (K_0F_2, K_0F_1) \\ &= 2i\mathrm{Im}(K_0F_1, K_0F_2) \\ &= i\sigma(F_1, F_2)\end{aligned} \tag{8.53}$$

The time evolution is

$$\begin{aligned}\sigma(\Phi_t, F) &= \sigma(\Phi, T(-t)F) \\ &= i(a(K_0T(-t)F) - a^*(K_0T(-t)F)) \\ &= i(a(e^{i\omega t}K_0F) - a^*(e^{i\omega t}K_0F)) \\ &= e^{iHt}\sigma(\Phi, F)e^{-iHt}\end{aligned} \tag{8.54}$$

the last step by $e^{-iHt} = \Gamma(e^{-i\omega t})$ and (5.69). □

The field $\phi(t, h) = \sigma(\Phi_t, (0, h))$ and the momentum $\pi(t, h) = \sigma(\Phi_t, (-h, 0))$ are given by

$$\begin{aligned}\phi(t, h) &= a\left(\frac{e^{i\omega t}\tilde{h}}{\sqrt{2\omega}}\right) + a^*\left(\frac{e^{i\omega t}\tilde{h}}{\sqrt{2\omega}}\right) \\ \pi(t, h) &= -ia\left(\sqrt{\omega/2}e^{i\omega t}\tilde{h}\right) + ia^*\left(\sqrt{\omega/2}e^{i\omega t}\tilde{h}\right)\end{aligned} \tag{8.55}$$

The first can also be written as the distribution identity

$$\phi(t, \mathbf{x}) = (2\pi)^{-3/2}\int \left(a(\mathbf{p})e^{-i(\omega(\mathbf{p})t - \mathbf{px})} + a^*(\mathbf{p})e^{i(\omega(\mathbf{p})t - \mathbf{px})}\right)\frac{d\mathbf{p}}{\sqrt{2\omega(\mathbf{p})}} \tag{8.56}$$

That is $\phi(t, h) = \int \phi(t, \mathbf{x})h(\mathbf{x})d\mathbf{x}$ has a meaning and the meaning is (8.55). But (8.56) can also be understood pointwise as a bilinear form on $\mathcal{D}_\mathcal{S} \times \mathcal{D}_\mathcal{S}$ as in section 5.4.2.

Our time evolution is generated by $H = d\Gamma(\omega)$. But is this the quantization of the classical Hamiltonian (8.30)? Indeed it is, suitably interpreted. The claim is that with $\phi(\mathbf{x}) = \phi(0, \mathbf{x})$ and $\pi(\mathbf{x}) = (d\phi/dt)(0, \mathbf{x})$ given by (8.56)

$$\frac{1}{2}\int :(\pi^2(\mathbf{x}) + (\nabla\phi(\mathbf{x}))^2 + m^2\phi^2(\mathbf{x})): d\mathbf{x} = \int a^*(\mathbf{p})\omega(\mathbf{p})a(\mathbf{p})d\mathbf{p} \tag{8.57}$$

Here on the left the symbol $: \cdots :$ is *Wick ordering* which means move all creation operators to the left and all annihilation operators to the right. Then both sides of the equation are well-defined as bilinear forms on $\mathcal{D}_{\mathcal{S}} \times \mathcal{D}_{\mathcal{S}}$ and the claim is that they are equal. We leave the details as a problem.

Problem 8.7 Verify (8.50), (8.51). (For the latter use (8.38) or $K_0 J\hat{H} = -i\omega K_0$.)

Problem 8.8 Verify (8.57).

8.2.3 Generalization

Next we digress and give a more abstract version of the construction we used in theorem 8.1. Let (S, σ) be a real vector space with symplectic form σ, that is σ is a bilinear form on S which is skew-symmetric and nondegenerate. Let $\mathcal{H}$ be a complex Hilbert space taken as a real vector space with symplectic form which is twice the imaginary part of the inner product. Let $K : S \to \mathcal{H}$ be symplectic with dense range. Then for $F \in S$

$$\sigma(\Phi, F) = i(a(KF) - a^*(KF)) \tag{8.58}$$

on $\mathcal{D}_0 \subset \mathcal{F}^+(\mathcal{H})$ satisfies the CCR (8.44).

Further suppose $T(t)$ is a linear symplectic time evolution on $\mathcal{S}$ and that $KT(t) = e^{-iht}K$ for some self-adjoint operator h on $\mathcal{H}$. Then the time evolution defined by $\sigma(\Phi_t, F) = \sigma(\Phi, T(-t)F)$ is unitarily implemented with $U(t) = \Gamma(e^{-iht})$ in the sense that

$$\sigma(\Phi_t, F) = U(t)^{-1}\sigma(\Phi, F)U(t) \tag{8.59}$$

The following problems refer to this abstract construction.

Problem 8.9 Let $(\mathcal{H}, K, h)$ and $(\mathcal{H}', K', h')$ be two such structures for $(S, \sigma, T(t))$ with fields $\sigma(\Phi, F)$, $\sigma(\Phi', F)$. Establish the following:

1. If there is a unitary operator $V : \mathcal{H} \to \mathcal{H}'$ such that $VK = K'$, then there is a unitary $U : \mathcal{F}^+(\mathcal{H}) \to \mathcal{F}^+(\mathcal{H}')$ so
$$U\sigma(\Phi_t, F)U^{-1} = \sigma(\Phi'_t, F) \tag{8.60}$$
2. If h, h' are both positive, then V exists.[1] (So a positive energy representation is essentially unique.)

[1] See Kay (1979).

Problem 8.10 Suppose we want to quantize a field which satisfies the equation $(-\Box + m^2 + V)\phi = 0$ where $V = V(\mathbf{x})$ in $L^2(\mathbb{R}^3)$ only depends on the spatial variables. Construct a suitable field operator satisfying the CCR at time zero, and show that the time evolution is unitarily implemented.

8.2.4 Covariant quantization

Now we give a more covariant construction of the scalar quantum field. The quantization problem is just as before but now we define the field operator on a different space. We start with the relativistic single particle space $\mathcal{H} = L^2(V_m^+, d\mu_m)$ which is identified with $\mathcal{H} = L^2(\mathbb{R}^3, d\mathbf{p}/2\omega(\mathbf{p}))$. Our map K from $\mathcal{S}(\mathbb{R}^3)$ to $\mathcal{H}$ is now taken to be

$$K(f,g) = \omega\tilde{f} + i\tilde{g} \tag{8.61}$$

and we still have

$$\begin{aligned} 2\,\mathrm{Im}(KF_1, KF_2) &= \sigma(F_1, F_2) \\ KF_t &= e^{-i\omega t}KF \end{aligned} \tag{8.62}$$

The field operator is again

$$\sigma(\Phi, F) = i(a(KF) - a^*(KF)) \tag{8.63}$$

on $\mathcal{D}_0 \subset \mathcal{F}^+(\mathcal{H})$. Time evolution is defined by $\sigma(\Phi_t, F) = \sigma(\Phi, T(-t)F)$ as before and we find that $\phi(t,h) = \sigma(\Phi_t, (0,h))$ is given by

$$\phi(t,h) = a(e^{i\omega t}\tilde{h}) + a^*(e^{i\omega t}\tilde{h}) \tag{8.64}$$

Since the operator $V = [\sqrt{2\omega(\mathbf{p})}]$ is unitary from $\mathcal{H}_0$ to $\mathcal{H}$ and satisfies $VK_0 = K$, the field operators are unitarily equivalent to those of the previous section by $\Gamma(V)$. (This is a special case of problem 8.9.)

It is also useful to consider a field smeared in space and time. Formally we have $\phi(t,h) = \int \phi(t,\mathbf{x})h(\mathbf{x})d\mathbf{x}$, so for real $f \in \mathcal{S}(\mathbb{R}^4)$ the operator $\phi(f) = \int \phi(t,\mathbf{x})f(t,\mathbf{x})dtd\mathbf{x}$ would be defined by

$$\phi(f) = \int \phi(t, f(t,\cdot))dt \tag{8.65}$$

This is evaluated as

$$\phi(f) = a(\Pi_+ f) + a^*(\Pi_+ f) \tag{8.66}$$

where

$$(\Pi_+ f)(\mathbf{p}) = \sqrt{2\pi}\hat{f}(\omega(\mathbf{p}), \mathbf{p}) \tag{8.67}$$

and where $\hat{f}$ is again the Fourier transform with the Lorentz inner product (see (7.88)). Thus $\Pi_+ f$ is essentially the restriction of the Fourier transform of f to the mass shell V_m^+. Since $(-\Box + m^2)f$ has a Fourier transform which vanishes on the mass shell we have

$$\phi((-\Box + m^2)f) = 0 \tag{8.68}$$

This says that ϕ satisfies the Klein–Gordon equation $(-\Box + m^2)\phi = 0$ in the sense of distributions.

We can also exhibit the covariance under proper Poincaré transformations (a, Λ). If $f_{a,\Lambda}(x) = f(\Lambda^{-1}(x - a))$ is the transformed test function, then $\Pi_+(f_{a,\Lambda}) = u(a, \Lambda)(\Pi_+ f)$ where $u(a, \Lambda)$ is the unitary representation defined in (8.16). Hence with $U(a, \Lambda) = \Gamma(u(a, \Lambda))$ we have

$$\phi(f_{a,\Lambda}) = U(a, \Lambda)\phi(f)U(a, \Lambda)^{-1} \tag{8.69}$$

This shows that the Poincaré transformed field is unitarily equivalent to the original field.

Finally we consider the commutator of two fields. We compute taking into account the reality condition $\overline{\hat{f}(p)} = \hat{f}(-p)$

$$\begin{aligned} &[\phi(g), \phi(f)] \\ &= (\Pi_+ g, \Pi_+ f)_{\mathcal{H}} - (\Pi_+ f, \Pi_+ g)_{\mathcal{H}} \\ &= 2\pi \int \left(\hat{g}(-\omega(\mathbf{p}), -\mathbf{p})\hat{f}(\omega(\mathbf{p}), \mathbf{p}) - \hat{f}(-\omega(\mathbf{p}), -\mathbf{p})\hat{g}(\omega(\mathbf{p}), \mathbf{p})\right) \frac{d\mathbf{p}}{2\omega(\mathbf{p})} \\ &= \frac{1}{i} < g, Ef > \end{aligned} \tag{8.70}$$

where $< g, f >= \int g(x)f(x)dx$ and $E = E^+ - E^-$ is the propagator (7.91). If f, g have spacelike separated supports, that is if supp $g \cap J^{\pm}(\text{supp}\, f) = \emptyset$, then supp $g \cap$ supp $E^{\pm} f = \emptyset$; hence $< g, Ef >= 0$, and so $[\phi(g), \phi(f)] = 0$. This strict locality result is a manifestation of the basic fact that influence cannot propagate faster than the speed of light.

Let us summarize writing (8.68), (8.69), (8.70) in distribution form:

Theorem 8.2 *(The free field) Let ϕ be the field operator by (8.66). Then in the sense of distributions*

1. *(Field equation)*

$$(-\Box + m^2)\phi = 0 \tag{8.71}$$

2. *(Covariance) For any proper Poincaré transformation $(a, \Lambda) \in \mathcal{P}_+^{\uparrow}$*

$$U(a, \Lambda)\phi(x)U(a, \Lambda)^{-1} = \phi(\Lambda x + a) \tag{8.72}$$

3. *(Locality) In the region of spacelike separation $(x - y) \cdot (x - y) > 0$*

$$[\phi(x), \phi(y)] = 0 \tag{8.73}$$

8.3 Charged scalar field

8.3.1 Hamiltonian formulation

First we write the equations for the classical charged scalar field as a Hamiltonian system. We have four functions (ϕ_1, π_1) and (ϕ_2, π_2) each pair of which satisfies (8.29). This is a Hamiltonian system with Hamiltonian

$$H(\phi, \pi) = \frac{1}{2} \sum_{i=1,2} \int (\pi_i^2 + (\nabla \phi_i)^2 + m^2 \phi_i^2) d\mathbf{x} \tag{8.74}$$

Since $\partial^0 \phi_i = -\partial_0 \phi_i = -\pi_i$, the charge (7.41) is

$$Q = \int (\phi_2 \pi_1 - \pi_2 \phi_1) d\mathbf{x} \tag{8.75}$$

It is essentially the symplectic form (8.35) but has a new interpretation in this context.

Here is another formal way to arrive at a conserved charge. Note that the Hamiltonian is invariant under rotations in $\mathbb{R}^2$, that is with $\phi = (\phi_1, \phi_2)$ and $\pi = (\pi_1, \pi_2)$ we have $H(\phi(\theta), \pi(\theta)) = H(\phi, \pi)$ where $\phi(\theta) = R(-\theta)\phi$ and $\pi(\theta) = R(-\theta)\pi$ and

$$R(-\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{8.76}$$

This is a global (that is constant) gauge transformation. But $\phi(\theta), \pi(\theta)$ are the solutions of the equations

$$\begin{aligned} \frac{d\phi_1}{d\theta} &= \phi_2 & \frac{d\pi_1}{d\theta} &= \pi_2 \\ \frac{d\phi_2}{d\theta} &= -\phi_1 & \frac{d\pi_2}{d\theta} &= -\pi_1 \end{aligned} \tag{8.77}$$

This is a Hamiltonian system with generator Q given by (8.75), that is $\partial Q/\partial \pi_1 = \phi_2$ and $-\partial Q/\partial \phi_1 = \pi_2$ and so forth. As explained in section 2.4 the invariance of H under the flow of Q implies that $\{H, Q\} = 0$, which in turn implies that Q is invariant under the flow of H, that is charge is conserved.

It will also be useful to regard the field as complex valued and introduce

$$\begin{aligned} \phi &= \frac{\phi_1 + i\phi_2}{\sqrt{2}} & \phi^* &= \frac{\phi_1 - i\phi_2}{\sqrt{2}} \\ \pi &= \frac{\pi_1 - i\pi_2}{\sqrt{2}} & \pi^* &= \frac{\pi_1 + i\pi_2}{\sqrt{2}} \end{aligned} \tag{8.78}$$

Then (ϕ, π) and (ϕ^*, π^*) are pairs of conjugate variables. The Hamiltonian is

$$H = \int (\pi^* \pi + \nabla \phi^* \cdot \nabla \phi + m^2 \phi^* \phi) d\mathbf{x} \tag{8.79}$$

and the charge is

$$Q = i \int (\phi^* \pi^* - \phi \pi) d\mathbf{x} \tag{8.80}$$

The gauge transformations generated by Q now take the form $\phi \to e^{-i\theta}\phi$, $\pi^* \to e^{-i\theta}\pi^*$, and $\phi^* \to e^{i\theta}\phi^*$, $\pi \to e^{i\theta}\pi$ and they are still a symmetry of the Hamiltonian.

8.3.2 Canonical quantization

For quantization we look for symmetric operators ϕ_1, π_1 and ϕ_2, π_2, each pair satisfying the field equations (8.29) and the CCR, and commuting self-adjoint operators H, Q implementing time evolution and gauge transformations. With $\Phi_i = (\phi_i, \pi_i)$, $i = 1, 2$ and $F = (f, g)$ in real $\mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$ the combination $\sigma(\Phi_i, F) = \phi_i(g) - \pi_i(f)$ should satisfy

$$\begin{aligned}
[\sigma(\Phi_i, F), \sigma(\Phi_i, F')] &= i\sigma(F, F') \\
[\sigma(\Phi_1, F), \sigma(\Phi_2, F')] &= 0 \\
e^{iHt}\sigma(\Phi_i, F)e^{-iHt} &= \sigma(\Phi_i, T(-t)F) \\
e^{iQ\theta}\begin{pmatrix} \sigma(\Phi_1, F) \\ \sigma(\Phi_2, F) \end{pmatrix} e^{-iQ\theta} &= R(-\theta)\begin{pmatrix} \sigma(\Phi_1, F) \\ \sigma(\Phi_2, F) \end{pmatrix}
\end{aligned} \tag{8.81}$$

We can also formulate this in terms of $\Phi = (\Phi_1 + i\Phi_2)/\sqrt{2} = (\phi, \pi^*)$ and its adjoint $\Phi^* = (\Phi_1 - i\Phi_2)/\sqrt{2} = (\phi^*, \pi)$. It is equivalent to ask for operators $\sigma(\Phi, F)$ and $\sigma(\Phi^*, F)$, which are adjoint to each other and self-adjoint H, Q which satisfy

$$\begin{aligned}
[\sigma(\Phi, F), \sigma(\Phi^*, F')] &= i\sigma(F, F') \\
[\sigma(\Phi, F), \sigma(\Phi, F')] &= 0 \\
e^{iHt}\sigma(\Phi, F)e^{-iHt} &= \sigma(\Phi, T(-t)F) \\
e^{iQ\theta}\sigma(\Phi, F)e^{-iQ\theta} &= e^{-i\theta}\sigma(\Phi, F)
\end{aligned} \tag{8.82}$$

The first three equations in (8.81) or (8.82) are familiar from the discussion of the neutral scalar field. The last is the quantum version of the classical statement that charge generates global gauge transformations.

To realize this structure we introduce a second particle and follow the treatment of section 8.2.2. On the Hilbert space $\mathcal{F}^+(\mathcal{H}_0) \otimes \mathcal{F}^+(\mathcal{H}_0)$ we define

$$a(f) = a(f) \otimes I \qquad b(f) = I \otimes a(f) \tag{8.83}$$

The operators $a(f), a^*(f), b(f), b^*(f)$ satisfy

$$[a(f), a^*(g)] = (f, g) \qquad\qquad [b(f), b^*(g)] = (f, g) \tag{8.84}$$

with all other commutators equal to zero. Then we define

$$\begin{aligned}\sigma(\Phi, F) &= i(a(K_0F) - b^*(K_0F))\\ \sigma(\Phi^*, F) &= i(b(K_0F) - a^*(K_0F))\end{aligned} \tag{8.85}$$

These are adjoint to each other and the identities (8.82) are satisfied with $e^{-iHt} = \Gamma(e^{-i\omega t}) \otimes \Gamma(e^{-i\omega t})$ and $e^{-iQ\theta} = \Gamma(e^{-i\theta}) \otimes \Gamma(e^{i\theta})$. The first three follow as before. For the last, note that

$$\begin{aligned}e^{iQ\theta}\sigma(\Phi, F)e^{-iQ\theta} &= i\Big(a(e^{i\theta}K_0F) - b^*(e^{-i\theta}K_0F)\Big)\\ &= e^{-i\theta}i\Big(a(K_0F) - b^*(K_0F)\Big)\\ &= e^{-i\theta}\sigma(\Phi, F)\end{aligned} \tag{8.86}$$

The Hamiltonian and the charge are given by

$$\begin{aligned}H &= d\Gamma(\omega) \otimes I + I \otimes d\Gamma(\omega)\\ &= \int \omega(\mathbf{p})[a^*(\mathbf{p})a(\mathbf{p}) + b^*(\mathbf{p})b(\mathbf{p})]d\mathbf{p}\\ Q &= N \otimes I - I \otimes N\\ &= \int [a^*(\mathbf{p})a(\mathbf{p}) - b^*(\mathbf{p})b(\mathbf{p})]d\mathbf{p}\end{aligned} \tag{8.87}$$

The charge can also be written as $Q = N_a - N_b$ where $N_a = N \otimes I$ is the number of a particles and $N_b = I \otimes N$ is the number of b particles. Thus we see that the a particles carry charge $+1$ (in natural units) while the b particles carry charge -1. Charge takes integer values – "charge is quantized." The b particles are called the *anti-particles* of the a particles. Both are necessary to get the structure we want.

Problem 8.11

1. Find expressions for $\phi(t, \mathbf{x}), \pi(t, \mathbf{x}), \phi^*(t, \mathbf{x}), \pi^*(t, \mathbf{x})$ as bilinear forms on $\mathcal{D}_\mathcal{S} \times \mathcal{D}_\mathcal{S}$. as in (8.56).
2. Show that the quantum charge Q in (8.87) regarded as a bilinear form on $\mathcal{D}_\mathcal{S} \times \mathcal{D}_\mathcal{S}$ has the same form as the classical charge (8.80), that is

$$Q = i\int : (\phi^*(\mathbf{x})\pi^*(\mathbf{x}) - \phi(\mathbf{x})\pi(\mathbf{x})) : d\mathbf{x} \tag{8.88}$$

where $\phi(\mathbf{x}) = \phi(0, \mathbf{x})$, etc. (Wick ordering was defined after equation (8.57).)

Notes on chapter 8: In general charge can be defined as the labels for the irreducible representations of the internal symmetry group, see Haag (1992). In our case the symmetry group is $U(1)$ and the irreducible representations $e^{i\theta} \to e^{in\theta}$ are labeled by integers n.

Everything in this chapter could be done in a spacetime $(\mathbb{R}^d, \eta)$ with $d \geq 2$ instead of specifically $d = 4$.

There are many physics texts on quantum field theory. Early works by Schweber (1962) and Roman (1969) are still useful. Some more recent books are Itzykson and Zuber (1980), Weinberg (1995) and Peskin and Schroeder (1995). Our notation mostly agrees with Weinberg.

A mathematical treatment of free fields can be found in Baez *et al.* (1992).

9 Electrons and photons

In this chapter we explore the quantum mechanical description of free electrons and photons and associated fields.

However we start with a digression developing the transformation properties of spinors under the Lorentz group, or more precisely the covering group of the Lorentz group. The discussion extends the discussion of spin in section 4.6 and the Dirac equation in section 7.3.

9.1 Spinors

Given 4×4 gamma matrices satisfying $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$, let V be the real vector space of all matrices of the form $\gamma \cdot a = \eta_{\mu\nu}\gamma^\mu a^\mu$ with $a \in \mathbb{R}^4$. Consider the group of nonsingular 4×4 matrices S so that $S(\gamma \cdot a)S^{-1}$ is again in V, called $Pin(1,3)$.[1] Since the γ matrices are a basis (problem 7.6) this means that $S(\gamma \cdot a)S^{-1} = \gamma \cdot a'$ for a unique $a' \in \mathbb{R}^4$. The map $a \to a'$ is linear and so $a' = \Lambda(S)a$ for some matrix $\Lambda(S)$. Thus $\Lambda(S)$ is defined by

$$S(\gamma \cdot a)S^{-1} = \gamma \cdot \Lambda(S)a \tag{9.1}$$

Lemma 9.1 *The map $S \to \Lambda(S)$ is a two-to-one homomorphism from $Pin(1,3)$ onto the Lorentz group $\mathcal{L} = O(1,3)$.*

Proof The map is a homomorphism since

$$\begin{aligned} \gamma \cdot \Lambda(ST)a &= ST(\gamma \cdot a)(ST)^{-1} \\ &= S(\gamma \cdot \Lambda(T)a)S^{-1} \\ &= \gamma \cdot (\Lambda(S)\Lambda(T))a \end{aligned} \tag{9.2}$$

which implies $\Lambda(ST) = \Lambda(S)\Lambda(T)$. The matrix $\Lambda(S)$ is a Lorentz transformation since

$$\Lambda(S)a \cdot \Lambda(S)a = (\gamma \cdot \Lambda(S)a)^2 = (S(\gamma \cdot a)S^{-1})^2 = a \cdot a \tag{9.3}$$

[1] The name $Pin(1,3)$ is a truncation of $Spin(1,3)$ defined below.

To see that the mapping is onto note that for $\Lambda \in \mathcal{L}$ we have that $(\gamma')^\mu = (\Lambda^{-1})^\mu_{\ \nu}\gamma^\nu$ is also a set of gamma matrices. But any two such are related by $(\gamma')^\mu = M\gamma^\mu M^{-1}$ for some M and so we have

$$M(\gamma \cdot a)M^{-1} = \Lambda^{-1}\gamma \cdot a = \gamma \cdot (\Lambda a) \tag{9.4}$$

Hence $M \in Pin(1,3)$ and $\Lambda(M) = \Lambda$.

The mapping is at least two-to-one since $\Lambda(-S) = \Lambda(S)$. To see it is exactly two-to-one it suffices to show that $\Lambda(S) = I$ implies $S = \pm I$, which we omit. □

Recall that $\mathcal{L}$ has a subgroup $\mathcal{L}_+ = SO(1,3)$ defined by the condition $\det \Lambda = 1$. This is covered by

$$Spin(1,3) = \{S \in Pin(1,3) : \Lambda(S) \in \mathcal{L}_+\} \tag{9.5}$$

which is a subgroup of $Pin(1,3)$. Furthermore $\mathcal{L}_+$ has a connected subgroup $\mathcal{L}_+^\uparrow$ defined by $\Lambda^0_{\ 0} \geq 1$, which is covered by

$$Spin^\uparrow(1,3) = \{S \in Spin(1,3) : \Lambda(S) \in \mathcal{L}_+^\uparrow\} \tag{9.6}$$

This is a subgroup of $Spin(1,3)$ and is connected as well (problem 9.4).

Next consider the indefinite inner product $u^\dagger \beta v$ of (7.53) on the spinor space $\mathbb{C}^4$. We show it is invariant under $Spin^\uparrow(1,3)$.

Lemma 9.2 *For $S \in Spin^\uparrow(1,3)$, $u, v, \in \mathbb{C}^4$, and $\beta = i\gamma^0$*

$$(Su)^\dagger \beta S v = u^\dagger \beta v \tag{9.7}$$

Equivalently if S^* is the adjoint with the usual inner product

$$S^*\gamma^0 S = \gamma^0 \tag{9.8}$$

Proof Take the adjoint of $S(\gamma \cdot x)S^{-1} = \gamma \cdot \Lambda(S)x$. Then with $\hat{x} = (-x_0, \mathbf{x})$ we find

$$(S^{-1})^*(\gamma \cdot \hat{x})S^* = \gamma \cdot \widehat{(\Lambda(S)x)} \tag{9.9}$$

However $\gamma \cdot \hat{x} = \gamma^0(\gamma \cdot x)\gamma^0$, so we have

$$[\gamma^0(S^{-1})^*\gamma^0](\gamma \cdot x)[\gamma^0 S^*\gamma^0] = \gamma \cdot (\Lambda(S)x) \tag{9.10}$$

This implies that

$$\gamma^0(S^{-1})^*\gamma^0 = \pm S \tag{9.11}$$

and so

$$S^*\gamma^0 S = \pm\gamma^0 \tag{9.12}$$

The sign is a continuous function of S on the connected set $Spin^\uparrow(1,3)$. Since it is $+1$ at the identity, it is $+1$ everywhere. □

Problem 9.1 If $S = \gamma \cdot y$ with $y \cdot y \neq 0$, show that $S \in Pin(1,3)$ and find $\Lambda(S)$.

Problem 9.2 Show that

$$e^{-\gamma^2\gamma^3\theta/2} = \cos(\theta/2) - \gamma^2\gamma^3 \sin(\theta/2) \tag{9.13}$$

is in $Spin^{\uparrow}(1,3)$ and that

$$\Lambda(e^{-\gamma^2\gamma^3\theta/2}) = R(e_1, \theta) \tag{9.14}$$

where $R(e_1, \theta)$ is the rotation by θ around the first axis.

Problem 9.3 Show that

$$e^{\gamma^0\gamma^1\beta/2} = \cosh(\beta/2) + \gamma^0\gamma^1 \sinh(\beta/2) \tag{9.15}$$

is in $Spin^{\uparrow}(1,3)$ and that

$$\Lambda(e^{\gamma^0\gamma^1\beta/2}) = \Lambda_\beta \tag{9.16}$$

where Λ_β is the boost (7.36).

Problem 9.4 Show that $Spin^{\uparrow}(1,3)$ is connected by showing that every element can be joined to the identity by a continuous path.

Problem 9.5 Show that if ψ is a solution of $(\gamma \cdot \partial + m)\psi = 0$, then $\psi_{a,S}(x) = S\psi(\Lambda(S)^{-1}(x-a))$ is also a solution.

9.2 Electrons

9.2.1 Solutions of the Dirac equation

We consider the Dirac equation $(\gamma \cdot \partial + m)\psi = 0$ defined in section 7.3.3 and try to interpret it as describing the quantum time evolution for a single electron or proton or any other massive spin 1/2 particle. We begin by rewriting it and exhibiting some solutions.

Consider the self-adjoint 4×4 matrices

$$\beta = i\gamma^0 \qquad \alpha^k = -\gamma^0\gamma^k \qquad k = 1,2,3 \tag{9.17}$$

These have the anti-commutation relations

$$\{\alpha^j, \alpha^k\} = 2\delta^{jk} \qquad \{\alpha^k, \beta\} = 0 \qquad \beta^2 = I \tag{9.18}$$

and they are traceless. Multiplying the Dirac equation by $i\gamma^0$, it takes the form

$$i\frac{\partial \psi}{\partial t} = (-i\nabla \cdot \alpha + \beta m)\psi \tag{9.19}$$

To solve it, note that the partial Fourier transform $\tilde{\psi}(t, \mathbf{p})$ satisfies the equation

$$i\frac{\partial \tilde{\psi}}{\partial t} = (\mathbf{p} \cdot \alpha + \beta m)\tilde{\psi} \tag{9.20}$$

The matrix $\mathbf{p} \cdot \alpha + \beta m$ is self-adjoint and satisfies

$$(\mathbf{p} \cdot \alpha + \beta m)^2 = |\mathbf{p}|^2 + m^2 = \omega(\mathbf{p})^2 \tag{9.21}$$

It follows that the eigenvalues of $\mathbf{p} \cdot \alpha + \beta m$ can only be $\pm\omega(\mathbf{p})$. Furthermore since $tr(\mathbf{p}\cdot\alpha+\beta m) = 0$, both $\pm\omega(\mathbf{p})$ must occur, each with multiplicity 2. For each $\mathbf{p} \in \mathbb{R}^3$ let $W_{\mathbf{p},\pm}$ be the two-dimensional eigenspace for the positive or negative eigenvalue. Then we have

$$\mathbb{C}^4 = W_{\mathbf{p}}^+ \oplus W_{\mathbf{p}}^- \tag{9.22}$$

Now a general solution of (9.20) is

$$\tilde{\psi}(t, \mathbf{p}) = e^{-i\omega(\mathbf{p})t}\psi_+(\mathbf{p}) + e^{i\omega(\mathbf{p})t}\psi_-(\mathbf{p}) \tag{9.23}$$

with $\psi_\pm(\mathbf{p}) \in W_{\mathbf{p}}^\pm$. A general solution of the Dirac equation is

$$\psi(t, \mathbf{x}) = (2\pi)^{-3/2} \int e^{i\mathbf{p}\cdot\mathbf{x}} \Big(e^{-i\omega(\mathbf{p})t}\psi_+(\mathbf{p}) + e^{i\omega(\mathbf{p})t}\psi_-(\mathbf{p}) \Big) d\mathbf{p} \tag{9.24}$$

9.2.2 Quantum interpretation

Now we give the quantum interpretation. First we work with the Fourier transformed variable, which we interpret as momentum space. The Hilbert space is $\mathcal{H}_0 \equiv L^2(\mathbb{R}^3, \mathbb{C}^4, d\mathbf{p})$, the $\mathbb{C}^4$-valued square integrable functions on $\mathbb{R}^3$. Equation (9.20) has the form of a Schrödinger equation if we take as the Hamiltonian

$$\tilde{H} = [\mathbf{p} \cdot \alpha + \beta m] \tag{9.25}$$

The Hilbert space splits into positive and negative energy subspaces

$$\mathcal{H}_0 = \mathcal{H}_0^+ \oplus \mathcal{H}_0^- \tag{9.26}$$

where

$$\mathcal{H}_0^\pm = \{\psi \in \mathcal{H}_0 : \psi(\mathbf{p}) \in W_{\mathbf{p}}^\pm\} \tag{9.27}$$

With respect to this splitting the Hamiltonian is $\tilde{H} = [\omega(\mathbf{p})] \oplus [-\omega(\mathbf{p})]$, which is self-adjoint on its natural domain. Time evolution is $\exp(-i\tilde{H}t) = \exp(-i\omega(\mathbf{p})t) \oplus \exp(i\omega(\mathbf{p})t)$ as in (9.23). To write it another way, note that the projection onto $\mathcal{H}_0^{\pm}$ is

$$P^{\pm} = \left[\frac{\omega(\mathbf{p}) \pm (\mathbf{p} \cdot \alpha + \beta m)}{2\omega(\mathbf{p})}\right] \tag{9.28}$$

Then for any $\psi \in \mathcal{H}_0$ we have

$$e^{-i\tilde{H}t}\psi = e^{-i\omega t}P^{+}\psi + e^{i\omega t}P^{-}\psi \tag{9.29}$$

Return to configuration space $\hat{\mathcal{H}}_0 \equiv L^2(\mathbb{R}^3, \mathbb{C}^4, d\mathbf{x})$ with a Fourier transform $\mathcal{F}^{-1} : \mathcal{H}_0 \to \hat{\mathcal{H}}_0$. The Dirac Hamiltonian is $H_D = \mathcal{F}^{-1}\tilde{H}\mathcal{F}$, which on $\mathcal{S}(\mathbb{R}^3)$ is

$$H_D = -i\nabla \cdot \alpha + \beta m \tag{9.30}$$

The time evolution is $\exp(-iH_D t) = \mathcal{F}^{-1}\exp(-i\tilde{H}t)\mathcal{F}$ as in (9.24).

At this point we remember that we only want positive energy solutions. Thus instead of $\mathcal{H}_0$, the Hilbert space is $\mathcal{H}_0^{+}$, the Hamiltonian is $[\omega(\mathbf{p})]$, and the time evolution is $[\exp(-i\omega(\mathbf{p})t)]$. Taking the inverse Fourier transform gives a Hilbert space $\hat{\mathcal{H}}_0^{+} = \mathcal{F}^{-1}\mathcal{H}_0^{+}$, a Hamiltonian $H_D = \mathcal{F}^{-1}[\omega(\mathbf{p})]\mathcal{F}$, and a time evolution $\mathcal{F}^{-1}[\exp(-i\omega(\mathbf{p})t)]\mathcal{F}$. The wave function with initial condition $\psi \in \hat{\mathcal{H}}_0^{+}$ is

$$\psi(t, \mathbf{x}) = (2\pi)^{-3/2} \int e^{-i(\omega(\mathbf{p}) - \mathbf{p}\cdot\mathbf{x})}\tilde{\psi}(\mathbf{p})d\mathbf{p} \tag{9.31}$$

which still satisfies the Dirac equation.

But there is a problem with this restriction to positive energy. The operator $[x^k]$ can no longer be precisely interpreted as representing the kth coordinate of the particle. This is because it does not act on $\hat{\mathcal{H}}_0^{+}$. To put it another way, $\mathcal{F}[x^k]\mathcal{F}^{-1} = i\partial/\partial p_k$ does not act on the positive energy subspace $\mathcal{H}_0^{+}$. To remedy this we stay in momentum space and define a new coordinate operator – the Newton–Wigner operator. It involves the projection operator at zero momentum which is $(1 + \beta)/2$ and it is defined by

$$X^k = 2P^{+}\sqrt{\frac{\omega}{\omega + m}}\left(\frac{1+\beta}{2}\right)\left(i\frac{\partial}{\partial p_k}\right)\sqrt{\frac{\omega}{\omega + m}}\,P^{+} \tag{9.32}$$

This is a symmetric operator mapping $\mathcal{H}_0^{+} \cap \mathcal{S}(\mathbb{R}^3)$ to itself and is supposed to represent the kth coordinate of the particle. This interpretation is supported by the following result:

Problem 9.6 Show that $X^k, [p_j]$ satisfy the the canonical commutation relations.

9.2.3 Translations and rotations

Next we investigate the effect of spacetime translations and rotations on solutions, excluding boosts for now.

Note that the positive energy condition $(\mathbf{p}\cdot\alpha+\beta m-\omega(p))\psi(\mathbf{p})=0$ can be multiplied by $\beta=i\gamma_0$ and written $(i\gamma\cdot\phi(\mathbf{p})+m)\psi(\mathbf{p})=0$ where $\phi(\mathbf{p})=(\omega(\mathbf{p}),\mathbf{p})$. If $\psi(\mathbf{p})$ satisfies the condition and $S\in Spin^{\uparrow}(1,3)$ is a spinor rotation covering $R\in SO(3)$, then $S\psi(R^{-1}\mathbf{p})$ also satisfies the condition. This follows since $S^{-1}(\gamma\cdot a)S=\gamma\cdot R^{-1}a$ and so

$$\Big(i\gamma\cdot\phi(\mathbf{p})+m\Big)S\psi(R^{-1}\mathbf{p})=S\Big(i\gamma\cdot\phi(R^{-1}\mathbf{p})+m\Big)\psi(R^{-1}\mathbf{p})=0 \tag{9.33}$$

Now we can define a representation of translations and rotations on $\mathcal{H}_0^+$. For $a=(a^0,\mathbf{a})$ and a spinor rotation S covering a rotation R, we define

$$(u_0(a,S)\psi)(\mathbf{p})=e^{i(\omega(\mathbf{p})a^0-\mathbf{p}\cdot\mathbf{a})}S\,\psi(R^{-1}\mathbf{p}) \tag{9.34}$$

One checks that $u_0(a,S)$ is unitary using the facts that the Lebesgue measure is rotation invariant and that $S^*S=I$. (For a rotation, $[S,\beta]=0$ so $S^*S=S^*\beta S\beta=\beta^2=I$.) One also checks that $u_0(a,S)u_0(a',S')=u_0(a+Ra',SS')$ so we have a representation.

A unitary representation on $\hat{\mathcal{H}}_0^+$ is defined by $\hat{u}_0(a,S)=\mathcal{F}^{-1}u_0(a,S)\mathcal{F}$. If $\psi(t,\mathbf{x})$ is the complete wave function given by (9.31), then

$$(\hat{u}_0(a,S)\psi(t,\cdot))(\mathbf{x})=S\psi(t-a^0,R^{-1}(\mathbf{x}-\mathbf{a})) \tag{9.35}$$

Furthermore the Newton–Wigner operator $X=(X^1,X^2,X^3)$ satisfies

$$u_0((0,\mathbf{a}),S)^{-1}Xu_0((0,\mathbf{a}),S)=RX+\mathbf{a} \tag{9.36}$$

This is a consistency check on the interpretation of both $u_0(a,S)$ and X.

Now specialize to the rotation $R_\theta=R(e_1,\theta)$ around the first axis. By problem 9.2 this is covered by $S_\theta=e^{-\gamma_2\gamma_3\theta/2}$ and so (9.34) becomes

$$(u_0(0,S_\theta)\psi)(\mathbf{p})=e^{-\gamma_2\gamma_3\theta/2}\psi(R_\theta^{-1}(\mathbf{p})) \tag{9.37}$$

The first component of angular momentum is $i(J_1\psi)(\mathbf{p})=id/d\theta[\cdots]|_{\theta=0}$ and we find as in (4.74)

$$J_1=p_3\left(i\frac{\partial}{\partial p_2}\right)-p_2\left(i\frac{\partial}{\partial p_3}\right)-\frac{i}{2}\gamma_2\gamma_3 \tag{9.38}$$

The first component of spin is the internal part $\Sigma_1=-i\gamma_2\gamma_3/2$. Similarly the other components are $\Sigma_2=-i\gamma_3\gamma_1/2$ and $\Sigma_3=-i\gamma_1\gamma_2/2$. In each case we have $\Sigma_i^2=1/4$ so Σ_i has eigenvalues $\pm1/2$, which confirms that we are describing a particle of spin 1/2.

Now we have operators on $\mathcal{H}_0^+$ for position, momentum, and angular momentum (but not spin alone, which does not act on this space). This is all we need. It is not required or particularly useful to try to give a meaning to the individual components of our spinor-valued wave function.

9.2.4 A covariant formulation

Next we give an alternative more covariant construction for the free electron. Our goal is a unitary representation of the *extended proper Poincaré group* defined as all pairs $\{a, S\}$ with $a \in \mathbb{R}^4$ and $S \in Spin^{\uparrow}(1,3)$ with the group law

$$\{a, S\}\{a', S'\} = \{a + \Lambda(S)a', SS'\} \tag{9.39}$$

We know that $u^\dagger \beta v$ is invariant under $Spin^{\uparrow}(1,3)$. We next show that it is postive definite on the positive energy subspace.

Lemma 9.3 *For* $u, v \in W_{\mathbf{p}}^+$

$$\omega(\mathbf{p})\, u^\dagger \beta v = m\, u^\dagger v \tag{9.40}$$

Proof Since $(\alpha \cdot \mathbf{p} + \beta m)v = \omega(\mathbf{p})v$

$$\omega(\mathbf{p})\, u^\dagger \beta v = u^\dagger \beta(\alpha \cdot \mathbf{p} + \beta m)v \tag{9.41}$$

and since $(\alpha \cdot \mathbf{p} + \beta m)u = \omega(\mathbf{p})u$

$$\begin{aligned} \omega(\mathbf{p})\, u^\dagger \beta v &= ((\alpha \cdot \mathbf{p} + \beta m)u)^\dagger \beta v \\ &= u^\dagger(\alpha \cdot \mathbf{p} + \beta m)\beta v \\ &= u^\dagger \beta(-\alpha \cdot \mathbf{p} + \beta m)v \end{aligned} \tag{9.42}$$

Adding these equations and using $\beta^2 = I$ gives the result. □

Now consider $L^2\left(\mathbb{R}^3, \mathbb{C}^4, m\,\omega(\mathbf{p})^{-1} d\mathbf{p}\right)$ as a candidate for the single particle space. Here we have used a slightly different normalization for the Lorentz invariant measure; compare the $d\mathbf{p}/(2\omega(\mathbf{p}))$ that we used in the scalar case. There is again a positive energy subspace defined by the condition $\psi(\mathbf{p}) \in W_{\mathbf{p}}^+$. On this positive energy subspace we introduce the relativistic inner product

$$\begin{aligned} (\psi, \chi)_\beta &= \int \psi(\mathbf{p})^\dagger \beta \chi(\mathbf{p}) \frac{m}{\omega(\mathbf{p})} d\mathbf{p} \\ &= \int \psi(\mathbf{p})^\dagger \chi(\mathbf{p}) \left(\frac{m}{\omega(\mathbf{p})}\right)^2 d\mathbf{p} \end{aligned} \tag{9.43}$$

This is positive definite and we let $\mathcal{H}^+$ be the completion in the associated norm. We have the identification

$$\mathcal{H}^+ = \{\psi \in L^2\left(\mathbb{R}^3, \mathbb{C}^4, m^2\omega(\mathbf{p})^{-2} d\mathbf{p}\right) : \psi(\mathbf{p}) \in W_{\mathbf{p}}^+\} \tag{9.44}$$

An alternative is to identify $\mathbb{R}^3$ with the mass shell V_m^+ via the function $\phi(\mathbf{p}) = (\omega(\mathbf{p}), \mathbf{p})$. As in section 8.1 the measure $m\omega(\mathbf{p})^{-1}d\mathbf{p}$ becomes the measure $2m\,\mu_m$ on V_m^+. The condition $u \in W_{\mathbf{p}}^+$ says $(i\gamma \cdot \phi(\mathbf{p}) + m)u = 0$ as noted earlier, and thus we can define $\mathcal{H}^+$ as the completion of

$$\{\psi \in L^2(V_m^+, \mathbb{C}^4, 2m\,\mu_m) : (i\gamma \cdot p + m)\psi(p) = 0\} \tag{9.45}$$

with the inner product

$$(\psi, \chi)_\beta = \int \psi(p)^\dagger \beta \chi(p)\, 2m\, d\mu_m(p) \tag{9.46}$$

Theorem 9.1 *On $\mathcal{H}^+$ there is a unitary representation of the extended proper Poincaré group defined by*

$$(u(a,S)\psi)(p) = e^{-ip\cdot a} S\psi(\Lambda(S)^{-1}p) \tag{9.47}$$

Proof To see that $u(a,S)$ maps $\mathcal{H}^+$ to itself, we compute as in (9.33)

$$\Big(i\gamma \cdot p + m\Big)S\psi(\Lambda(S)^{-1}p) = S\Big(i\gamma \cdot (\Lambda(S)^{-1}p) + m\Big)\psi(\Lambda(S)^{-1}p) = 0 \tag{9.48}$$

To see that $U(a,S)$ is unitary, use $(Su)^\dagger\beta(Sv) = u^\dagger\beta v$ from (9.7) and the Lorentz invariance of the measure to compute

$$\begin{aligned} \|u(a,S)\psi\|_\beta^2 &= \int \Big(S\psi(\Lambda(S)^{-1}p)\Big)^\dagger \beta \Big(S\psi(\Lambda(S)^{-1}p)\Big)\, 2m\, d\mu_m(p) \\ &= \int \psi(p)^\dagger \beta \psi(p)\, 2m\, d\mu_m(p) \\ &= \|\psi\|_\beta^2 \end{aligned} \tag{9.49}$$

One also checks that $u(a,S)u(a',S') = u(a + \Lambda(S)a', SS')$ to finish the proof. □

Remark Identify the new Hilbert space with the old Hilbert space by the unitary operator $V : \mathcal{H}_0^+ \to \mathcal{H}^+$ defined by

$$(V\psi)(\mathbf{p}) = \frac{\omega(\mathbf{p})}{m}\,\psi(\mathbf{p}) \tag{9.50}$$

With this identification the new representation of translations and rotations in (9.47) extends the old representation (9.34).

9.2.5 Charge conjugation

We close this section with some special results about charge conjugation, which we need in the next section. Let us temporarily return to the classical Dirac equation with an external electromagnetic potential A. This has the form (from (7.77))

$$(\gamma \cdot (\partial - ieA) + m)\psi = 0 \tag{9.51}$$

We want to find a *charge conjugation* operator which maps solutions of this equation with charge e to solutions with the opposite charge $-e$. Start by taking the complex conjugate which yields

$$(\overline{\gamma} \cdot (\partial + ieA) + m)\overline{\psi} = 0 \tag{9.52}$$

Now let C be a nonsingular matrix which satisfies $\mathsf{C}\overline{\gamma^\mu}\mathsf{C}^{-1} = \gamma^\mu$. Such a matrix exists since $\overline{\gamma^\mu}$ is again the possible choice of the gamma matrices. In fact we can choose $C = \overline{C} = C^{-1} = C^*$. For example in the representation (7.49) where γ^2 is real and the other γ^μ are imaginary we could take $\mathsf{C} = \gamma^2$. Applying C we get

$$(\gamma \cdot (\partial + ieA) + m)\mathsf{C}\overline{\psi} = 0 \tag{9.53}$$

Thus the charge conjugation operator we seek is

$$\mathcal{C}\psi = \mathsf{C}\,\overline{\psi} \tag{9.54}$$

It is anti-linear and satisfies $\mathcal{C}\gamma^\mu = \gamma^\mu\mathcal{C}$ and $\mathcal{C}^2 = 1$.

Return now to the free case $A = 0$ and the quantum interpretation. Then $\mathcal{C}$ defined on $L^2(\mathbb{R}^3, \mathbb{C}^4, d\mathbf{x})$ by (9.54) is an anti-unitary operator which is a conjugation in the mathematical sense that $(\mathcal{C}\psi, \mathcal{C}\chi) = (\chi, \psi)$. It satisfies $\mathcal{C}\alpha^i\mathcal{C} = \alpha^i$ and $\mathcal{C}\beta\mathcal{C} = -\beta$ and hence we have

$$\mathcal{C}(-i\nabla \cdot \alpha + \beta m)\mathcal{C} = -(-i\nabla \cdot \alpha + \beta m) \tag{9.55}$$

This shows that $\mathcal{C}$ maps positive energy states into negative energy states and vice-versa. Thus it does not act on our positive energy Hilbert space. However charge conjugation will find a place in the quantum field theory.

We also will need the momentum space version of (9.55). Let $\tilde{\mathcal{C}} = \mathcal{F}\mathcal{C}\mathcal{F}^{-1}$ be conjugation in $L^2(\mathbb{R}^3, \mathbb{C}^4, d\mathbf{p})$. Explicitly

$$(\tilde{\mathcal{C}}\psi)(\mathbf{p}) = \mathsf{C}\,\overline{\psi(-\mathbf{p})} \tag{9.56}$$

Then we have

$$\tilde{\mathcal{C}}[\mathbf{p} \cdot \alpha + \beta m]\tilde{\mathcal{C}} = -[\mathbf{p} \cdot \alpha + \beta m] \tag{9.57}$$

Problem 9.7 Let $H = H_D + \beta V = -i\nabla \cdot \alpha + \beta(m + V)$ where V is is a real function with $\|V\|_\infty < m$.

1. Show that H is self-adjoint on $D(H_D) \subset \hat{\mathcal{H}}_0$.
2. Show that zero is in the resolvent set for H.
3. Show that $\mathcal{C}H = -H\mathcal{C}$.
4. (hard) Let $P^\pm$ be the projection onto the positive and negative energy subspaces for H. Show that $\mathcal{C}P^\pm = P^\mp\mathcal{C}$.

9.3 Dirac fields

9.3.1 The problem

Now we want solutions of the Dirac equation which are quantum field operators. We seek Dirac field operators $\psi_\alpha(x) = \psi_\alpha(t, \mathbf{x})$ indexed by $x \in \mathbb{R}^4, \alpha = 1, \ldots, 4$ such that

$$\begin{aligned} i\frac{\partial \psi}{\partial t} &= (-i\nabla \cdot \alpha + \beta m)\psi \\ \{\psi_\alpha(0, \mathbf{x}), \psi^*_\beta(0, \mathbf{y})\} &= \delta(\mathbf{x} - \mathbf{y})\delta_{\alpha\beta} \end{aligned} \tag{9.58}$$

This is roughly analogous to what we did with the scalar field, except that now we use the anti-commutator instead of the commutator. This turns out to be necessary for a consistent theory. It means that the associated particles are fermions as we shall see.

These equations should be interpreted in the sense of distributions. We formally integrate with complex test functions $h \in \mathcal{S}(\mathbb{R}^3, \mathbb{C}^4)$ in space only and define

$$\begin{aligned} \psi(t, h) &= \int \sum_\alpha \psi_\alpha(t, \mathbf{x})\overline{h_\alpha(\mathbf{x})}d\mathbf{x} \\ \psi^*(t, h) &= \int \sum_\alpha \psi^*_\alpha(t, \mathbf{x})h_\alpha(\mathbf{x})d\mathbf{x} \end{aligned} \tag{9.59}$$

Thus we seek operators $\psi(t, h)$ anti-linear in h with adjoints $\psi^*(t, h)$ linear in h such that

$$\begin{aligned} i\frac{\partial}{\partial t}\psi(t, h) &= \psi(t, (-i\nabla \cdot \alpha + \beta m)h) \\ \{\psi(0, h), \psi^*(0, g)\} &= (h, g) \end{aligned} \tag{9.60}$$

where (h, g) is the $L^2(\mathbb{R}^3, \mathbb{C}^4, d\mathbf{x})$ inner product.

Example 9.1 Here is a construction that accomplishes these goals. With $\mathcal{H}_0 = L^2(\mathbb{R}^3, \mathbb{C}^4, d\mathbf{p})$ take the anti-symmetric Fock space $\mathcal{F}^-(\mathcal{H}_0)$. Define the fields in terms of the bounded creation and annihilation operators a, a^* by

$$\psi(h) = a(\tilde{h}) \qquad \psi^*(h) = a^*(\tilde{h}) \tag{9.61}$$

Then $\{\psi(h), \psi^*(g)\} = (\tilde{h}, \tilde{g}) = (h, g)$ as required. Time evolved field operators are defined by $\psi(t, h) = \psi(e^{iH_D t}h)$ where $H_D = -i\nabla \cdot \alpha + \beta m$ is the self-adjoint Hamiltonian defined in section 9.2.2. Then $i\partial/\partial t\, \psi(t, h) = \psi(t, H_D h)$ as required.

This construction is not satisfactory because of the presence of the negative energy particles. If we suppress them entirely, we do not get the commutator we want. An alternative picture is that all the negative energy states are present, but they

are already filled, thereby stabilizing the theory. (The idea invokes the exclusion principle again.) Because this sea of negative energy fermions is homogeneous, its presence is not manifest. It is possible however to have unoccupied states or holes in this negative energy sea. These holes propagate as if they have positive energy and the opposite charge. They are identified as the anti-particles. If the particles are electrons, the anti-particles are called positrons. The observed electron-positron annihilation into photons is then identified with an electron falling to a negative energy state, filling the hole and emitting a photon. The theory is known as *hole theory* and the negative energy background is called the *Dirac sea*. It is reminiscent of the Fermi sea encountered in section 6.4, but here it is a more radical concept since there is no lowest energy state.

We do not attempt to make direct mathematical sense of this picture. Instead take a standard shortcut and give a construction in which the anti-particles are introduced as elementary particles, just as for the charged scalar field. Our goal is again field operators $\psi(t,h)$ satisfying (9.60) but now such that time evolution is implemented with positive energy.

9.3.2 The field operator

Let $\mathcal{H}_0^+$ be the positive energy Hilbert space (9.27) and consider the fermion Fock space

$$\mathcal{F}_0 = \mathcal{F}^-(\mathcal{H}_0^+) \otimes \mathcal{F}^-(\mathcal{H}_0^+) \tag{9.62}$$

The first factor represents particles and the second factor represents anti-particles. We introduce annihilation operators for $h \in \mathcal{H}_0^+$ by

$$a(h) = a(h) \otimes I \qquad\qquad b(h) = (-1)^N \otimes a(h) \tag{9.63}$$

where N is the number operator. These satisfy

$$\{a(h), a^*(g)\} = (h,g) \qquad\qquad \{b(h), b^*(g)\} = (h,g) \tag{9.64}$$

with all other anti-commutators equal to zero. In particular we have

$$\{a(h), b(g)\} = 0 \tag{9.65}$$

This is the reason we have introduced the factor $(-1)^N$ in the definition; otherwise the commutator would be zero rather than the anti-commutator. This anti-commutator is a reflection of the deep connection between particles and anti-particles, which is natural in hole theory.

Now for $h \in \mathcal{S}(\mathbb{R}^3, \mathbb{C}^4)$ define the time zero field operator on $\mathcal{F}_0$ by

$$\psi(h) = a(P^+\tilde{h}) + b^*(\tilde{C}P^-\tilde{h}) \tag{9.66}$$

Here $P^{\pm}$ is the projection onto $\mathcal{H}_0^{\pm}$ as in (9.28) and $\tilde{\mathcal{C}}$ is charge conjugation in momentum space as in (9.56). Note that $\tilde{\mathcal{C}}P^- h$ does have positive energy since it follows from (9.57) that $\tilde{\mathcal{C}}P^{\pm} = P^{\mp}\tilde{\mathcal{C}}$. The field has the desired anti-commutator since

$$\begin{aligned}\{\psi(h), \psi^*(g)\} &= (P^+\tilde{h}, P^+\tilde{g}) + (\tilde{\mathcal{C}}P^-\tilde{g}, \tilde{\mathcal{C}}P^-\tilde{h}) \\ &= (P^+\tilde{h}, P^+\tilde{g}) + (P^-\tilde{h}, P^-\tilde{g}) \\ &= (\tilde{h}, \tilde{g}) = (h, g)\end{aligned} \tag{9.67}$$

To satisfy the field equation, time evolution is again defined by

$$\psi(t, h) = \psi(e^{iH_D t}h) \tag{9.68}$$

But $\mathcal{F}(e^{iH_D t}h) = e^{i\tilde{H}t}\tilde{h}$ and $P^+e^{i\tilde{H}t} = e^{i\omega t}P^+$ and $\tilde{\mathcal{C}}P^-e^{i\tilde{H}t} = \tilde{\mathcal{C}}P^-e^{-i\omega t} = e^{i\omega t}\tilde{\mathcal{C}}P^-$ and so we have

$$\psi(t, h) = a(e^{i\omega t}P^+\tilde{h}) + b^*(e^{i\omega t}\tilde{\mathcal{C}}P^-\tilde{h}) \tag{9.69}$$

In this form we can see that time evolution is generated by a Hamiltonian with positive energy. We collect this and other properties of the field in a theorem.

Theorem 9.2

1. *$\psi(t, h)$ satisfies the field equation and anti-commutation relations (9.60).*
2. *Time evolution is unitarily implemented by*

$$\psi(t, h) = e^{iHt}\psi(h)e^{-iHt} \tag{9.70}$$

where $e^{-iHt} = \Gamma(e^{-i\omega t}) \otimes \Gamma(e^{-i\omega t})$ and

$$H = d\Gamma(\omega) \otimes I + I \otimes d\Gamma(\omega) \geq 0 \tag{9.71}$$

3. *Gauge transformations are implemented by*

$$e^{-i\theta}\psi(t, h) = e^{iQ\theta}\psi(t, h)e^{-iQ\theta} \tag{9.72}$$

where $e^{-iQ\theta} = \Gamma(e^{-i\theta}) \otimes \Gamma(e^{i\theta})$ and the charge is

$$Q = N_a - N_b = N \otimes I - I \otimes N \tag{9.73}$$

The last point follows just as for the charged scalar field, see (8.86). As in problem 8.11, Q can be identified with $\int_{x^0=t} : \psi^*\psi :$ which is the quantization of (7.56). From (9.73) we see that each particle contributes charge $+1$ (in natural units), while each anti-particle contributes charge -1. These are the conventions for a positively charged fermion like a proton. For electrons replace Q by $-Q$ and change signs elsewhere.

Note the reversal from the construction of example 9.1. There Q is positive and H is not. Now H is positive and Q is not.

Problem 9.8 Consider the Dirac equation with a bounded real potential $V = V(\mathbf{x})$ satisfying $\|V\|_\infty < m$

$$i\frac{\partial \psi}{\partial t} = (-i\alpha \cdot \nabla + \beta(m + V))\psi = 0 \tag{9.74}$$

Using the results of problem 9.7 construct a distribution solution $\psi(t, h)$ with the canonical anti-commutator at $t = 0$ and with positive energy.

9.3.3 Locality

Let us also smear in space and time and define for $f \in \mathcal{S}(\mathbb{R}^4, \mathbb{C}^4)$

$$\psi(f) = \int \psi(t, f(t, \cdot))dt = a(P^+\Pi_+ f) + b^*(\tilde{C}P^-\Pi_- f) \tag{9.75}$$

where as in (8.67)

$$(\Pi_\pm f)(\mathbf{p}) = \sqrt{2\pi}\hat{f}(\pm\omega(\mathbf{p}), \mathbf{p}) \tag{9.76}$$

Note that $\psi(f)$ is anti-linear in f.

Then we compute

$$\begin{aligned}\{\psi(f), \psi^*(g)\} &= (P^+\Pi_+ f, P^+\Pi_+ g) + (\tilde{C}P^-\Pi_- g, \tilde{C}P^-\Pi_- f) \\ &= (P^+\Pi_+ f, P^+\Pi_+ g) + (\,P^-\Pi_- f, P^-\Pi_- g) \\ &= (P^+\Pi_+ f, \Pi_+ g) + (\,P^-\Pi_- f, \Pi_- g)\end{aligned} \tag{9.77}$$

Now use the identity

$$(P^\pm\Pi_\pm f)(\mathbf{p}) = \pm\frac{1}{2\omega(\mathbf{p})}(\Pi_\pm(i\partial_t + H_D)f)(\mathbf{p}) \tag{9.78}$$

where again $H_D = -i\alpha \cdot \nabla + \beta m$. Then (9.77) becomes

$$\left(\Pi_+(i\partial_t + H_D)f, \frac{1}{2\omega}\Pi_+ g\right) - \left(\Pi_-(i\partial_t + H_D)f, \frac{1}{2\omega}\Pi_- g\right) \tag{9.79}$$

However

$$\left(\Pi_+ f, \frac{1}{2\omega}\Pi_+ g\right) - \left(\Pi_- f, \frac{1}{2\omega}\Pi_- g\right) = \frac{1}{i}(f, Eg) \tag{9.80}$$

where E is the scalar propagator (7.91) and on the right side the inner product is in $L^2(\mathbb{R}^4, \mathbb{C}^4)$. Then we have

$$\{\psi(f), \psi^*(g)\} = \frac{1}{i}\big((i\partial_t + H_D)f, Eg\big) = \frac{1}{i}\big(f, (i\partial_t + H_D)Eg\big) \tag{9.81}$$

Since $(i\partial_t + H_D)$ does not enlarge supports, this vanishes if f, g have spacelike separated supports. This is our locality result.

9.3.4 A covariant formulation

Now we give an alternate construction in which the Lorentz covariance is manifest. Again let $\mathcal{H}^+$ be the positive energy relativistic Hilbert space defined in (9.43)–(9.46) and set

$$\mathcal{F} = \mathcal{F}^-(\mathcal{H}^+) \otimes \mathcal{F}^-(\mathcal{H}^+) \tag{9.82}$$

The unitary map $V : \mathcal{H}_0^+ \to \mathcal{H}^+$ defined in (9.50) induces a unitary map $\Gamma^2(V) = \Gamma(V) \otimes \Gamma(V)$ from $\mathcal{F}_0$ to $\mathcal{F}$. Thus starting with the field operator on $\mathcal{F}_0$ defined in (9.75) and now denoted $\psi_0(f)$ we define an equivalent field operator $\psi(f)$ on $\mathcal{F}$ by

$$\psi(f) = \Gamma^2(V)\psi_0(f)\Gamma^2(V)^{-1} \tag{9.83}$$

Then

$$\psi(f) = a(VP^+\Pi_+ f) + b^*(\tilde{\mathcal{C}}VP^-\Pi_- f) \tag{9.84}$$

where now the creation and annihilation operators are on $\mathcal{F}$. However by (9.78)

$$VP^\pm\Pi_\pm = \pm\frac{1}{2m}\Pi_\pm(i\partial_t + H_D) = \pm\Pi_\pm\Lambda\beta \tag{9.85}$$

where

$$\Lambda = \frac{-\gamma\cdot\partial + m}{2m} \tag{9.86}$$

Therefore

$$\psi(f) = a(\Pi_+\Lambda\beta f) - b^*(\tilde{\mathcal{C}}\Pi_-\Lambda\beta f) \tag{9.87}$$

These have the following properties:

Theorem 9.3 *The Dirac fields $\psi(f) = \int \psi(x)\overline{f(x)}dx$ satisfy in the sense of distributions:*

1. *(Field equation)* $(\gamma\cdot\partial + m)\psi = 0$
2. *(Covariance) There is a unitary representation $U(a,S)$ of the extended proper Poincaré group such that*

$$U(a,S)\psi(x)U(a,S)^{-1} = S^{-1}\ \psi(\Lambda(S)x + a) \tag{9.88}$$

3. *(Locality) With $\tilde{\psi}(f) = \psi^*(\beta f)$ (the "Dirac adjoint") and scalar propagator E*

$$\{\psi(f), \tilde{\psi}(g)\} = \frac{1}{i}(f, (-\gamma\cdot\partial + m)Eg) \tag{9.89}$$

Proof These operations translate nicely to the test function if we consider $\psi(\beta f)$ rather than $\psi(f)$. For the field equation $((\gamma\cdot\partial + m)\psi)(\beta f) = 0$ says that

$\psi(\beta(\gamma\cdot\partial+m)f)=0$. (Think of Green's identity (7.55) at times $\pm\infty$.) We compute using $\Pi_\pm(-\Box+m^2)f=0$

$$\begin{aligned}&\psi(\beta(\gamma\cdot\partial+m)f)\\&=a(\Pi_+\Lambda(\gamma\cdot\partial+m)f)-b^*(\tilde{\mathcal{C}}\Pi_-\Lambda(\gamma\cdot\partial+m)f)\\&=\Big(a(\Pi_+(-\Box+m^2)f)-b^*(\tilde{\mathcal{C}}\Pi_-(-\Box+m^2)f)\Big)/2m\\&=0\end{aligned}\tag{9.90}$$

For the second we set $U(a,S)=\Gamma(u(a,S))$ where $u(a,S)$ is defined in (9.47). Then with $f_{a,S}(x)=Sf(\Lambda(S)^{-1}(x-a))$ we have

$$\begin{aligned}U(a,S)\psi(\beta f)U(a,S)^{-1}&=a(u(a,S)\Pi_+\Lambda f)-b^*(u(a,S)\tilde{\mathcal{C}}\Pi_-\Lambda f)\\&=a(\Pi_+\Lambda f_{a,S})-b^*(\tilde{\mathcal{C}}\Pi_-\Lambda f_{a,S})\\&=\psi(\beta f_{a,S})\end{aligned}\tag{9.91}$$

This is the meaning of (9.88). We have also used $[\tilde{\mathcal{C}},u(a,S)]=0$.

For the third point refer back to the result (9.81) for $\psi_0(f)$. Then $\psi(f)$ must have the same anti-commutator and so

$$\{\psi(f),\tilde{\psi}(g)\}=\frac{1}{i}(f,(i\partial_t+H_D)\beta Eg)=\frac{1}{i}(f,(-\gamma\cdot\partial+m)Eg)\tag{9.92}$$

□

9.4 Photons

Recall that in the absence of charges the electromagnetic potential A satisfies the wave equation $\Box A=0$ and a constraint $\partial^\mu A_\mu=0$. Following the scalar case we want to interpret positive energy solutions as wave functions for a massless spin one particle – the photon. Again there is more than one way to accomplish this.

9.4.1 Coulomb gauge

Our first quantization starts with the observation that potentials A,A' which are gauge equivalent $A'_\mu=A_\mu+\partial_\mu\lambda$ represent the same physical situation. Any potential A is gauge equivalent to a potential A' with $A'_0=0$; one has only to take $\lambda(t,\mathbf{x})=-\int_0^t A_0(s,\mathbf{x})ds$. Thus it suffices to consider solutions with constraint $A_0=0$ and $\mathbf{A}=(A_1,A_2,A_3)$. Making this choice we are working in the *Coulomb gauge*.

In the Coulomb gauge the equations for $\mathbf{A}(t, \mathbf{x})$ become

$$\left(\frac{\partial^2}{\partial t^2} - \Delta\right)\mathbf{A} = 0 \qquad \nabla \cdot \mathbf{A} = 0 \tag{9.93}$$

The partial Fourier transform $\tilde{\mathbf{A}}(t, \mathbf{p})$ satisfies

$$\left(\frac{\partial^2}{\partial t^2} + |\mathbf{p}|^2\right)\tilde{\mathbf{A}} = 0 \qquad \mathbf{p} \cdot \tilde{\mathbf{A}} = 0 \tag{9.94}$$

For quantization we consider complex square integrable positive energy solutions. In the Fourier transform variable, identified as momentum, the Hilbert space is the closed subspace

$$\mathcal{H}_0 = \{\Psi \in L^2(\mathbb{R}^3, \mathbb{C}^3, d\mathbf{p}) : \mathbf{p} \cdot \Psi(\mathbf{p}) = 0\} \tag{9.95}$$

The time evolution is $\Psi(t, \mathbf{p}) = e^{-i|\mathbf{p}|t}\Psi(\mathbf{p})$. It does preserve the constraint $\mathbf{p} \cdot \Psi(\mathbf{p}) = 0$ and is unitary on the Hilbert space. The Hamiltonian is $[|\mathbf{p}|]$. In configuration space the Hilbert space is $\hat{\mathcal{H}}_0 = \mathcal{F}^{-1}\mathcal{H}_0$, the time evolution is

$$U(t) = \mathcal{F}^{-1}[e^{-i|\mathbf{p}|t}]\mathcal{F} \tag{9.96}$$

and the Hamiltonian is

$$H = \mathcal{F}^{-1}[|\mathbf{p}|]\mathcal{F} = (-\Delta)^{1/2} \tag{9.97}$$

We define a representation of spacetime translations and rotations on $\mathcal{H}_0$ by

$$(u_0(a, R)\Psi)(\mathbf{p}) = e^{i(|\mathbf{p}|a^0 - \mathbf{p}\cdot\mathbf{a})}R\Psi(R^{-1}\mathbf{p}) \tag{9.98}$$

This preserves the constraint since if $\mathbf{p} \cdot \Psi(\mathbf{p}) = 0$, then $\mathbf{p} \cdot R\Psi(R^{-1}\mathbf{p}) = R^{-1}\mathbf{p} \cdot \Psi(R^{-1}\mathbf{p}) = 0$. It is unitary since $R^T R = I$. On $\hat{\mathcal{H}}_0$ the representation is

$$\hat{u}_0(a, R) = \mathcal{F}^{-1}\, u_0(a, R)\mathcal{F} \tag{9.99}$$

The definitions implement these symmetries since if $\Psi(t, \mathbf{x}) = (U(t)\Psi)(\mathbf{x})$ is a complete time evolved wave function, then

$$(\hat{u}_0(a, R)\Psi(t, \cdot))(\mathbf{x}) = R\,\Psi(t - a_0, R^{-1}(\mathbf{x} - \mathbf{a})) \tag{9.100}$$

The representation of the rotation group is characteristic of a spin one particle.

However we have again lost the position operator. Multiplication by x^k does not preserve the Fourier transformed space $\hat{\mathcal{H}}_0$. In this case the difficulty cannot be avoided. There is no Newton–Wigner operator and photons cannot be precisely localized.

Problem 9.9

1. Show that $\mathcal{H}_0$ is a closed subspace of $L^2(\mathbb{R}^3, \mathbb{C}^3, d\mathbf{p})$.

2. Show that the projection onto $\mathcal{H}_0$ is

$$(P\Psi)_i(\mathbf{p}) = \sum_j \left(\delta_{ij} - \frac{p_i p_j}{|\mathbf{p}|^2}\right) \Psi_j(\mathbf{p}) \tag{9.101}$$

9.4.2 A covariant formulation

Next we give a more covariant construction with the zeroth component restored. Start with the light cone

$$V_0^+ = \{p \in \mathbb{R}^4 : p \cdot p = 0, p^0 > 0\} \tag{9.102}$$

Consider the Hilbert space

$$\mathcal{H} = L^2(V_0^+, \mathbb{C}^4, \mu_0) \tag{9.103}$$

where μ_0 is the invariant measure on V_0^+. Let $\mathcal{H}'$ be the closed subspace

$$\mathcal{H}' = \{\psi \in \mathcal{H} : p^\mu \psi_\mu(p) = 0\} \tag{9.104}$$

We want a representation of the Poincaré group on this space consistent with (9.98). Our first choice would be to take on $\mathcal{H}$

$$(u(a,\Lambda)\psi)_\mu(p) = e^{-ip\cdot a}(\Lambda^{-1})^\nu{}_\mu \psi_\nu(\Lambda^{-1}p) \tag{9.105}$$

which can also be written $(u(a,\Lambda)\psi)(p) = e^{-ip\cdot a}(\Lambda^{-1})^T\psi(\Lambda^{-1}p)$. This preserves $\mathcal{H}'$ since if $p^\mu\psi_\mu(p) = 0$ for all p, then

$$p^\mu(\Lambda^{-1})^\nu{}_\mu \psi_\nu(\Lambda^{-1}p) = (\Lambda^{-1}p)^\nu \psi_\nu(\Lambda^{-1}p) = 0 \tag{9.106}$$

If $\Lambda = R$ is a rotation, then $(R^{-1})^T = R$ and the representation is unitary. However for boosts the representation is not unitary since $\Lambda^{-1}(\Lambda^{-1})^T = I$ fails.

To fix this we introduce an indefinite inner product on $L^2(V_0^+, \mathbb{C}^4, \mu_0)$ by

$$(\psi, \chi)_\eta = \int_{V_0^+} \overline{\psi(p)} \cdot \chi(p)\, d\mu_0(p) = \int_{V_0^+} \overline{\psi_\mu(p)}\eta^{\mu\nu}\chi_\nu(p)\, d\mu_0(p) \tag{9.107}$$

This is well-defined and since $\Lambda^{-1}\eta(\Lambda^{-1})^T = \eta$ we do have the invariance

$$(u(a,\Lambda)\psi, u(a,\Lambda)\chi)_\eta = (\psi,\chi)_\eta \tag{9.108}$$

But if we adopt the indefinite inner product, we no longer have a Hilbert space. We proceed as follows. Consider the closed subspace

$$\mathcal{H}'' = \{\psi \in \mathcal{H}' : \psi_\mu(p) = p_\mu f(p)\} \tag{9.109}$$

where f is some measurable function from V_0^+ to $\mathbb{C}$. These are null vectors: $(\psi,\psi)_\eta = 0$ for $\psi \in \mathcal{H}''$. Then consider the factor space $\mathcal{H}'/\mathcal{H}''$. Elements of this space are equivalence classes $[\psi]$ of vectors in $\mathcal{H}'$ with $\psi \sim \psi'$ if $\psi - \psi' \in \mathcal{H}''$.

Theorem 9.4

1. *The indefinite inner product is well-defined and positive definite on the factor space $\mathcal{H}'/\mathcal{H}''$ and with this inner product it is a Hilbert space denoted*

$$\mathcal{H}_{phys} = \mathcal{H}'/\mathcal{H}'' \tag{9.110}$$

2. *The operators $u(a,\Lambda)$ defined by (9.105) on $\mathcal{H}'$ determine unitary operators $u(a,\Lambda)$ on $\mathcal{H}_{phys}$, which give an (irreducible) representation of the proper Poincaré group.*

Proof With the usual L^2 inner product let $\mathcal{M}$ be the orthogonal complement of $\mathcal{H}''$ in $\mathcal{H}'$ so that $\mathcal{H}' = \mathcal{M} \oplus \mathcal{H}''$. The projection from $\mathcal{H}'$ onto $\mathcal{M}$ has kernel $\mathcal{H}''$ and so gives an identification of $\mathcal{H}'/\mathcal{H}''$ and $\mathcal{M}$ as vector spaces. The inverse sends $\psi \in \mathcal{M}$ to $[\psi] \in \mathcal{H}'/\mathcal{H}''$.

To identify $\mathcal{M}$, note that the condition that $\psi = (\psi_0, \Psi) \in \mathcal{H}'$ be orthogonal to $\mathcal{H}''$ is the condition that $p_\mu \psi_\mu(p) = 0$. Since $p_0 = -p^0 = -|\mathbf{p}|$, this says $-|\mathbf{p}|\psi_0(p) + \mathbf{p}\cdot\Psi(p) = 0$. On the other hand because $\psi \in \mathcal{H}'$ we have $p^\mu \psi_\mu(p) = 0$ or $|\mathbf{p}|\psi_0(p) + \mathbf{p}\cdot\Psi(p) = 0$. Together these imply that $\psi_0(p) = 0$ and $\mathbf{p}\cdot\Psi(p) = 0$. Thus we have

$$\mathcal{M} = \{\psi = (\psi_0, \Psi) \in \mathcal{H}' : \psi_0 = 0,\ \mathbf{p}\cdot\Psi(p) = 0\} \tag{9.111}$$

Now consider the indefinite inner product. If ψ' is in $\mathcal{H}'$ and $\psi''_\mu(p) = p_\mu f(p)$ is in $\mathcal{H}''$, then

$$(\psi', \psi'')_\eta = \int_{V_0^+} \overline{p^\mu \psi'_\mu(p)} f(p) d\mu_0(p) = 0 \tag{9.112}$$

It follows that the indefinite inner product is well defined on $\mathcal{H}'/\mathcal{H}''$ and for $\psi_1, \psi_2 \in \mathcal{H}'$ we have $(\psi_1, \psi_2)_\eta = ([\psi_1], [\psi_2])_\eta$. So the map from $\mathcal{M}$ to $\mathcal{H}'/\mathcal{H}''$ preserves the indefinite inner product. But the indefinite inner product on $\mathcal{M}$ coincides with the L^2 inner product since there is no zeroth component and hence it is positive definite. Hence the indefinite inner product is positive definite on $\mathcal{H}'/\mathcal{H}''$. Furthermore $\mathcal{M}$ is complete with the indefinite inner product since $\mathcal{M}$ is a closed subspace of L^2. Hence the same is true for $\mathcal{H}'/\mathcal{H}''$ and it is a Hilbert space. Thus we have the first result and have identified $\mathcal{H}_{phys}$ and $\mathcal{M}$ as Hilbert spaces.

We have already noted that $u(a,\Lambda)$ preserves $\mathcal{H}'$. It also preserves $\mathcal{H}''$ since if $\psi_\nu(p) = p_\nu f(p) = \eta_{\nu\sigma} p^\sigma f(p)$ is in $\mathcal{H}'$, then

$$\begin{aligned}(\Lambda^{-1})^\nu{}_\mu \psi_\nu(\Lambda^{-1}p) &= (\Lambda^{-1})^\nu{}_\mu \eta_{\nu\sigma} (\Lambda^{-1})^\sigma{}_\rho p^\rho f(\Lambda^{-1}p) \\ &= \eta_{\mu\rho} p^\rho f(\Lambda^{-1}p) = p_\mu f(\Lambda^{-1}p)\end{aligned} \tag{9.113}$$

is again in $\mathcal{H}''$. It follows that we have a representation on $\mathcal{H}_{phys}$, and since it preserves the inner product, it is unitary. □

Remarks

1. $\mathcal{H}_{phys}$ is the basic Hilbert space, called the physical Hilbert space. This procedure of going to the factor space is known as *Gupta–Bleuler* quantization.

2. When we form the factor space, equivalence classes are formed with $\psi \sim \psi'$ if $\psi_\mu = \psi'_\mu + p_\mu f$. Returning to spacetime with an inverse Fourier transform this says that ψ and ψ' are related by a gauge transformation. Thus the mathematical equivalence in the construction of the Hilbert space is mirrored by physical equivalence in the sense that equivalent vectors have the same field strength.
3. The identification of $\mathcal{H}_{phys}$ with $\mathcal{M}$ also gives the equivalence with the quantization in the Coulomb gauge on $\mathcal{H}_0$. First identify $\mathbb{R}^3$ with V_0^+ by $\phi(\mathbf{p}) = (|\mathbf{p}|, \mathbf{p})$. Then define a unitary map $V : \mathcal{H}_0 \to \mathcal{M}$ given by

$$(V\Psi)(\phi(\mathbf{p})) = \sqrt{2|\mathbf{p}|}\Big(0, \Psi(\mathbf{p})\Big) \tag{9.114}$$

One checks that for spacetime translations and rotations

$$V^{-1}u(a,R)V = u_0(a,R) \tag{9.115}$$

which establishes the equivalence.

9.5 Electromagnetic field

Now we undertake the quantization of the electromagnetic field, again as represented by a potential A satisfying the wave equation $\Box A = 0$ and a constraint $\partial^\mu A_\mu = 0$. We skip a Coulomb gauge construction and proceed with a covariant construction. The strategy is to first ignore the constraint, carry out the quantization, and then impose the constraint after quantization.

The quantization of $\Box A = 0$ proceeds as in the scalar case except that now we use the indefinite inner product. Start with the single particle Hilbert space $\mathcal{H} = L^2(V_0^+, \mathbb{C}^4, \mu_0)$ and form the symmetric Fock space $\mathcal{F} = \mathcal{F}^+(\mathcal{H})$. The indefinite inner product $(h,g)_\eta = (h, \eta g)$ on $\mathcal{H}$ induces an indefinite inner product $(\psi, \chi)_\eta = (\psi, \Gamma(\eta)\chi)$ on $\mathcal{F}$. If we then define creation and annihilation operators $a^\dagger(h), a(h)$ as in (5.54), but now with the indefinite inner product, we find that on $\mathcal{D}_0$

$$\begin{aligned} [a(h), a^\dagger(h')] &= (h, h')_\eta \\ (a^\dagger(h)\psi, \chi)_\eta &= (\psi, a(h)\chi)_\eta \end{aligned} \tag{9.116}$$

So $a^\dagger$ is the formal adjoint with respect to the indefinite inner product.

Now as in (8.66) we define field operator $A(f) = \int A_\mu(x)\eta^{\mu\nu}f_\nu(x)\,dx$ by

$$A(f) = a^\dagger(\Pi_+ f) + a(\Pi_+ f) \tag{9.117}$$

where $\Pi_+ f \in \mathcal{H}$ is defined by

$$(\Pi_+ f)_\mu(p) = \sqrt{2\pi}\hat{f}_\mu(p) \tag{9.118}$$

We summarize the results for this field.

Theorem 9.5 *For real $f \in \mathcal{S}(\mathbb{R}^4, \mathbb{R}^4)$ let $A(f)$ be the symmetric operator valued distribution defined on $(\mathcal{F}, (\cdot,\cdot)_\eta)$ by (9.117). Then*

1. *(Field equation)* $(-\Box + m^2)A = 0$.
2. *(Covariance) There is a representation $U(a,\Lambda)$ of the proper Poincaré group preserving the indefinite inner product $(\cdot,\cdot)_\eta$ such that*

$$U(a,\Lambda)A_\mu(x)U(a,\Lambda)^{-1} = \Lambda^\nu{}_\mu A_\nu(\Lambda x + a) \tag{9.119}$$

3. *(Locality)*

$$[A(f), A(g)] = \frac{1}{i}\eta^{\mu\nu} < f_\mu, Eg_\nu > \tag{9.120}$$

The representation of the Poincaré group is $U(a,\Lambda) = \Gamma(u(a,\Lambda))$ where $u(a,\Lambda)$ is defined in (9.105) and the covariance (9.119) is equivalent to

$$U(a,\Lambda)A(f)U(a,\Lambda)^{-1} = A(f_{a,\Lambda}) \tag{9.121}$$

where $(f_{a,\Lambda})_\mu(x) = (\Lambda^{-1})^\nu{}_\mu f_\nu(\Lambda^{-1}(x-a))$. All statements can be checked as for the scalar field, Theorem 8.2.

Now we impose the constraint $\partial^\mu A_\mu = 0$ not as an operator identity but on wavefunctions by passing to the subspace

$$\mathcal{F}' \equiv \mathcal{F}^+(\mathcal{H}') \tag{9.122}$$

where $\mathcal{H}'$ defined in (9.104) has $p^\mu \psi_\mu(p) = 0$. The indefinite inner product is nonnegative on $\mathcal{F}'$. We consider the null subspace

$$\mathcal{F}'' = \{\psi \in \mathcal{F}' : (\psi,\psi)_\eta = 0\} \tag{9.123}$$

By the Schwarz inequality, elements of $\mathcal{F}''$ are orthogonal to any element of $\mathcal{F}'$. Then the indefinite inner product is well-defined on $\mathcal{F}'/\mathcal{F}''$ with $([\psi_1],[\psi_2])_\eta = (\psi_1,\psi_2)_\eta$. Furthermore it is positive definite. The physical Hilbert space is the completion

$$\mathcal{F}_{phys} = \overline{\mathcal{F}'/\mathcal{F}''} \tag{9.124}$$

We will see later that the completion is unnecessary.

Theorem 9.6

1. *The representation $U(a,\Lambda)$ on $\mathcal{F}$ determines a unitary representation of the Poincaré group on $\mathcal{F}_{phys}$.*
2. *If $\partial^\mu f_\mu = 0$, then the field operator $A(f)$ acts on a dense domain in $\mathcal{F}_{phys}$.*

Proof $U(a,\Lambda)$ preserves $\mathcal{F}'$ with its inner product and $\mathcal{F}''$ and hence determines an inner product preserving operator on $\mathcal{F}'/\mathcal{F}''$ which extends to a unitary on $\mathcal{F}_{phys}$.

If $\partial^\mu f_\mu = 0$, then

$$p^\mu(\Pi_+ f_\mu)(p) = (\Pi_+(-i\partial^\mu f_\mu))(p) = 0 \tag{9.125}$$

so $\Pi_+ f \in \mathcal{H}'$. Hence $A(f)$ preserves finite particle vectors in $\mathcal{F}'$. Furthermore it preserves finite particle vectors in $\mathcal{F}''$ since if $(\psi,\psi)_\eta = 0$, then

$$\begin{aligned}(A(f)\psi, A(f)\psi)_\eta &= (\psi, A(f)^2\psi)_\eta \\ &\leq (\psi,\psi)_\eta^{1/2}(A(f)^2\psi, A(f)^2\psi)_\eta^{1/2} \\ &= 0\end{aligned} \tag{9.126}$$

Thus $A(f)$ gives an operator on finite particle vectors in $\mathcal{F}'/\mathcal{F}''$. □

Remark Without the restriction $\partial^\mu f_\mu = 0$, field operators $A(f)$ do not act on $\mathcal{F}_{phys}$. However the field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ does act on $\mathcal{F}_{phys}$. For a family of test functions $h_{\mu\nu}$ we have $F(h) = A(\delta h)$ where $(\delta h)_\mu = \partial^\nu(h_{\mu\nu} - h_{\nu\mu})$ and this does satisfy $\partial^\mu(\delta h)_\mu = 0$.

We also can identify our physical Hilbert space $\mathcal{F}_{phys}$ as a multi-photon Fock space.

Theorem 9.7 *There is a natural identification*

$$\mathcal{F}_{phys} \approx \mathcal{F}^+(\mathcal{H}_{phys}) \tag{9.127}$$

where $\mathcal{H}_{phys} = \mathcal{H}'/\mathcal{H}''$ is the single photon Hilbert space.

Proof First work with the usual inner product. Since $\mathcal{H}' = \mathcal{M} \oplus \mathcal{H}''$ we have the identification of Fock spaces (see problem 9.10)

$$\mathcal{F}' = \mathcal{F}^+(\mathcal{H}') \approx \mathcal{F}^+(\mathcal{M}) \otimes \mathcal{F}^+(\mathcal{H}'') \tag{9.128}$$

Separate off the vacuum component and write $\mathcal{F}^+(\mathcal{H}'') = \mathbb{C} \oplus \mathcal{F}^+_{\geq 1}(\mathcal{H}'')$ and make the identification $\mathbb{C} \otimes \mathcal{F}^+(\mathcal{M}) \approx \mathcal{F}^+(\mathcal{M})$ and we have

$$\mathcal{F}' \approx \mathcal{F}^+(\mathcal{M}) \oplus (\mathcal{F}^+(\mathcal{M}) \otimes \mathcal{F}^+_{\geq 1}(\mathcal{H}'')) \tag{9.129}$$

Every element of $\mathcal{F}^+(\mathcal{M}) \otimes \mathcal{F}^+_{\geq 1}(\mathcal{H}'')$ is in $\mathcal{F}''$ since all entries have at least one element in $\mathcal{H}''$. No element of $\mathcal{F}^+(\mathcal{M})$ is in $\mathcal{F}''$. We conclude that $\mathcal{F}'' = \mathcal{F}^+(\mathcal{M}) \otimes \mathcal{F}^+_{\geq 1}(\mathcal{H}'')$ and have the identification

$$\mathcal{F}' \approx \mathcal{F}^+(\mathcal{M}) \oplus \mathcal{F}'' \tag{9.130}$$

It follows that $\mathcal{F}^+(\mathcal{M}) \approx \mathcal{F}'/\mathcal{F}''$ as vector spaces under the map $\psi \to [\psi]$. The identification preserves the indefinite inner product as we have seen. Thus $\mathcal{F}'/\mathcal{F}''$ with the indefinite inner product inherits completeness from $\mathcal{F}^+(\mathcal{M})$, the completion in (9.124) was unnecessary, and $\mathcal{F}^+(\mathcal{M}) \approx \mathcal{F}'/\mathcal{F}'' = \mathcal{F}_{phys}$ as Hilbert spaces. Since also $\mathcal{F}^+(\mathcal{M}) \approx \mathcal{F}(\mathcal{H}_{phys})$ we have the result. □

Problem 9.10 Show that if a Hilbert space splits as $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$, then there is a unitary operator $V : \mathcal{F}^+(\mathcal{H}) \mapsto \mathcal{F}^+(\mathcal{H}_1) \otimes \mathcal{F}^+(\mathcal{H}_2)$ such that

1. If $U = U_1 \oplus U_2$ is unitary on $\mathcal{H}$, then
$$V\Gamma(U)V^{-1} = \Gamma(U_1) \otimes \Gamma(U_2)$$
2. If $h = (h_1, h_2)$, then
$$Va(h)V^{-1} = a(h_1) \otimes I + I \otimes a(h_2)$$

Problem 9.11 Show that $U(a, \Lambda)F_{\mu\nu}(x)U(a, \Lambda)^{-1} = \Lambda^{\sigma}{}_{\mu}\Lambda^{\rho}{}_{\nu}F_{\sigma\rho}(\Lambda x + a)$.

Remarks For both electrons and photons we have two kinds of fields. The fields ψ, A are basic for the construction, but are not self-adjoint operators on a Hilbert space and are not regarded as observables. From them one can construct current densities $j^{\mu} = i\tilde{\psi}\gamma^{\mu}\psi$ (we only did $Q = \int j^0$) and field strengths $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ which are self-adjoint operators on a Hilbert space and are regarded as observables.

The natural next step would be to quantize the combined system of electrons and photons as described by the equations (7.77). This is known as quantum electrodynamics. There has been very little mathematical progress on this problem. There are expressions for scattering amplitudes given as formal expansions in the charge e. The coefficients in these expansions are given by formal integrals, some of which are badly divergent. Nevertheless physicists have developed consistent methods for interpreting these integrals known as *renormalization*. When renormalized, the scattering amplitudes for the interactions of electrons and photons agree with experiments with an accuracy of up to 11 significant figures. This spectacular agreement is one reason that quantum field theory continues to be a fascinating subject.

Notes on chapter 9: For a discussion of the groups $O(n, m)$, $\mathrm{Spin}(n, m)$, see Choquet-Bruhat and DeWitt-Morette (1989).

The universal covering group of $\mathcal{L}_+^{\uparrow}$ can be identified as $SL(2, \mathbb{C})$ the complex 2×2 matrices with determinant 1. Our group $Spin^{\uparrow}(1, 3)$ is the so-called $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ reducible representation of $SL(2, \mathbb{C})$.

The general treatment of irreducible representations of the Poincaré group was originally due to Wigner and can be found in many places, for example Ohnuki (1988).

For Newton–Wigner coordinates see Newton and Wigner (1949).

There is a fairly large literature on quantum fields with external potentials, but no definitive reference. The examples of problem 8.10 and problem 9.8 were chosen more for simplicity than physical relevance.

For a discussion of the quantization of the electromagnetic field in various gauges, see Stocchi and Wightman (1974).

10 Field theory on a manifold

In this chapter we construct a scalar quantum field operator on a general class of manifolds.[1] This is the mathematics appropriate for modeling the propagation of quantum particles and fields in a gravitational field. It is particularly relevant for extreme situations like black holes and the early universe where quantum effects are important. However we confine the treatment to basic constructions.

10.1 Lorentzian manifolds

A Lorentzian manifold is a pair $(\mathcal{M}, g)$ consisting of a d-dimensional manifold $\mathcal{M}$ and a *Lorentzian metric* g. The latter is a symmetric nondegenerate 2-tensor with signature $(-, +, \ldots, +)$. This means that for each point $p \in \mathcal{M}$ there is a symmetric nondegenerate bilinear form $g_p(v, w)$ on the tangent space $\mathcal{M}_p$ such that the associated matrix has one negative and $d - 1$ positive eigenvalues. Local coordinates $\{x^\mu\}$ give a basis $\{\partial/\partial x^\mu\}$ for $\mathcal{M}_p$ and the matrix in this basis is

$$g_{\mu\nu}(p) = g_p(\partial/\partial x^\mu, \partial/\partial x^\nu) \tag{10.1}$$

A general tangent vector $v \in \mathcal{M}_p$ is written $v = v^\mu \partial/\partial x^\mu$ (summation convention) and then

$$g_p(v, v) = g_{\mu\nu}(p) v^\mu v^\nu \tag{10.2}$$

In terms of the dual basis $\{dx^\mu\}$ for the cotangent space we have

$$g_p = g_{\mu\nu}(p) dx^\mu dx^\nu \tag{10.3}$$

Minkowski space is the special case $\mathcal{M} = \mathbb{R}^4, g_{\mu\nu} = \eta_{\mu\nu}$.

We classify the tangent vectors according to the sign of $g_p(v, v)$. The vector $v \in \mathcal{M}_p$ is *spacelike* if $g_p(v, v) > 0$, it is *lightlike* if $g_p(v, v) = 0$, and it is *timelike* if $g_p(v, v) < 0$. The timelike vectors form a cone with two components. One component is designated as future directed and one component is designated as

[1] This chapter assumes more knowledge about manifolds than elsewhere in the book. In particular we assume the reader is familiar with the definitions and basic properties of manifolds, tensors, metrics, etc. This chapter is not referred to elsewhere and can be skipped.

past directed. We assume that this can be done in a continuous way over the whole manifold, that is the manifold has a *time orientation*. There is an associated splitting of lightlike vectors.

A curve $x : [a, b] \to \mathcal{M}$ is *timelike* if the tangent vectors $dx/d\tau = x_*(d/d\tau)$ are timelike at all points on the curve. Similarly we define lightlike and spacelike curves. A curve is *causal* if its tangent vectors are either timelike or lightlike. A causal curve is *future directed* if the tangent vectors are all future directed. Future directed causal curves are the possible world lines of particles. If x is future directed and timelike, the elapsed proper time is defined as

$$\int_a^b \sqrt{-g_{x(\tau)}\Big(\frac{dx}{d\tau}, \frac{dx}{d\tau}\Big)}\, d\tau \tag{10.4}$$

Timelike curves with fixed endpoints minimizing proper time are geodesics. These are the trajectories of freely falling objects.

For any point $p \in \mathcal{M}$ define the future and past of the point by

$$J^{\pm}(p) = \{q \in \mathcal{M} : \exists \text{ a future/past causal curve from p to q}\} \tag{10.5}$$

We also define $J^{\pm}(A) = \cup_{p\in A} J^{\pm}(p)$ for any $A \subset \mathcal{M}$.

A hypersurface Σ is a submanifold of dimension $d - 1$. For $p \in \Sigma$ the tangent space Σ_p is identified as a $d - 1$-dimensional subspace of $\mathcal{M}_p$. There is also a one-dimensional subspace of normal vectors N defined by the condition $g_p(N, v) = 0$ for all $v \in \Sigma_p$. The hypersurface Σ is spacelike if the nonzero tangent vectors Σ_p are spacelike for all $p \in \Sigma$. It is equivalent to say that the nonzero normal vectors N are timelike.

A spacelike hypersurface is a *Cauchy surface* if every endless causal curve intersects it exactly once.

Theorem 10.1 *The following conditions on a Lorentzian manifold* $(\mathcal{M}, g)$ *are equivalent.*

1. *There exist no closed causal curves*[2] *and the set* $J^+(p) \cap J^-(q)$ *is compact for all* $p, q \in \mathcal{M}$.
2. *There is Cauchy surface* Σ.
3. $(\mathcal{M}, g)$ *is diffeomorphic to a manifold* $(\mathbb{R} \times \Sigma, g')$ *for which* $\Sigma_t = \{t\} \times \Sigma$ *is a Cauchy surface for all* t.

If these conditions hold, we say that $(\mathcal{M}, g)$ is *globally hyperbolic*.

[2] This includes a prohibition on "almost closed" causal curves; see Bär Ginoux and Pfäffle (2007) for the definition.

10.2 Classical fields on a manifold

We study the Klein–Gordon equation on a globally hyperbolic Lorentzian manifold $(\mathcal{M}, g)$. It has the form

$$(-\Box + m^2)u = 0 \tag{10.6}$$

where in local coordinates

$$\Box u = |\det g|^{-1/2} \frac{\partial}{\partial x_\mu} \left(|\det g|^{1/2} g^{\mu\nu} \frac{\partial u}{\partial x_\nu} \right) \tag{10.7}$$

and $\{g^{\mu\nu}\}$ is the inverse matrix to $\{g_{\mu\nu}\}$.

Let Σ be a spacelike hypersurface and let n be the vector field of forward directed unit normal vectors on Σ. A real solution u of $(-\Box + m^2)u = 0$ has data on (Σ, n) consisting of the restriction u_Σ and the normal derivative given in local coordinates by

$$\frac{\partial u}{\partial n} \equiv n^\mu \frac{\partial u}{\partial x^\mu}\Big|_\Sigma \tag{10.8}$$

We call a solution *regular* if it is $\mathcal{C}^\infty$ and if the data on any Cauchy surface have compact support.

The basic existence and uniqueness theorem is:

Theorem 10.2 *Let $(\mathcal{M}, g)$ be globally hyperbolic and let (Σ, n) be a Cauchy surface with forward unit normal n. Then for any $f, g \in \mathcal{C}_0^\infty(\Sigma)$ there exists a unique regular solution u of $(-\Box + m^2)u = 0$ such that $u_\Sigma = f$ and $\partial u/\partial n = g$.*

For any smooth functions u, v on $\mathcal{M}$ and any Cauchy surface Σ define

$$\sigma_\Sigma(u, v) = \int_\Sigma \left(u \frac{\partial v}{\partial n} - \frac{\partial u}{\partial n} v \right) d\mu_\Sigma \tag{10.9}$$

where μ_Σ is the measure on Σ induced by the Riemannian volume form on Σ. Let $\mathcal{O}$ be an open set in $\mathcal{M}$ bounded by two Cauchy surfaces Σ_1, Σ_2 and say Σ_2 lies to the future of Σ_1. Then Green's identity says that for any smooth functions u, v

$$\int_\mathcal{O} \left[u(-\Box + m^2)v - v(-\Box + m^2)u \right] d\mu_\mathcal{M} = \sigma_{\Sigma_2}(u, v) - \sigma_{\Sigma_1}(u, v) \tag{10.10}$$

Here $\mu_\mathcal{M}$ is the measure on $\mathcal{M}$ induced by the Lorentzian volume form, in local coordinates $d\mu_\mathcal{M} = |\det g|^{1/2} dx$. If u, v are regular solutions, then the left side vanishes. Hence $\sigma_\Sigma(u, v)$ is independent of Σ and we can denote it just by $\sigma(u, v)$.

Next we need advanced and retarded fundamental solutions. In theorem 7.2 these have already been constructed for Minkowski space.

Theorem 10.3 *Let $(\mathcal{M}, g)$ be globally hyperbolic. Then there exist linear operators $E^\pm : \mathcal{C}_0^\infty(\mathcal{M}) \to \mathcal{C}^\infty(\mathcal{M})$ such that*

$$(-\Box + m^2)E^{\pm}f = E^{\pm}(-\Box + m^2)f = f \tag{10.11}$$

and such that

$$\text{supp}\,(E^{\pm}f) \subset J^{\pm}(\text{supp}\,f) \tag{10.12}$$

From these fundamental solutions we construct the propagator

$$E = E^{+} - E^{-} \tag{10.13}$$

For any $f \in \mathcal{C}_0^{\infty}(\mathcal{M})$ the function $u = Ef$ is a regular solution of the Klein–Gordon equation, and it turns out every regular solution has this form.

Let u be a regular solution regarded as a distribution. For a test function $f \in \mathcal{C}_0^{\infty}(\mathcal{M})$ we have

$$< u, f >= \int_{\mathcal{M}} uf \, d\mu_{\mathcal{M}} \tag{10.14}$$

Then we have the following result which expresses the solution in terms of its data on any Cauchy surface.

Lemma 10.1 *Let u be a regular solution of $(-\Box+m^2)u = 0$ on a globally hyperbolic manifold. Then for any $f \in \mathcal{C}_0^{\infty}(\mathcal{M})$*

$$< u, f >= \sigma(u, Ef) \tag{10.15}$$

Proof We can assume that the manifold has the form $\mathbb{R} \times \Sigma$ with Cauchy surfaces $\Sigma_t = \{t\} \times \Sigma$. By Green's identity we have

$$\sigma_{\Sigma_0}(u, E^{+}f) - \sigma_{\Sigma_{-t}}(u, E^{+}f) = \int_{(-t,0)\times\Sigma} uf \, d\mu_{\mathcal{M}} \tag{10.16}$$

But $E^{+}f$ vanishes on Σ_{-t} for t sufficiently large. Thus taking the limit $t \to \infty$ we have

$$\sigma_{\Sigma_0}(u, E^{+}f) = \int_{(-\infty,0)\times\Sigma} uf \, d\mu_{\mathcal{M}} \tag{10.17}$$

Similarly

$$\sigma_{\Sigma_0}(u, E^{-}f) = -\int_{(0,\infty)\times\Sigma} uf \, d\mu_{\mathcal{M}} \tag{10.18}$$

Taking the difference of the last two equations gives the result. □

10.3 Quantum fields on a manifold

We want to quantize the classical field theory just discussed. Generalizing the Minkowski space case we seek to solve the Klein–Gordon equation with data on some Cauchy surface Σ which satisfy the CCR.

For the Cauchy data we suppose we have a pair of operator valued distributions $\phi(h), \pi(h)$. For each $h \in \mathcal{C}_0^\infty(\Sigma)$ these are symmetric operators on a dense domain in some complex Hilbert space $\mathcal{H}$. They are required to satisfy

$$\begin{aligned} [\phi(h_1), \phi(h_2)] &= 0 \\ [\pi(h_1), \pi(h_2)] &= 0 \\ [\phi(h_1), \pi(h_2)] &= i < h_1, h_2 >_\Sigma \end{aligned} \tag{10.19}$$

where $< h_1, h_2 >_\Sigma = \int h_1 h_2 d\mu_\Sigma$. Equivalently if we define for $\Phi = (\phi, \pi)$ and $H = (h_1, h_2)$

$$\sigma(\Phi, H) = \phi(h_2) - \pi(h_1) \tag{10.20}$$

and define

$$\sigma_\Sigma(H, H') = < h_1, h_2' >_\Sigma - < h_2, h_1' >_\Sigma \tag{10.21}$$

then

$$[\sigma(\Phi, H), \sigma(\Phi, H')] = i\sigma_\Sigma(H, H') \tag{10.22}$$

Now we evolve in time and define $\phi(f)$ for $f \in \mathcal{C}_0^\infty(\mathcal{M})$ by analogy with (10.15) as

$$\phi(f) = \sigma(\Phi, \rho_\Sigma(Ef)) \tag{10.23}$$

where

$$\rho_\Sigma(u) = \left(u_\Sigma, \frac{\partial u}{\partial n} \right) \tag{10.24}$$

are the data on Σ for a solution u.

Theorem 10.4 *For $f \in \mathcal{C}_0^\infty(\mathcal{M})$ the field operator $\phi(f)$ satisfies the field equation*

$$(-\Box + m^2)\phi = 0 \tag{10.25}$$

and

$$[\phi(f_1), \phi(f_2)] = -i < f_1, Ef_2 > \tag{10.26}$$

In particular the commutator vanishes if f_1, f_2 have spacelike separated supports, that is if $\mathrm{supp} f_1 \cap J^\pm(\mathrm{supp} f_2) = \emptyset$ *(locality).*

Proof The operator $\Box$ is symmetric for the inner product (10.14) so the meaning of the field equation is $\phi((-\Box + m^2)f) = 0$. This follows from $E(-\Box + m^2) = 0$.

For the second point we compute

$$\begin{aligned} [\phi(f_1), \phi(f_2)] &= \left[\sigma(\Phi, \rho_\Sigma(Ef_1)), \sigma(\Phi, \rho_\Sigma(Ef_2)) \right] \\ &= i\sigma_\Sigma \left(\rho_\Sigma(Ef_1), \rho_\Sigma(Ef_2) \right) \\ &= i\sigma(Ef_1, Ef_2) \\ &= -i < f_1, Ef_2 > \end{aligned} \tag{10.27}$$

where the last step follows from (10.15). □

Remarks

1. The construction does not depend on the choice of Cauchy surface. Indeed we can formulate it so that there is no reference to a particular Cauchy surface as follows. Consider the space of regular solutions u of the Klein–Gordon equation with symplectic form $\sigma(u, v)$. Let $R(u)$ be a representation of the CCR over this space, that is we have a family of densely defined operators $R(u)$ on a Hilbert space $\mathcal{H}$ such that $R(u)$ is linear in u and

$$[R(u), R(v)] = i\sigma(u, v) \tag{10.28}$$

   Define the field operator $\phi(f)$ by

$$\phi(f) = R(Ef) \tag{10.29}$$

   This satisfies the field equation (10.25) and has the commutator (10.26) just as in the theorem.

   Now if we pick a Cauchy surface Σ, we can define $\sigma(\Phi, H) = R(\rho_\Sigma^{-1} H)$ where $\rho_\Sigma^{-1} H$ is the unique regular solution with data H on Σ. Then

$$[\sigma(\Phi, H), \sigma(\Phi, H')] = i\sigma(\rho_\Sigma{}^{-1} H, \rho_\Sigma{}^{-1} H') = i\sigma_\Sigma(H, H') \tag{10.30}$$

   and $\phi(f) = \sigma(\Phi, \rho_\Sigma(Ef))$ so we recover our earlier construction.
2. There is still the question of which representation of the CCR to take. In this generality there is no definitive answer. Without a timelike symmetry there is no Hamiltonian and so we cannot ask for a positive energy representation as we did before. There is also no vacuum to use as a reference point so it is also difficult to identify particles. One has to proceed on a case by case basis.
3. The construction does respect the principle of general covariance mentioned at the beginning of chapter 7. We made no special choice of coordinates.

Problem 10.1 Show that representations of the CCR as defined by (10.19) or (10.22) exist by making an explicit construction.

Problem 10.2 In the special case of Minkowksi space $(\mathbb{R}^4, \eta)$, take $\Sigma = \{0\} \times \mathbb{R}^3$ and take $\sigma(\Phi, H)$ given by (8.63). Show that the new definition of the field $\phi(f) = \sigma(\Phi, \rho_\Sigma(Ef))$ agrees with the old definition (8.66).

Notes on chapter 10: These topics are covered by Wald (1994) and by Bär Ginoux and Pfäffle (2007).

In spite of the uncertainty in the choice of a representation of the CCR it has been possible to identify a class of representations with desirable physical properties. These are characterized by the requirement that correlation functions have certain prescribed "Hadamard singularities" at coinciding points, see Wald (1994).

For a mathematical formulation of the principle of general covariance, see Dimock (1980) or Brunetti Fredenhagen and Verch (2003).

By studying a scalar field on the spacetime manifold for a collapsing black hole, Hawking was led to his famous prediction that black holes emit thermal radiation. See Bachelot (1999).

Part III

Probabilistic methods

11 Path integrals

As we have seen quantum mechanics is fundamentally probabilistic, but it is a special kind of noncommutative probability. Nevertheless the techniques of standard probability theory can also be useful. In the remainder of the book we explore some of the ways this occurs.

In this chapter we return to the consideration of a single non-relativistic particle and develop some new representations of the dynamics. These are the Feynman path integrals which express the quantum dynamics as an integral over all possible classical paths with a special weighting. For the single particle Hamiltonian $H = -\Delta/2 + V$ and $\psi, \chi \in L^2(\mathbb{R}^3)$ a typical integral is

$$(\psi, e^{-iHt}\chi) = \int \overline{\psi(\omega(0))} \exp\left(-i\int_0^t V(\omega(s))ds\right) \chi(\omega(t))d\omega \tag{11.1}$$

Here the "integral" is over all possible paths $\omega : [0,t] \to \mathbb{R}^3$ and "$d\omega$" is supposed to be some kind of measure on these paths. Actually it has not been possible to make sense of this within the context of standard measure theory. But if one replaces the time evolution e^{-iHt} by e^{-Ht}, that is if we go to imaginary time, then one can give a rigorous formulation. This is what we study in this chapter. This gets us away from the basic dynamics but still can be useful in an indirect way. For example one can study properties of the Hamiltonian through the semi-group e^{-Ht} as represented by path integrals. Similar representations occur in quantum field theory where they are crucial for a mathematical analysis. We explore this in subsequent chapters.

11.1 Probability

We start by reviewing some definitions. A measure space is a triple $(\mathcal{M}, \Sigma, \mu)$ consisting of a set $\mathcal{M}$, a σ-algebra of subsets Σ, and a measure $\mu : \Sigma \to [0, \infty]$. We consider probability measure spaces which have $\mu(\mathcal{M}) = 1$. In applications $\mathcal{M}$ represents all possible outcomes, $A \in \Sigma$ represent events, and $\mu(A)$ is interpreted as the probability that A occurs.

Random variables are measurable functions $X : \mathcal{M} \to \mathbb{R}$. The probability that X takes values in a Borel set B is

$$P(X \in B) = \mu(X^{-1}[B]) \equiv m(B) \tag{11.2}$$

The set function $m(B)$ is a probability measure on $\mathbb{R}$ called the *distribution* of X. It contains all relevant information about X. For any Borel function f the function $f(X) = f \circ X$ is again a random variable and we have the *expectation* or *mean*

$$E(f(X)) \equiv \int_{\mathcal{M}} f(X) d\mu = \int_{\mathbb{R}} f(x) dm(x) \tag{11.3}$$

if the integrals exist. The identity of the two integrals is a standard argument. First verify it for simple functions, then by monotone limits for positive functions, and finally for any integrable function. In particular there is the *characteristic function* of X, which is a function on $\mathbb{R}$ defined by

$$\Phi(s) = E(e^{isX}) = \int e^{isx} dm(x) \tag{11.4}$$

As the Fourier transform of the measure m it uniquely determines m. Indeed the probability distribution m determines a tempered distribution m and the Fourier transform is bijective on $\mathcal{S}'(\mathbb{R})$ – see appendix C.

More generally suppose we have n random variables $X_1, \ldots, X_n$ on a probability measure space $(\mathcal{M}, \Sigma, \mu)$. Equivalently we have a vector-valued random variable $X = (X_1, \ldots, X_n) : \mathcal{M} \to \mathbb{R}^n$. Again the probability that the random variables take values in a Borel set $B \subset \mathbb{R}^n$ is

$$P\Big((X_1, \ldots, X_n) \in B\Big) = \mu(X^{-1}[B]) \equiv m(B) \tag{11.5}$$

The distribution $m(B)$ is now a Borel measure on $\mathbb{R}^n$. For any Borel function f on $\mathbb{R}^n$

$$E(f(X)) = \int_{\mathcal{M}} f(X_1, \ldots, X_n) d\mu = \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) dm(x) \tag{11.6}$$

The characteristic function is $\Phi : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\Phi(s_1, \ldots, s_n) = E(e^{i \sum_i s_i X_i}) = \int e^{i \sum_i s_i x_i} dm(x) \tag{11.7}$$

and it determines the distribution.

We generalize still further and suppose we have an infinite family $\{X_\alpha\}_{\alpha \in \mathcal{A}}$ of random variables on a space $(\mathcal{M}, \Sigma, \mu)$, called a *stochastic process*. For any finite ordered subset $I = (\alpha_1, \ldots, \alpha_n)$ from $\mathcal{A}$ we have the family of random variables $X_I = (X_{\alpha_1}, \ldots, X_{\alpha_n})$ and we consider their joint distribution

$$P(X_I \in B) = \mu(X_I^{-1}(B)) \equiv m_I(B) \tag{11.8}$$

Then

$$E(f(X_I)) = \int_{\mathcal{M}} f(X_I) d\mu = \int_{\mathbb{R}^n} f(x) dm_I(x) \tag{11.9}$$

The joint distributions must satisfy some consistency conditions. Let π be a permutation of $(1,\dots,n)$. This induces mappings $\pi I = (\alpha_{\pi(1)},\dots,\alpha_{\pi(n)})$ on the index set and $\pi x = (x_{\pi(1)},\dots,x_{\pi(n)})$ on $\mathbb{R}^n$. Then the event $X_{\pi(I)} \in \pi B$ is the same as the event $X_I \in B$ and hence

$$m_{\pi(I)}(\pi(B)) = m_I(B) \tag{11.10}$$

Also if $I' = (I,\alpha)$, then the event $X_{I'} = (X_I, X_\alpha) \in B \times \mathbb{R}$ is the same as the event $X_I \in B$ and hence

$$m_{I'}(B \times \mathbb{R}) = m_I(B) \tag{11.11}$$

The family of all finite-dimensional distributions is enough to specify the full structure for one has the following result on the existence of stochastic processes

Theorem 11.1 (Kolmogorov) *Let $\mathcal{A}$ be an index set and suppose for each finite ordered subset I in $\mathcal{A}$ there is a Borel probability measure m_I on $\mathbb{R}^{|I|}$. The measures are assumed to satisfy the consistency conditions (11.10), (11.11). Then there exists a probability measure space $(\mathcal{M},\Sigma,\mu)$ and a family of random variables $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ such that for any finite ordered subset I in $\mathcal{A}$ the random variables X_I have the distributions m_I.*

There is also a uniqueness result which says that (under some further conditions) any two realizations are equivalent by an isomorphism of measure spaces.

11.2 Gaussian processes

Now consider a special class of random variables, the Gaussian random variables. A random variable X is *Gaussian* if there are constants c, a such that for any Borel $B \subset \mathbb{R}$ the distribution is the normal distribution

$$P(X \in B) = m(B) = (2\pi c)^{-1/2} \int_B \exp\left(-\frac{(x-a)^2}{2c}\right) dx \tag{11.12}$$

Then for any Borel function $f : \mathbb{R} \to \mathbb{R}$ we have that $f(X)$ is integrable iff $f(x)\exp(-(x-a)^2/2c)$ is integrable on $\mathbb{R}$ in which case

$$E(f(X)) = \int f(x)dm(x) = (2\pi c)^{-1/2} \int_{\mathbb{R}} f(x)\exp\left(-\frac{(x-a)^2}{2c}\right) dx \tag{11.13}$$

For the second step one again verifies the identity successively for simple functions, positive functions, and integrable functions. In particular the characteristic function is (see problem 1.3)

$$\Phi(s) = E(e^{isX}) = \exp\left(ixa - \frac{1}{2}cs^2\right) \tag{11.14}$$

Taking derivatives at $s = 0$ we find that X has mean $E(X) = a$ and $E(X^2) = c + a^2$. The variance is then

$$Var(X) \equiv E(X^2) - E(X)^2 = c \tag{11.15}$$

More generally a family of random variables $X = (X_1, \ldots, X_n)$ is *jointly Gaussian* if there is an $n \times n$ positive definite symmetric matrix $C = \{C_{ij}\}$ and $a \in \mathbb{R}^n$ such that for any Borel $B \subset \mathbb{R}^n$ we have the joint distribution

$$\begin{aligned} &P(X \in B) = m(B) \\ &= (2\pi)^{-n/2} (\det C)^{-1/2} \int_B \exp\left(-\frac{1}{2}(x-a)\cdot C^{-1}(x-a)\right) dx \end{aligned} \tag{11.16}$$

(Positive definite means $x \cdot Cx \geq 0$ and $x \cdot Cx = 0$ iff $x = 0$.) It is clear that such structures exist. Indeed we can take $(\mathcal{M}, \mu) = (\mathbb{R}^n, m)$ and $X_i(x) = x_i$. For any Borel function $f : \mathbb{R}^n \to \mathbb{R}$ we have

$$\begin{aligned} &E(f(X)) \\ &= (2\pi)^{-n/2} (\det C)^{-1/2} \int_{\mathbb{R}^n} f(x) \exp\left(-\frac{1}{2}(x-a)\cdot C^{-1}(x-a)\right) dx \end{aligned} \tag{11.17}$$

if the integral exists. In particular we find for the characteristic function $\Phi(s) = \Phi(s_1, \ldots, s_n)$

$$\begin{aligned} \Phi(s) &= E\left(e^{is\cdot X}\right) \\ &= (2\pi)^{-n/2} (\det C)^{-1/2} \int_{\mathbb{R}^n} e^{is\cdot x} \exp\left(-\frac{1}{2}(x-a)\cdot C^{-1}(x-a)x\right) dx \\ &= \exp\left(is\cdot a - \frac{1}{2} s\cdot Cs\right) \end{aligned} \tag{11.18}$$

Taking derivatives at $s = 0$ we find the mean and covariance are

$$\begin{aligned} E(X_i) &= a_i \\ E(X_i X_j) - E(X_i)E(X_j) &= C_{ij} \end{aligned} \tag{11.19}$$

The mean and the covariance completely characterize the family of Gaussian random variables.

Finally suppose we have an infinite set $\mathcal{A}$ and functions $a : \mathcal{A} \to \mathbb{R}$ and symmetric $C : \mathcal{A} \times \mathcal{A} \to \mathbb{R}$ such that for any $\alpha_1, \ldots, \alpha_n \in \mathcal{A}$ the matrix $C_{ij} = C(\alpha_i, \alpha_j)$ is positive definite. Then we can define a *Gaussian process* with mean a and covariance C to be a collection of random variables $\{X_\alpha\}$ indexed by $\alpha \in \mathcal{A}$ such that for any finite collection $I = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ the random variables $X_I = (X_{\alpha_1}, \ldots, X_{\alpha_n})$ are jointly Gaussian with means $a_I = (a(\alpha_1), \ldots, a(\alpha_n))$ and covariance $C_I = \{C(\alpha_i, \alpha_j)\}$. In particular then

$$\begin{aligned} E(X_\alpha) &= a(\alpha) \\ E(X_\alpha X_\beta) - E(X_\alpha)E(X_\beta) &= C(\alpha, \beta) \end{aligned} \tag{11.20}$$

Such a Gaussian process exists as a consequence of the theorem 11.1 once we show that the consistency conditions on the finite-dimensional measures m_I are satisfied. To show such measures are equal it suffices to show that the characteristic functions are equal. Thus it suffices to consider the characteristic function Φ_I for X_I, which is

$$\Phi_I(s) = \exp\left(i \sum_{i=1}^{n} s_i a(\alpha_i) - \frac{1}{2} \sum_{i,j=1}^{n} s_i C(\alpha_i, \alpha_j) s_j \right) \tag{11.21}$$

The permutation condition (11.10) translates as $\Phi_{\pi I}(\pi s) = \Phi_I(s)$. The extension condition (11.11) translates as $\Phi_{(I,\alpha_{n+1})}(s, 0) = \Phi_I(s)$. Both are easily checked and hence the process exists.

Problem 11.1 Let C be a symmetric positive definite matrix.

1. Show that C is invertible and that C^{-1} is symmetric and positive definite.
2. Show that there exists $M > 0$ so that $x \cdot Cx \geq M|x|^2$, hence also for C^{-1}.

Problem 11.2 Do the integral in (11.18). (Hint: diagonalize C.)

Problem 11.3

1. Show that a Gaussian random variable is in L^p for all $1 \leq p < \infty$.
2. Justify the differentiations in computing the mean and covariance in (11.19).

11.3 Brownian motion

This is a particular example of a Gaussian process indexed by $\mathbb{R}^+ = [0, \infty)$. A family of random variables X_t, $t \geq 0$ is a *Brownian motion* if it is Gaussian with mean and covariance

$$E(X_t) = 0 \qquad E(X_t X_s) = \min(t, s) \tag{11.22}$$

For this to be well defined we need for any distinct $t_1, t_2, \ldots, t_n$ that the matrix $\{\min(t_i, t_j)\}$ is positive definite. To verify this it suffices to assume $t_1 < t_2 < \cdots < t_n$. Then we have the identity (with $t_0 = 0$)

$$\sum_{ij} x_i x_j \min(t_i, t_j) = \sum_{i=1}^{n} (t_i - t_{i-1}) \left(\sum_{j=i}^{n} x_j \right)^2 \tag{11.23}$$

This vanishes iff $\sum_{j=i}^{n} x_j = 0$ for $i = 1, \ldots, n$ which occurs iff $x_i = 0$. Hence the matrix is positive definite.

Now X_t has mean zero and variance $E(X_t^2) = t$. Hence for $t = 0$, X_0 takes the constant value $X_0 = 0$. If $t > 0$, we compute the probability that X_t is in a Borel set B as

$$P(X_t \in B) = \int_B p_t(x)dx \tag{11.24}$$

where p_t is the even function

$$p_t(x) = (2\pi t)^{-1/2} e^{-x^2/2t} \tag{11.25}$$

This is saying that for t small X_t takes values near the origin with high probability, while for t large that probability is widely spread around the origin. These features allow the interpretation that X_t describes the location at time t of a diffusing particle which starts at the origin at $t = 0$ and moves randomly as time evolves.

More generally a Brownian motion starting at $x \in \mathbb{R}$ is a family of Gaussian random variables X_t^x with mean x and variance t defined by

$$X_t^x = X_t + x \tag{11.26}$$

Then we find

$$\begin{aligned} P(X_t^x \in B) &= P(X_t \in B - x) \\ &= \int_{B-x} p_t(y)dy = \int_B p_t(y - x)dy \end{aligned} \tag{11.27}$$

Similarly if $p_t(x - \cdot)f$ is integrable, we compute

$$\begin{aligned} E(f(X_t^x)) &= E(f(X_t + x)) \\ &= \int f(y + x)p_t(y)dy = \int p_t(x - y)f(y)dy \end{aligned} \tag{11.28}$$

Note the following facts. The sum of two Gaussian random variables is again Gaussian. Hence for $s < t$, $X_t^x - X_s^x = X_t - X_s$ is Gaussian with mean zero and variance

$$E((X_t^x - X_s^x)^2) = E((X_t - X_s)^2) = t - s - s + s = t - s \tag{11.29}$$

Furthermore for $s_1 \leq t_1 \leq s_2 \leq t_2$ we have that $X_{t_1}^x - X_{s_1}^x$ and $X_{t_2}^x - X_{s_2}^x$ are uncorrelated since

$$E\Big((X_{t_1}^x - X_{s_1}^x)(X_{t_2}^x - X_{s_2}^x)\Big) = t_1 - t_1 - s_1 + s_1 = 0 \tag{11.30}$$

For Gaussian random variables uncorrelated means independent so $X_{t_1}^x - X_{s_1}^x$ and $X_{t_2}^x - X_{s_2}^x$ are independent, which means that the joint distribution is the product of the individual distributions. One says that the process has independent increments.

In particular for $0 < s < t$ the random variables $X_t^x - X_s^x$ and $X_s^x = X_s^x - X_0^x$ are independent and so

$$E\Big(f(X_s^x, X_t^x - X_s^x)\Big) = \int p_s(x-y_1)\, p_{t-s}(y_2)\, f(y_1,y_2)\, dy_1 dy_2 \tag{11.31}$$

Then we compute

$$\begin{aligned} E(f_1(X_s^x) f_2(X_t^x)) &= E(f_1(X_s^x) f_2(X_s^x + (X_t^x - X_s^x))) \\ &= \int p_s(x-y_1)\, p_{t-s}(y_2)\, f_1(y_1) f_2(y_1+y_2)\, dy_1 dy_2 \\ &= \int p_s(x-y_1)\, f_1(y_1)\ \ p_{t-s}(y_1-y_2)\ \ f_2(y_2)\, dy_1 dy_2 \end{aligned} \tag{11.32}$$

Similarly for $0 < t_1 < t_2 < \cdots < t_n$

$$\begin{aligned} &E\big(f_1(X_{t_1}^x) \cdots f_n(X_{t_n}^x)\big) \\ &= \int p_{t_1}(x-y_1) f_1(y_1) \cdots p_{t_n - t_{n-1}}(y_{n-1} - y_n) f_n(y_n) dy_1 \cdots dy_n \end{aligned} \tag{11.33}$$

Specializing to characteristic functions we have for the joint distribution

$$\begin{aligned} &P(X_{t_1}^x \in B_1, \ldots, X_{t_n}^x \in B_n) \\ &= \int_{B_1 \times \cdots \times B_n} p_{t_1}(x-y_1) \cdots p_{t_n - t_{n-1}}(y_{n-1} - y_n) dy_1 \cdots dy_n \end{aligned} \tag{11.34}$$

This gives an idea of the character of the measure on paths.

We quote the following regularity result which shows that Brownian paths are continuous but nowhere differentiable.

Theorem 11.2 *There is a construction of Brownian motion X_t on a measure space $(\mathcal{M}, \mu)$ with the following properties:*

1. *Let $\alpha < 1/2$. Then for almost every $\omega \in \mathcal{M}$ the path $t \to X_t(\omega)$ is Holder continuous with exponent α, that is there is a constant C_ω such that*

$$|X_t(\omega) - X_s(\omega)| < C_\omega |t-s|^\alpha \tag{11.35}$$

2. *Let $\alpha > 1/2$. Then for almost every $\omega \in \mathcal{M}$ the path $t \to X_t(\omega)$ is nowhere Holder continuous with exponent α.*

The construction needs more than the Kolmogorov theorem. Just for the continuity one way to proceed is to take as the basic probability space the continuous functions $\mathcal{M} = \{\omega \in \mathcal{C}(\mathbb{R}^+) : \omega(0) = 0\}$ and then construct a measure so that the evaluation maps $X_t(\omega) = \omega(t)$ give a Brownian motion. This gives a strong meaning to the idea that we are integrating over a space of paths.

All the above is easily generalized to $\mathbb{R}^d$. We set

$$X_t = (X_t^1, \ldots, X_t^d) \tag{11.36}$$

where the X_t^i are independent one-dimensional Brownian motions. All the above formulas generalize. For example we have for a function f on $\mathbb{R}^d$ and a point $x \in \mathbb{R}^d$

$$E(f(X_t^x)) = (2\pi t)^{-d/2} \int \exp\left(-\frac{|x-y|^2}{2t}\right) f(y) dy \tag{11.37}$$

Problem 11.4 Verify (11.23).

11.4 The Feynman–Kac formula

The connection with single particle quantum mechanics comes via the Laplacian. We work in d dimensions. With mass $m = 1$ the free Hamiltonian is $H_0 = -\Delta/2$ and for $f \in L^2(\mathbb{R}^d)$ we have from (4.8)

$$(e^{-H_0 t} f)(x) = (2\pi t)^{-d/2} \int \exp\left(-\frac{|x-y|^2}{2t}\right) f(y) dy \tag{11.38}$$

This is the same as (11.37). Thus if X_t^x is Brownian motion in $\mathbb{R}^d$ starting at x, then

$$(e^{-H_0 t} f)(x) = E\big(f(X_t^x)\big) \tag{11.39}$$

Taking into account that $X_0^x = x$ another way to write this is

$$(g, e^{-H_0 t} f) = \int E\left(\overline{g(X_0^x)} f(X_t^x)\right) dx \tag{11.40}$$

Now we add a bounded potential to the Hamiltonian

Theorem 11.3 *(Feynman–Kac formula) Let V be bounded and continuous on $\mathbb{R}^d$ and let $H = H_0 + V$. Then for $f, g \in L^2(\mathbb{R}^d)$*

$$(g, e^{-Ht} f) = \int E\left(\overline{g(X_0^x)} \exp\left(-\int_0^t V(X_s^x) ds\right) f(X_t^x)\right) dx \tag{11.41}$$

Remark For the proof we use the *Trotter product formula*[1] which says that if S and T are self-adjoint and bounded below and $S+T$ defined on $D(T) \cap D(S)$ is self-adjoint, then as a strong limit

$$\lim_{n\to\infty} \left(e^{-St/n} e^{-Tt/n}\right)^n = e^{-(S+T)t} \tag{11.42}$$

We apply this with $S = H_0$, $T = V$. The sum $H = H_0 + V$ is self-adjoint on $D(H_0) \cap D(V) = D(H_0)$ by theorem 4.1.

Proof By (11.33) for $t_1 < t_2 < \cdots < t_n$

$$\begin{aligned} &E\Big(f_1(X_{t_1}^x) \cdots f_n(X_{t_n}^x)\Big) \\ &= \left(e^{-t_1 H_0} f_1 e^{-(t_2-t_1)H_0} f_2 \ldots e^{-(t_n - t_{n-1})H_0} f_n\right)(x) \end{aligned} \tag{11.43}$$

[1] See for example Reed and Simon (1980: 295).

Thus with $t_j = jt/n$

$$\begin{aligned}(g, e^{-Ht}f) &= \lim_{n\to\infty}\left(g, \left(e^{-H_0 t/n}e^{-Vt/n}\right)^n f\right)\\ &= \lim_{n\to\infty}\int \overline{g(x)}E\left(\prod_{j=1}^{n}\exp\left(-\frac{t}{n}V(X_{t_j}^x)\right)f(X_t^x)\right)\,dx\\ &= \lim_{n\to\infty}\int E\left(\overline{g(X_0^x)}\exp\left(-\frac{t}{n}\sum_{j=1}^{n}V(X_{t_j}^x)\right)f(X_t^x)\right)\,dx\\ &= \int E\left(\overline{g(X_0^x)}\exp\left(-\int_0^t V(X_s^x)\,ds\right)f(X_t^x)\right)dx\end{aligned} \tag{11.44}$$

Here in the last step we have used that

$$\lim_{n\to\infty}\frac{t}{n}\sum_{j=1}^{n}V(X_{t_j}^x) = \int_0^t V(X_s^x)\,ds \tag{11.45}$$

holds almost everywhere. This follows by the continuity of V and the continuity of paths with the integral interpreted as a Riemann integral. We have also used

$$\left|\overline{g(X_0^x)}\exp\left(-\frac{t}{n}\sum_{j=1}^{n}V(X_{t_j}^x)\right)f(X_t^x)\right| \le |g(X_0^x)||f(X_t^x)|e^{t\|V\|_\infty} \tag{11.46}$$

Then since

$$\int E\Big(|g(X_0^x)||f(X_t^x)|\Big)dx = (|g|, e^{-H_0 t}|f|) < \infty \tag{11.47}$$

we can use the dominated convergence theorem to take the limit inside the integrals in (11.44). □

Problem 11.5 For $h \in L^\infty(\mathbb{R}^d)$ and $0 \le u \le t$ show that

$$\begin{aligned}&(g, e^{-uH}\, h\, e^{-(t-u)H}f)\\ &= \int E\left(\overline{g(X_0^x)}h(X_u^x)\exp\left(-\int_0^t V(X_s^x)ds\right)f(X_t^x)\right)\,dx\end{aligned} \tag{11.48}$$

11.5 Oscillator process

The *oscillator process* (also called the Ornstein–Uhlenbeck process) is defined to be the Gaussian process X_t indexed by $t \in \mathbb{R}$ with mean and covariance

$$E(X_t) = 0 \qquad E(X_s X_t) = C(s,t) \tag{11.49}$$

where

$$C(s,t) = (2\pi)^{-1} \int \frac{e^{ip(s-t)}}{p^2+1} dp = \frac{1}{2} e^{-|t-s|} \tag{11.50}$$

The second version follows by closing the contour in the upper or lower half plane depending on this sign of $s - t$. This is positive definite since for any sequence $x_1, \dots, x_n$ and any choice of points $t_1, \dots, t_n$

$$\sum_{ij} x_i x_j C(t_i, t_j) = (2\pi)^{-1} \int \frac{|\sum_i x_i e^{ipt_i}|^2}{p^2+1} dp \ \geq \ 0 \tag{11.51}$$

and it vanishes iff $\sum_i x_i e^{ipt_i} = 0$ for all p which occurs iff $x_i = 0$.

Note that X_t has mean zero and constant variance $1/2$. Thus if it is describing the motion of a particle, it is not diffusive like Brownian motion but stays localized around the origin. This process also has continuous paths.

The oscillator process is related to the semi-group e^{-Ht} generated by the harmonic oscillator Hamiltonian which we considered in section 4.4 and which is given by

$$H = \frac{1}{2}\Big(-\frac{d^2}{dx^2} + x^2 \Big) \tag{11.52}$$

Recall that the operator has discrete spectrum and that the lowest eigenvalue is $1/2$ with eigenvector $\Omega_0(x) = \pi^{-1/4} e^{-x^2/2}$.

Theorem 11.4 *Let X_t be the oscillator process and let f, g be polynomially bounded functions on $\mathbb{R}$. Then for $t > 0$*

$$(g\Omega_0, e^{-(H-\frac{1}{2})t} f\Omega_0) = E(\overline{g(X_0)} f(X_t)) \tag{11.53}$$

Proof The covariance matrix for X_0, X_t is

$$C = \begin{pmatrix} C(0,0) & C(0,t) \\ C(t,0) & C(t,t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & e^{-t} \\ e^{-t} & 1 \end{pmatrix} \tag{11.54}$$

Thus we compute by (11.17)

$$\begin{aligned}
&E(\overline{g(X_0)} f(X_t)) \\
&= (2\pi)^{-1} (\det C)^{-1/2} \int \overline{g(x_1)} \exp\left(-\frac{1}{2} x \cdot C^{-1} x \right) f(x_2) dx \\
&= \pi^{-1} (1 - e^{-2t})^{-1/2} \\
&\quad \int \overline{g(x_1)} \exp\Big(-(1 - e^{-2t})^{-1} (x_1^2 + x_2^2 - 2x_1 x_2 e^{-t}) \Big) f(x_2) dx \\
&= e^{t/2} (2\pi \sinh t)^{-1/2} \\
&\quad \int \overline{(g\Omega_0)(x_1)} \exp\left(-\frac{1}{2} (\coth t)(x_1^2 + x_2^2) + (\sinh t)^{-1} x_1 x_2 \right) (f\Omega_0)(x_2) dx \\
&= (g\Omega_0, e^{-(H-\frac{1}{2})t} f\Omega_0)
\end{aligned} \tag{11.55}$$

The last step follows by Mehler's formula (4.38). □

Remarks

1. There is also a Feynman–Kac formula for the oscillator process. If V is a bounded continuous function on $\mathbb{R}$ and

$$H' = H + V = \frac{1}{2}\Big(-\frac{d^2}{dx^2} + x^2\Big) + V \tag{11.56}$$

   Then

$$(g\Omega_0, e^{-(H'-\frac{1}{2})t} f\Omega_0) = E\left(\overline{g(X_0)} \exp\left(-\int_0^t V(X_s)ds\right) f(X_t)\right) \tag{11.57}$$

2. The fact that Brownian motion and the oscillator process are related to semi-groups is not accidental. In fact both are (time homogeneous) Markov processes which roughly means that the future depends only on the present and not on the past. Such Markov processes always give rise to semi-groups of operators.

11.6 Application: ground states

If a quantum mechanical system has a Hamiltonian for which the bottom of the spectrum is an eigenvalue, then states in the corresponding eigenspace are called *ground states*. These are the states most likely to be occupied. In this section we give some results about ground states for non-relativistic single particle systems using the Feynman–Kac formula.

First some definitions. A function on a measure space $(\mathcal{M}, \mu)$ is *positive*, written $f \geq 0$, if $f(m) \geq 0$ for almost every m and f is not identically zero. A function is *strictly positive*, written $f > 0$, if $f(m) > 0$ for almost every m. A function $f \in L^2$ is strictly positive iff $(f, g) > 0$ for every positive $g \in L^2$ (see problem below). A bounded operator A on $L^2(\mathcal{M}, d\mu)$ is *positivity improving* $Af > 0$ whenever $f \geq 0$. This is true iff $(g, Af) > 0$ whenever $f \geq 0, g \geq 0$.

Problem 11.6 Let $(\mathcal{M}, \mu)$ be a σ-finite measure space.[2] Show that $f \in L^2$ is strictly positive iff $(f, g) > 0$ for every positive $g \in L^2$.

For example on $L^2(\mathbb{R}^d)$ consider e^{-tH_0} where H_0 is the free Hamiltonian. The operator e^{-tH_0} has a strictly positive kernel $(2\pi t)^{-d/2} \exp(-|x-y|^2/2t)$ and hence it is positivity improving.

Lemma 11.1 *Let A be a bounded self-adjoint operator on $L^2(\mathcal{M}, \mu)$ which is positivity improving. If $\|A\|$ is an eigenvalue, then the eigenspace is spanned by a single strictly positive function.*

[2] σ-finite means there is a sequence of subsets $\mathcal{M}_i$ with finite measure so $\cup_i \mathcal{M}_i = \mathcal{M}$.

Remark Since $\sigma(A) \subset (-\|A\|, \|A\|)$, $\|A\|$ is the largest possible eigenvalue.

Proof Let ψ be an eigenvector for A with eigenvalue $\|A\|$. Since A is reality preserving both the real and imaginary parts are eigenvectors with eigenvalues $\|A\|$. Hence we may as well assume ψ is real. Then $|\psi| \pm \psi \geq 0$ hence $A(|\psi| \pm \psi) > 0$ and hence

$$|A\psi| \leq A|\psi| \tag{11.58}$$

It follows that

$$\|A\|\|\psi\|^2 = (A\psi, \psi) \leq (|A\psi|, |\psi|) \leq (A|\psi|, |\psi|) \leq \|A\|\|\psi\|^2 \tag{11.59}$$

and hence

$$(A\psi, \psi) = (A|\psi|, |\psi|) \tag{11.60}$$

Write $\psi = \psi_+ - \psi_-$ where $\psi_\pm \geq 0$. Then $|\psi| = \psi_+ + \psi_-$ and the last identity implies that

$$(A\psi_+, \psi_-) + (A\psi_-, \psi_+) = 0 \tag{11.61}$$

If $\psi_\pm$ are both nonzero, this contradicts the strict positivity. Thus one of them must be zero and we may assume that $\psi_- = 0$. Thus $\psi \geq 0$. Since $\psi = \|A\|^{-1}A\psi$ we have $\psi > 0$. Finally if ψ' is another eigenvector, then by the same argument conclude $\psi' > 0$. Then ψ' cannot be orthogonal to ψ so the eigenspace is one dimensional. □

The next result shows that ground states are unique.

Theorem 11.5 *Let $H_0 = -\Delta/2$ and let V be a bounded continuous function on $\mathbb{R}^d$ so that $H = H_0 + V$ is self-adjoint.*

1. *The operators e^{-tH} are positivity improving for all $t > 0$.*
2. *If H has an eigenvalue at the bottom of the spectrum, then the eigenspace is spanned by a single strictly positive function.*

Proof For $f, g \geq 0$ we have by the Feynman–Kac formula for Brownian motion (11.41)

$$\begin{aligned}(g, e^{-tH}f) &= \int E\left(g(X_0^x)\exp\left(-\int_0^t V(X_s^x)ds\right)f(X_t^x)\right)\,dx \\ &\geq e^{-t\|V\|_\infty}\int E\left(g(X_0^x)f(X_t^x)\right)\,dx \\ &= e^{-t\|V\|_\infty}(g, e^{-tH_0}f)\end{aligned} \tag{11.62}$$

Hence $(g, e^{-tH}f) > 0$ and e^{-tH} is positivity improving.

If E is a lowest eigenvalue for H, then $e^{-tE} = \|e^{-tH}\|$ is a highest eigenvalue for e^{-tH} and the eigenspace is the same. The result now follows by lemma 11.1. This proves the second point. □

Problem 11.7 Let $H = H_0 + V$ and let $H' = H_0 + V'$ be as in the theorem with isolated lowest eigenvalues $E, E' < 0$ and with normalized eigenvectors ψ, ψ'.

1. Show that $(\psi, \psi') > 0$.
2. Show that
$$E' = \lim_{t\to\infty} -\frac{1}{t} \log(\psi, e^{-tH'}\psi) \tag{11.63}$$
3. Show that
$$\psi' = \lim_{t\to\infty} \frac{e^{-tH'}\psi}{\|e^{-tH'}\psi\|} \tag{11.64}$$

Problem 11.8 (Perturbation theory) As in the previous problem suppose ψ, E are known and $V' = V + \lambda V_1$. We want to compute $E' = E'(\lambda)$ for λ small. More precisely we want to compute the first-order term in an expansion
$$E'(\lambda) = E + \left(\frac{dE'}{d\lambda}(0)\right)\lambda + \cdots \tag{11.65}$$
Assuming one can exchange limits and derivatives use the representation (11.63) and the Feynman–Kac formula to compute the first-order term.
(Answer: $(dE'/d\lambda)(0) = (\psi, V_1\psi)$.)

Notes on chapter 11: There are many books on the fundamentals of probability and stochastic processes, for example Billingsley (1979) or Durrett (1996).

For more on path integrals and their application to physics, see Simon (1979) or Glimm and Jaffe (1987).

It is possible to make some sense of Feynman's original real time path integral (11.1). See Albeverio *et al.* (2008).

12 Fields as random variables

Now we return to quantum field theory. The time zero scalar fields are a family of commuting symmetric operators on Fock space. Thinking of the spectral theorem, this suggests that it may be possible to represent them all as functions on some measure space. In this chapter we develop this representation, known as the Schrödinger representation. This representation also leads to a path space representation for the imaginary time dynamics analogous to that for a single particle.

12.1 More on Gaussian processes

12.1.1 Indexing by an inner product space

We consider Gaussian processes indexed by a real vector space $\mathcal{S}$. We are particularly interested in the case when $\mathcal{S}$ is the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ but proceed generally. The covariance is a function $C : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$, which we suppose is an inner product on $\mathcal{S}$. For any collection $h_1, \dots, h_n$ of linearly independent elements of $\mathcal{S}$ define an $n \times n$ matrix $\hat{C}$ by

$$\hat{C}_{ij} = C(h_i, h_j) \tag{12.1}$$

Then $\hat{C}$ is positive definite since for any $s_1, \dots, s_n$

$$\sum_{ij} s_i s_j \hat{C}_{ij} = C\left(\sum_i s_i h_i,\ \sum_j s_j h_j\right) \geq 0 \tag{12.2}$$

and since it equals zero iff $\sum_i s_i h_i = 0$ which occurs iff $s_i = 0$. We define a *Gaussian random process with covariance C* (and mean zero) to be a probability measure space $(\mathcal{M}, \Sigma, \mu)$ and a family of random variables $\{\phi(h)\}_{h \in \mathcal{S}}$ linear in h such that for any finite collection of linearly independent elements $h_1, \dots, h_n$, the random variables $\phi(h_1), \dots, \phi(h_n)$ are jointly Gaussian with mean zero and covariance matrix $\hat{C}$. Thus

$$\begin{aligned} &E\Big(f(\phi(h_1), \dots, \phi(h_n))\Big) \\ &= (2\pi)^{-n/2} (\det \hat{C})^{-1/2} \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \exp\left(-\frac{1}{2} x \cdot \hat{C}^{-1} x\right) dx \end{aligned} \tag{12.3}$$

The characteristic function is

$$E(e^{i\phi(h)}) = \exp\left(-\frac{1}{2}C(h,h)\right) \tag{12.4}$$

and it follows that for any $h_1, \ldots, h_n \in \mathcal{H}$ (not necessarily linearly independent)

$$E\left(\exp\left(i\sum_{i=1}^{n} t_i\phi(h_i)\right)\right) = \exp\left(-\frac{1}{2}\sum_{i,j=1}^{n} t_i t_j C(h_i,h_j)\right) \tag{12.5}$$

We also have $E(\phi(h)) = 0$ and $E(\phi(h_1)\phi(h_2)) = C(h_1,h_2)$.

In fact the characteristic function is enough to determine the process. Indeed if a family of random variables $\{\phi(h)\}_{h\in\mathcal{S}}$ satisfies (12.4) and is linear in h, then (12.5) with $h_1, \ldots, h_n$ linearly independent says that the joint distribution for $\phi(h_1), \ldots, \phi(h_n)$ is Gaussian with covariance $\hat{C}_{ij} = C(h_i,h_j)$ as required.

Theorem 12.1 *Let $\mathcal{S}$ be a real vector space with inner product C. Then a Gaussian random process $\{\phi(h)\}_{h\in\mathcal{S}}$ with covariance C exists.*

Remarks

1. Because of the linearity requirement and the restriction to linearly independent elements, the existence does not follow directly from the Kolmogorov theorem.
2. Just as before (problem 11.3) the random variables $\phi(h)$ are in $L^p(\mathcal{M},\mu)$ for all $1 \le p < \infty$. Hence the same is true for polynomials in the $\phi(h)$. If we assume that Σ is the smallest σ-algebra with respect to which the $\phi(h)$ are measurable, then polynomials are dense in $L^2(\mathcal{M},\mu)$, a result we need later.[1]
3. In the proof we show more. Let $\mathcal{H}$ be the real Hilbert space which is the completion of $\mathcal{S}$ in the inner product C. We construct a family of random variables $\phi(h)$ indexed by $h \in \mathcal{H}$ with the stated properties.

Proof Pick an orthonormal basis $\{e_i\}$ for $\mathcal{H}$. For any finite collection of basis elements the matrix $C(e_i,e_j) = \delta_{ij}$ is positive definite. Hence there exists a Gaussian process $\{\phi(e_i)\}_{i=1}^{\infty}$ with identity covariance by the Kolmogorov theorem as explained in section 11.2. We have

$$E(\phi(e_i)\phi(e_j)) = C(e_i,e_j) = \delta_{ij} \tag{12.6}$$

Any $h \in \mathcal{H}$ has the expansion $h = \sum_i C(e_i,h)e_i$. Hence we define

$$\phi(h) = \sum_{i=1}^{\infty} C(e_i,h)\phi(e_i) \tag{12.7}$$

[1] For this result see Segal (1956).

The limit exists in $L^2(\mathcal{M},\mu)$ since if $h_N = \sum_{i=1}^N C(e_i,h)e_i$, then for $N > M$

$$\|\phi(h_N) - \phi(h_M)\|^2 = \|\sum_{i=M+1}^{N} C(e_i,h)\phi(e_i)\|^2 = \sum_{i=M+1}^{N} C(e_i,h)^2 \tag{12.8}$$

and this converges to zero as $N, M \to \infty$. Note that $\phi(h)$ is linear in h. Also we compute

$$\begin{aligned} E(e^{i\phi(h_N)}) &= E\left(\exp\left(i\sum_{i=1}^{N} C(e_i,h)\phi(e_i)\right)\right) \\ &= \exp\left(-\frac{1}{2}\sum_{i=1}^{N} C(e_i,h)^2\right) \\ &= \exp\left(-\frac{1}{2}C(h_N,h_N)\right) \end{aligned} \tag{12.9}$$

There is a subsequence $\phi(h_{N_j})$ that converges to $\phi(h)$ almost everywhere and passing to this subsequence and using the dominated convergence theorem we take the limit in (12.9) and conclude that $E(e^{i\phi(h)}) = \exp(-C(h,h)/2)$. This is sufficient to show that $\phi(h)$ is the desired Gaussian process. □

Theorem 12.2

$$E(\phi(h_1)\dots\phi(h_n)) = \left\{ \begin{array}{rl} 0 & n \text{ odd} \\ \sum_P \prod_{\{i,j\}\in P} C(h_i,h_j) & n \text{ even} \end{array} \right\} \tag{12.10}$$

where the sum is over pairings $P = \{\{i_1,j_1\},\dots,\{i_{n/2},j_{n/2}\}\}$ *of* $(1,\dots,n)$.

Proof Take the partial derivative $\partial^n/\partial_{t_1}\cdots\partial_{t_n}[\dots]_{t_i=0}$ of (12.5). On the left we get $i^n E(\phi(h_1)\dots\phi(h_n))$. On the right we get the coefficient of $t_1\cdots t_n$ in the power series expansion of $\exp\left(-\frac{1}{2}\sum_{i,j=1}^n t_i t_j C(h_i,h_j)\right)$. For this we can ignore terms with $i = j$. Also we can drop the factor $1/2$ and write it as a sum over unordered pairs $\{i,j\}$ (two element subsets) from $(1,\dots,n)$. Then we have

$$\begin{aligned} \exp\left(-\sum_{\{i,j\}} t_i t_j C(h_i,h_j)\right) &= \prod_{\{i,j\}} \exp\left(-t_i t_j C(h_i,h_j)\right) \\ &= \prod_{\{i,j\}} \left(1 - t_i t_j C(h_i,h_j) + \dots\right) \\ &= 1 + \sum_{Q} \prod_{\{i,j\}\in Q} \left(-t_i t_j C(h_i,h_j)\right) + \dots \end{aligned} \tag{12.11}$$

where the sum over Q is over collections of pairs $\{i,j\}$ from $(1,\dots,n)$. But only collections which give a partition of $(1,\dots,n)$ will contribute. This is only possible if n is even in which case we get the announced sum over pairings. There is also a factor $(-1)^{n/2}$ which matches the i^n on the left. Hence the result. □

12.1.2 Wick monomials

We next introduce Wick monomials.[2] These are polynomials in the random variables $\phi(h)$ with the property that monomials of different degree are orthogonal. First define

$$: e^{i\phi(h)} :_C = e^{i\phi(h)} \exp\left(\frac{1}{2}C(h,h)\right) \tag{12.12}$$

This is defined so that

$$E(: e^{i\phi(h)} :_C) = 1 \tag{12.13}$$

Then we define *Wick monomials* by

$$\begin{aligned} : \phi(h_1)\dots\phi(h_n) :_C &= \frac{1}{i^n}\frac{\partial^n}{\partial t_1 \dots \partial t_n}\left[: \exp\left(i\sum_i t_i\phi(h_i)\right) :_C \right]_{t=0} \\ &= \frac{1}{i^n}\frac{\partial^n}{\partial t_1 \dots \partial t_n}\left[\exp\left(i\sum_i t_i\phi(h_i) + \frac{1}{2}\sum_{ij} t_i t_j C(h_i,h_j)\right)\right]_{t=0} \end{aligned} \tag{12.14}$$

This looks like the generating function for Hermite polynomials and indeed Wick monomials are a generalization of Hermite polynomials. We have explicitly

$$\begin{aligned} : \phi(h) :_C &= \phi(h) \\ : \phi(h_1)\phi(h_2) :_C &= \phi(h_1)\phi(h_2) - C(h_1,h_2) \\ : \phi(h_1)\phi(h_2)\phi(h_3) :_C &= \phi(h_1)\phi(h_2)\phi(h_3) - \phi(h_1)C(h_2,h_3) \\ &\quad - \phi(h_2)C(h_1,h_3) - \phi(h_3)C(h_1,h_2) \end{aligned} \tag{12.15}$$

and in general

$$: \phi(h_1)\dots\phi(h_n) :_C = \phi(h_1)\dots\phi(h_n) + \text{ lower order terms} \tag{12.16}$$

It follows that Wick monomials span the dense subspace of polynomials in $L^2(\mathcal{M},\mu)$. Also note that $: \phi(h_1)\dots\phi(h_n) :_C$ is linear in each h_i and is invariant under permutations of the h_i.

Lemma 12.1

1. *For a single Wick monomial*

$$E\Big(: \phi(h_1)\dots\phi(h_n) :_C \Big) = 0 \tag{12.17}$$

2. *For a pair of Wick monomials*

$$\begin{aligned} &E\Big(: \phi(h_1)\dots\phi(h_n) :_C : \phi(g_1)\dots\phi(g_m) :_C \Big) \\ &= \left\{ \begin{array}{rl} 0 & n \neq m \\ \sum_\pi C(h_1, g_{\pi(1)})\dots C(h_n, g_{\pi(n)}) & n = m \end{array} \right\} \end{aligned} \tag{12.18}$$

[2] There is a connection with Wick ordering explained in the next section

where the sum is over permutations π of $(1, 2, \dots, n)$.

Proof The first follows by taking derivatives of

$$E\Big(: \exp(i \sum_i t_j \phi(h_j)) :_C \Big) = 1 \tag{12.19}$$

For the second note the identity.

$$: e^{i\phi(h)} :_C : e^{i\phi(g)} :_C =: e^{i\phi(h+g)} :_C \, e^{-C(h,g)} \tag{12.20}$$

Let $h = \sum_{i=1}^n t_i h_i$ and $g = \sum_{j=1}^m s_j g_j$ and take the expectation which gives

$$\begin{aligned} & E\left(: \exp\big(i \sum_i t_i \phi(h_i)\big) :_C : \exp\big(i \sum_j s_j \phi(g_j)\big) :_C \right) \\ & = \exp\left(- \sum_{ij} t_i s_j C(h_i, g_j) \right) \end{aligned} \tag{12.21}$$

Now take a single derivative in each of t_i, s_j at $t_i = s_j = 0$. On the left we get i^{n+m} times the desired expectation. On the right we get the coefficient of $t_1 \cdots t_n s_1 \cdots s_m$ in a power series expansion in s, t. To identify this coefficient we write

$$\begin{aligned} \exp\left(- \sum_{(i,j)} t_i s_j C(h_i, g_j) \right) &= \prod_{(i,j)} \exp\Big(- t_i s_j C(h_i, g_j) \Big) \\ &= \prod_{(i,j)} \big(1 - t_i s_j C(h_i, g_j) + \dots \big) \\ &= 1 + \sum_{\Pi} \prod_{(i,j) \in \Pi} \big(- t_i s_j C(h_i, g_j) \big) + \dots \end{aligned} \tag{12.22}$$

where the sum over Π is over collections of elements (i, j) from $(1, \dots, n) \times (1, \dots, m)$. But only collections in which each element of $(1, \dots, n)$ appears exactly once in the first position and each element of $(1, \dots, m)$ appears exactly once in the second position will contribute. This is only possible if $n = m$ and then the sum over such terms can be identified with a sum over permutations π of $(1, \dots, n)$. Hence we get $(-1)^n = i^{2n}$ times the right side of (12.18). □

Problem 12.1 Show that

$$\begin{aligned} \phi(h) : \phi(h_1) \dots \phi(h_n) :_C =: \phi(h)\phi(h_1) \dots \phi(h_n) :_C \\ + \sum_{j=1}^n C(h, h_j) : \phi(h_1) \dots \widehat{\phi(h_j)} \dots \phi(h_n) :_C \end{aligned} \tag{12.23}$$

where the hat on $\phi(h_j)$ means "omit this factor".

Problem 12.2 Establish the identity

$$: \phi(h_1) \dots \phi(h_n) :_C = \sum_Q \prod_{\{i,j\} \in Q} \Big(- C(h_i, h_j) \Big) \prod_{k \notin Q} \phi(h_k) \tag{12.24}$$

where Q is a (possibly empty) collection of pairs from $(1, \dots, n)$. Then establish the special case

$$: \phi(h)^n :_C = \sum_{j=0}^{[n/2]} \frac{(-1)^j n!}{(n-2j)!\, j!\, 2^j} C(h,h)^j \phi(h)^{n-2j} \tag{12.25}$$

Problem 12.3 Let $\{h_i^\alpha\}$ be a collection from $\mathcal{S}$ indexed by pairs (α, i) with $1 \le \alpha \le r$ and $1 \le i \le n_\alpha$. Show that

$$E\left(\prod_{\alpha=1}^{r} : \prod_{i=1}^{n_\alpha} \phi(h_i^\alpha) : \right) = \sum_G \prod_{\{(\alpha,i),(\beta,j)\} \in G} C(h_i^\alpha, h_j^\beta) \tag{12.26}$$

where the sum is over all graphs G on r vertices with legs (α, i) at the αth vertex and lines $\{(\alpha, i); (\beta, j)\}$. Each leg must belong to exactly one line and the lines must have $\alpha \neq \beta$, that is lines cannot join legs at the same vertex.

12.1.3 Realization on $\mathcal{S}'$

Depending on the real vector space $\mathcal{S}$ and the covariance C there may be more concrete representations of the Gaussian process. In particular suppose that $\mathcal{S}$ is the Schwartz space $\mathcal{S}(\mathbb{R}^d)$. It turns out that in this case we can take our basic measure space to be (Q, Σ, μ_C) where $Q = \mathcal{S}'(\mathbb{R}^d)$ is the space of real tempered distributions, Σ is the σ-algebra generated by the functions $q \to < q, f >, f \in \mathcal{S}(\mathbb{R}^d)$, and μ_C is a Gaussian measure. The random variables $\phi(f)$ are given by the evaluation map

$$\big(\phi(f)\big)(q) = < q, f > \qquad\qquad q \in Q, \tag{12.27}$$

so the distributions q are the fields. With this choice, expectations will be written out explicitly as

$$E(F) = \int_Q F \, d\mu_C \tag{12.28}$$

An advantage of this representation is that there is a natural definition of derivatives with respect to the field. For any function F on $Q = \mathcal{S}'(\mathbb{R}^d)$ we define the derivative along $h \in Q$ by

$$(\nabla_h F)(q) = \frac{d}{dt} [F(q + th)]_{t=0} \tag{12.29}$$

if it exists. We consider in particular functions on Q of the form

$$F = \hat{F}(\phi(f_1), \ldots, \phi(f_n)) \tag{12.30}$$

where $\hat{F}$ is a complex Borel function on $\mathbb{R}^n$ and $f_i \in \mathcal{S}(\mathbb{R}^d)$. Equivalently we can write

$$F(q) = \hat{F}(< q,f_1 >, \ldots, < q,f_n >) \tag{12.31}$$

Such functions are called *cylinder functions*. If $\hat{F}$ is differentiable, then $\nabla_h F$ does exist and by the chain rule we have

$$\nabla_h F = \sum_{i=1}^{n} \frac{\partial \hat{F}}{\partial x_i}(\phi(f_1), \ldots, \phi(f_n)) < h,f_i > \tag{12.32}$$

Note that this is linear in h.

As a special case we can take $h = \delta_x$, the delta function at x. Then $\nabla_{\delta_x} F$ is denoted $\partial F/\partial\phi(x)$. If F has the form (12.30), then

$$\frac{\partial F}{\partial \phi(x)} = \sum_{i=1}^{n} \frac{\partial \hat{F}}{\partial x_i}(\phi(f_1), \ldots, \phi(f_n)) f_i(x) \tag{12.33}$$

Hence $\partial F/\partial\phi(x)$ is in $\mathcal{S}(\mathbb{R}^d)$ and for $h \in \mathcal{S}'(\mathbb{R}^d)$

$$\nabla_h F =< h, \frac{\partial F}{\partial \phi} >= \int h(x) \frac{\partial F}{\partial \phi(x)} dx \tag{12.34}$$

Problem 12.4 Establish the identities

$$\begin{aligned} \nabla_h \Big(\phi(f_1) \ldots \phi(f_n) \Big) &= \sum_{j=1}^{n} < h,f_j > \phi(f_1) \ldots \widehat{\phi(f_j)} \ldots \phi(f_n) \\ \nabla_h \Big(:\phi(f_1) \ldots \phi(f_n): _C \Big) &= \sum_{j=1}^{n} < h,f_j > :\phi(f_1) \ldots \widehat{\phi(f_j)} \ldots \phi(f_n): _C \end{aligned} \tag{12.35}$$

Next we develop an integration by parts formula for ∇_h.

Lemma 12.2 *Let $C(f,g) =< f, Cg >$ where C is a bijection on $\mathcal{S}(\mathbb{R}^d)$. Let $F = \hat{F}(\phi(f_1), \ldots, \phi(f_n))$ be a cylinder function on $Q = \mathcal{S}'(\mathbb{R}^d)$ with $\hat{F}$ and its partial derivatives continuous and exponentially bounded. Then for $h \in \mathcal{S}(\mathbb{R}^d)$*

$$\int_Q \nabla_h F \, d\mu_C = \int_Q F \, \phi(C^{-1}h) \, d\mu_C \tag{12.36}$$

Proof Let $e_1, \ldots, e_m$ be a basis for the subspace spanned by $f_1, \ldots, f_n$ and $C^{-1}h$ such that $C(e_i, e_j) = \delta_{ij}$. Such a basis can be constructed by the Gram-Schmidt process. We can write $F = \tilde{F}(\phi(e_1), \ldots, \phi(e_m))$ and evaluate the derivative by (12.32). Then write the integral in $\mathbb{R}^m$ by (12.3), and integrate by parts to obtain

$$
\begin{aligned}
&\int_Q \nabla_h F \, d\mu_C \\
&= \sum_i < h, e_i > \int_Q \frac{\partial \tilde{F}}{\partial x_i}(\phi(e_1), \ldots, \phi(e_m)) \, d\mu_C \\
&= \sum_i < h, e_i > (2\pi)^{-m/2} \int_{\mathbb{R}^m} \frac{\partial \tilde{F}}{\partial x_i}(x_1, \ldots, x_m) e^{-|x|^2/2} dx \\
&= \sum_i < h, e_i > (2\pi)^{-m/2} \int_{\mathbb{R}^m} \tilde{F}(x_1, \ldots, x_m) \, x_i \, e^{-|x|^2/2} dx \\
&= \sum_i < h, e_i > \int_Q \tilde{F}(\phi(e_1), \ldots, \phi(e_m)) \phi(e_i) \, d\mu_C \\
&= \int_Q F \, \phi(C^{-1}h) d\mu_C
\end{aligned}
\tag{12.37}
$$

The last step follows by $C^{-1}h = \sum_i < h, e_i > e_i$. □

Corollary 12.1 *(Integration by parts) With F, G as in the lemma*

$$
\int_Q F(\nabla_h G) \, d\mu_C = -\int_Q (\nabla_h F) \, G \, d\mu_C + \int_Q F \, G \, \phi(C^{-1}h) d\mu_C \tag{12.38}
$$

Proof In the lemma replace F by FG and use

$$
\nabla_h(FG) = (\nabla_h F)G + F(\nabla_h G) \tag{12.39}
$$

□

12.2 The Schrödinger representation

12.2.1 Definitions and equivalence

As an application of the previous section we give another representation of the time zero free scalar field. This is known as the *Schrödinger* representation and is characterized by the feature that the field operators are all multiplication operators.

For the free scalar field in space dimension d we generalize the results of section 8.2.2 for $d = 3$. The symmetric Fock space is $\mathcal{F}^+(\mathcal{H}_0)$ where $\mathcal{H}_0 = L^2(\mathbb{R}^d, d\mathbf{p})$. On finite particle vectors $\mathcal{D}_0 \subset \mathcal{F}^+(\mathcal{H}_0)$ the field operator is defined as in (8.55) and is now denoted $\phi_0(t, h)$. Thus

$$
\phi_0(t, h) = a\left(\frac{e^{i\omega t}\tilde{h}}{\sqrt{2\omega}}\right) + a^*\left(\frac{e^{i\omega t}\tilde{h}}{\sqrt{2\omega}}\right) \tag{12.40}
$$

where $h \in \mathcal{S}(\mathbb{R}^d)$ is real and $\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}$. The field and its time derivative at $t = 0$ are

$$\begin{aligned}\phi_0(h) &= a\left((2\omega)^{-1/2}\tilde{h}\right) + a^*\left((2\omega)^{-1/2}\tilde{h}\right)\\ \pi_0(h) &= -ia\left((\omega/2)^{1/2}\,\tilde{h}\right) + ia^*\left((\omega/2)^{1/2}\,\tilde{h}\right)\end{aligned} \tag{12.41}$$

and these satisfy the canonical commutation relations. We will also want to consider Wick ordered products $: \phi_0(h_1)\cdots\phi_0(h_n) :$ on the vacuum Ω_0. These satisfy the identity

$$: \phi_0(h_1)\cdots\phi_0(h_n) : \Omega_0 = a^*\left((2\omega)^{-1/2}\tilde{h}_1\right)\cdots a^*\left((2\omega)^{-1/2}\tilde{h}_n\right)\Omega_0 \tag{12.42}$$

This follows since Wick ordering means move annihilation operators to the right, and annihilation operators on Ω_0 give zero.

To represent the $\phi_0(h)$ as multiplication operators we consider

$$\hat{\omega} = \sqrt{-\Delta + m^2} = \mathcal{F}^{-1}[\omega(\mathbf{p})]\mathcal{F} \tag{12.43}$$

and introduce the Gaussian process $\phi(h)$ indexed by real $h \in \mathcal{S}(\mathbb{R}^d)$ with mean zero and covariance $(2\hat{\omega})^{-1}$. Thus[3]

$$E(\phi(g)\phi(h)) = < g, (2\hat{\omega})^{-1}h > = (\tilde{g}, (2\omega)^{-1}\tilde{h}) \tag{12.44}$$

Theorem 12.3 *Let $\phi(h)$ be a Gaussian process with covariance $(2\hat{\omega})^{-1}$ on a measure space $(\mathcal{M}, \Sigma, \mu)$. There is a unitary operator $V : \mathcal{F}^+(\mathcal{H}_0) \to L^2(\mathcal{M}, \mu)$ such that $V\Omega_0 = 1$ and*

$$V\Big(: \phi_0(h_1)\cdots\phi_0(h_n) : \Omega_0\Big) =: \phi(h_1)\cdots\phi(h_n) :_{(2\hat{\omega})^{-1}} \tag{12.45}$$

Proof First define V on complex linear combinations of the vectors (12.42). These can be written in the form $\sum_\alpha c_\alpha : \phi_0(h_{\alpha_1})\cdots\phi_0(h_{\alpha_n}) : \Omega_0$ with $\alpha = (\alpha_1, \ldots, \alpha_n)$ in some index set and complex c_α. We want to define

$$V\left(\sum_\alpha c_\alpha : \phi_0(h_{\alpha_1})\cdots\phi_0(h_{\alpha_n}) : \Omega_0\right) = \sum_\alpha c_\alpha : \phi(h_{\alpha_1})\cdots\phi(h_{\alpha_n}) :_{(2\hat{\omega})^{-1}} \tag{12.46}$$

To see this is well-defined we first claim that both vectors have the same norm. For this it suffices to show that

$$\begin{aligned}&(: \phi_0(g_1)\cdots\phi_0(g_m) : \Omega_0,\ : \phi_0(h_1)\cdots\phi_0(h_n) : \Omega_0)\\ &= E\left(: \phi(g_1)\cdots\phi(g_m) :_{(2\hat{\omega})^{-1}}\ : \phi(h_1)\cdots\phi(h_n) :_{(2\hat{\omega})^{-1}}\right)\end{aligned} \tag{12.47}$$

In fact each side is zero if $n \neq m$. If $n = m$ the left side of (12.47) is computed as

$$\sum_\pi (\tilde{g}_1, (2\omega)^{-1}\tilde{h}_{\pi(1)})\cdots(\tilde{g}_n, (2\omega)^{-1}\tilde{h}_{\pi(n)}) \tag{12.48}$$

[3] Here and elsewhere we write $< g, h > = \int g(x)h(x)dx$ as a reminder that it is the real inner product. But since g, h are real it is the same as $(g, h) = \int \overline{g(x)}h(x)dx$.

where the sum is over permutations π of $(1,\dots,n)$. This follows from (12.42) and the commutation relations for a, a^*, or more directly from (5.60). This is the same as the right side of (12.47) by (12.18) and (12.44).

Since both sides have the same norm, it follows that if a sum is the zero vector, then it is sent to the zero vector. Hence if a vector has two different representations they are sent to the same place. Thus the mapping is well-defined.

The mapping has a dense domain since vectors $a^*(f_1)\cdots a^*(f_n)\Omega_0$ with $f_i \in \mathcal{S}(\mathbb{R}^d)$ span a dense subspace of Fock space. Also the range is the subspace of all polynomials in the $\phi(h)$ which is dense in $L^2(\mathcal{M},\mu)$. Since V is norm preserving with dense domain and dense range it extends to a unitary operator (theorem 1.3). □

Remark Let $H_0 = d\Gamma(\omega)$ be the free Hamiltonian on Fock space so imaginary time evolution is $e^{-H_0 t} = \Gamma(e^{-\omega t})$. Then by (12.42) and $e^{-H_0 t}a^*(f) = a^*(e^{-\omega t}f)e^{-H_0 t}$ or directly from (5.60)

$$e^{-H_0 t} : \phi_0(h_1)\cdots\phi_0(h_n) : \Omega_0 =: \phi_0(e^{-\hat{\omega}t}h_1)\cdots\phi_0(e^{-\hat{\omega}t}h_n) : \Omega_0 \tag{12.49}$$

Note that $e^{-\hat{\omega}t}h_j$ is still real. Then $Ve^{-H_0 t}V^{-1}$ defines a contraction on $L^2(\mathcal{M},\mu)$, also denoted $e^{-H_0 t}$, and we have

$$e^{-H_0 t} : \phi(h_1)\cdots\phi(h_n) :_{(2\hat{\omega})^{-1}} =: \phi(e^{-\hat{\omega}t}h_1)\cdots\phi(e^{-\hat{\omega}t}h_n) :_{(2\hat{\omega})^{-1}} \tag{12.50}$$

Problem 12.5

1. Show that

$$\begin{aligned} &\phi_0(h) : \phi_0(h_1)\cdots\phi_0(h_n) : \Omega_0 =: \phi_0(h)\phi_0(h_1)\cdots\phi_0(h_n) : \Omega_0 \\ &+ \sum_{j=1}^{n} < h, (2\hat{\omega})^{-1}h_j > : \phi_0(h_1)\cdots\widehat{\phi_0(h_j)}\cdots\phi_0(h_n) : \Omega_0 \end{aligned} \tag{12.51}$$

2. Show that $\phi(h) = V\phi_0(h)V^{-1}$ on polynomials.

Problem 12.6 Show that

$$: \phi(h_1)\cdots\phi(h_n) :_{(2\hat{\omega})^{-1}} = V : \phi_0(h_1)\cdots\phi_0(h_n) : V^{-1} \tag{12.52}$$

12.2.2 The CCR

If the Gaussian process $\phi(h)$ is realized on the space $(Q, \Sigma, \mu_{(2\hat{\omega})^{-1}})$ with $Q = \mathcal{S}'(\mathbb{R}^d)$, then we can construct a representation of the CCR. This is a representation on $L^2(Q, \mu_{(2\hat{\omega})^{-1}})$ in which $\phi(h)$ is multiplication by $(\phi(h))(q) =< q, h >$ and $\pi(h)$ is a derivative operator. However we cannot take $\pi(h) = -i\nabla_h$ since this would not be symmetric with respect to the Gaussian measure. Instead we take for real $h \in \mathcal{S}(\mathbb{R}^d)$

$$\begin{aligned} \phi(h) &= [\phi(h)] \\ \pi(h) &= -i\nabla_h + i[\phi(\hat{\omega}h)] \end{aligned} \tag{12.53}$$

with the polynomials as the domain.

Lemma 12.3 *$\phi(h), \pi(h)$ are densely defined symmetric operators on the Hilbert space $L^2(Q, \mu_{(2\hat{\omega})^{-1}})$ which satisfy the canonical commutation relations.*

Proof $\phi(h)$ is symmetric. To see that $\pi(h)$ is symmetric use the integration by parts formula (12.38) with $C = (2\hat{\omega})^{-1}$ to obtain

$$\begin{aligned} (F, \pi(h)G) &= (F, (-i\nabla_h)G) + i(F, \phi(\hat{\omega}h)G) \\ &= ((-i\nabla_h)F, G) - i(F, \phi(\hat{\omega}h)G) \\ &= (\pi(h)F, G) \end{aligned} \tag{12.54}$$

For the commutators we compute that $[\phi(g), -i\nabla_h] = i < g, h >$. It follows that $[\phi(g), \pi(h)] = i < g, h >$ and $[\pi(g), \pi(h)] = 0$. Since $[\phi(g), \phi(h)] = 0$ is trivial, the proof is complete. □

The next problem combined with problem 12.5 shows that the representation $\phi(h), \pi(h)$ of the CCR on $L^2(Q, \mu_{(2\hat{\omega})^{-1}})$ is unitarily equivalent to the representation $\phi_0(h), \pi_0(h)$ on Fock space.

Problem 12.7 Show that $\pi(h) = V\pi_0(h)V^{-1}$ on polynomials.

12.3 Path integrals – free fields

We continue to consider the free scalar field. The Schrödinger representation opens the door for the representation of the imaginary time dynamics $e^{-H_0 t}$ in terms of random paths as in the Feynman–Kac formula. Recall from section 11.5 that the imaginary time dynamics for the harmonic oscillator with Hamiltonian $1/2(p^2 + x^2)$ could be represented by a Gaussian process X_t with covariance $E(X_{t_1}X_{t_2}) = e^{-|t_2 - t_1|/2}/2$. Our Hamiltonian $H_0 = 1/2 \int (\pi^2 + \phi\hat{\omega}^2\phi)$ is an infinite-dimensional analog of the harmonic oscillator. This suggests a similar representation which we now explain.

For $t \in \mathbb{R}$ and real $h \in \mathcal{S}(\mathbb{R}^d)$ let $\phi(t, h)$ be the Gaussian process with mean zero and covariance $E(\phi(t_1, h_1)\phi(t_2, h_2))$ given by

$$\begin{aligned} C(t_1, h_1; t_2, h_2) &= \left\langle h_1, \left(\frac{e^{-|t_2 - t_1|\hat{\omega}}}{2\hat{\omega}} \right) h_2 \right\rangle \\ &= \int \overline{\tilde{h}_1(\mathbf{p})} \, \frac{e^{-|t_2 - t_1|\omega(\mathbf{p})}}{2\omega(\mathbf{p})} \, \tilde{h}_2(\mathbf{p}) \, d\mathbf{p} \end{aligned} \tag{12.55}$$

This is positive definite since it can also be written in the form

$$C(t_1,h_1;t_2,h_2)=\frac{1}{2\pi}\int \overline{\tilde{h}_1(\mathbf{p})}\,\frac{e^{ip_0(t_1-t_2)}}{p_0^2+|\mathbf{p}|^2+m^2}\,\tilde{h}_2(\mathbf{p})\,dp_0\,d\mathbf{p} \tag{12.56}$$

To see that this is the same, evaluate the p_0 integral by closing the contour in the upper or lower half plane depending on the sign of t_1-t_2. The Gaussian process $\phi(t,h)$ exists by the Kolmogorov theorem as in our previous discussions. We also give an alternate construction shortly.

Note that for fixed t the random variables $\phi(t,h)$ are Gaussian with covariance $(2\hat{\omega})^{-1}$ and so are a realization of our basic scalar field on $\mathbb{R}^d$. Thus $t\to\phi(t,h)$ is a random path through random scalar fields. Furthermore our basic Hilbert space is square-integrable functions of these fixed time fields and so this structure is imbedded at various times in a larger Hilbert space of time dependent fields. This leads to a Feynman–Kac formula. The details are as follows:

Theorem 12.4 *Let $\phi(h)$ be a Gaussian process indexed by $h\in\mathcal{S}(\mathbb{R}^d)$ with covariance $<h_1,(2\hat{\omega})^{-1}h_2>$ on a measure space $(\mathcal{M},\Sigma,\mu)$. Furthermore let $\phi(t,h)$ be a Gaussian process indexed by $(t,h)\in\mathbb{R}\times\mathcal{S}(\mathbb{R}^d)$ with covariance $C(t_1,h_1;t_2,h_2)$ on a measure space $(\mathcal{M}',\Sigma',\mu')$. Then*

1. *For each t there is an isometry $J_t: L^2(\mathcal{M},\mu)\to L^2(\mathcal{M}',\mu')$ such that $J_t(1)=1$ and*

$$J_t\left(:\phi(h_1)\cdots\phi(h_n):_{(2\hat{\omega})^{-1}}\right)=:\phi(t,h_1)\cdots\phi(t,h_n):_C \tag{12.57}$$

2. *Let $F,G\in L^2(\mathcal{M},\mu)$ be states of the scalar field. Then for $t\geq 0$*

$$(G,e^{-H_0t}F)=E(\overline{J_0G}\,J_tF) \tag{12.58}$$

Proof For the first part we follow the strategy of theorem 12.3. We want to define the map by

$$J_t\left(\sum_\alpha c_\alpha:\phi(h_{\alpha_1})\dots\phi(h_{\alpha_n}):_{(2\hat{\omega})^{-1}}\right)=\sum_\alpha c_\alpha:\phi(t,h_{\alpha_1})\dots\phi(t,h_{\alpha_n}):_C \tag{12.59}$$

This is well-defined if both vectors have the same norm and this follows from

$$\begin{aligned}&E\Big(:\phi(g_1)\dots\phi(g_m):_{(2\hat{\omega})^{-1}}:\phi(h_1)\dots\phi(h_n):_{(2\hat{\omega})^{-1}}\Big)\\&=E\Big(:\phi(t,g_1)\dots\phi(t,g_m):_C:\phi(t,h_1)\dots\phi(t,h_n):_C\Big)\end{aligned} \tag{12.60}$$

Each side is evaluated by (12.18) and the result follows from the equal time identity

$$<g,(2\hat{\omega})^{-1}h>=C(t,g;t,h) \tag{12.61}$$

Hence J_t is well defined. Since the domain is dense and it is norm preserving it extends to an isometry.

For the second point first take

$$G=:\phi(g_1)\dots\phi(g_n):_{(2\hat{\omega})^{-1}}\qquad F=:\phi(h_1)\dots\phi(h_n):_{(2\hat{\omega})^{-1}} \tag{12.62}$$

Then we have using (12.50) and (12.18)

$$\begin{aligned}
&(G, e^{-H_0 t}F)\\
&= E\left(:\phi(g_1)\dots\phi(g_n):_{(2\hat{\omega})^{-1}} : \phi(e^{-t\hat{\omega}}h_1)\dots\phi(e^{-t\hat{\omega}}h_n):_{(2\hat{\omega})^{-1}}\right)\\
&= \sum_{\pi}\left\langle g_1, \left(\frac{e^{-t\hat{\omega}}}{2\hat{\omega}}\right)h_{\pi(1)}\right\rangle\cdots\left\langle g_n, \left(\frac{e^{-t\hat{\omega}}}{2\hat{\omega}}\right)h_{\pi(n)}\right\rangle \qquad (12.63)\\
&= E\Big(:\phi(0,g_1)\dots\phi(0,g_n):_C : \phi(t,h_1)\dots\phi(t,h_n):_C\Big)\\
&= E(\overline{J_0 G}\, J_t F)
\end{aligned}$$

The same holds for complex linear combinations of such vectors since monomials of different degree are orthogonal. This is a dense domain and since both sides of the equation are continuous bilinears on $L^2(\mathcal{M}, \mu)$, the result follows. □

Problem 12.8 Show that $J_t[\phi(h)] = [\phi(t,h)]J_t$ and hence

$$J_t\Big(\phi(h_1)\cdots\phi(h_n)\Big) = \phi(t_1,h_1)\cdots\phi(t_n,h_n) \qquad (12.64)$$

Remarks Starting with our Gaussian process $\phi(t,h)$ with mean zero and covariance $C(t_1,h_1,t_2,h_2)$ we can consider real test functions $f \in \mathcal{S}(\mathbb{R}^{d+1})$ and define[4]

$$\phi(f) = \int \phi(t, f(t,\cdot))dt \qquad (12.65)$$

Then the $\phi(f)$ are Gaussian with covariance from (12.56)

$$\begin{aligned}
E(\phi(f_1)\phi(f_2)) &= \int C(t_1, f(t_1,\cdot),\ t_2, f(t_2,\cdot))\, dt_1 dt_2\\
&= \int_{\mathbb{R}^{d+1}} \overline{\tilde{f}_1(p)}\, \frac{1}{p^2+m^2}\, \tilde{f}_2(p)\, dp \qquad (12.66)\\
&= < f_1, (-\Delta + m^2)^{-1} f_2 >
\end{aligned}$$

In fact we could have started with a Gaussian process $\phi(f)$ indexed by $\mathcal{S}(\mathbb{R}^{d+1})$ with covariance $(-\Delta+m^2)^{-1}$. The sharp time fields can then be recovered as follows. As in the proof of theorem 12.1 the $\phi(f)$ are naturally defined for f in the completion of $\mathcal{S}(\mathbb{R}^{d+1})$ in the norm $(f, (-\Delta + m^2)^{-1} f)^{1/2}$. This space can be identified as the Sobolev space

$$\mathcal{H}^{-1}(\mathbb{R}^{d+1}) = \{f \in \mathcal{S}'(\mathbb{R}^{d+1}) : \int |\tilde{f}(p)|^2 (p^2+m^2)^{-1} dp < \infty\} \qquad (12.67)$$

For $h \in \mathcal{S}(\mathbb{R}^d)$ the Fourier transform of the distribution $\delta_t \otimes h$ is the function $(2\pi)^{-1/2} e^{-ip_0 t}\tilde{h}(\mathbf{p})$. Since $\int |\tilde{h}(\mathbf{p})|^2 (p_0^2 + |\mathbf{p}|^2 + m^2)^{-1} dp$ is finite, $\delta_t \otimes h$ is in the Sobolev space and so $\phi(\delta_t \otimes h)$ is defined. If we set $\phi(t,h) = \phi(\delta_t \otimes h)$, we get Gaussian fields with covariance $C(t_1,h_1,t_2,h_2)$.

[4] Formally $\phi(t,h) = \int \phi(t,\mathbf{x})h(\mathbf{x})$ and so formally $\phi(f) = \int \phi(t,\mathbf{x})f(t,\mathbf{x})$.

Thus the basic free dynamics are encoded in a Gaussian process with covariance $(-\Delta + m^2)^{-1}$. Expectations for this process can be thought of as integrals with respect to the formal measure[5], which is a constant times

$$\begin{aligned} &e^{-S(\phi)} d\phi \\ &\equiv \exp\left(-\frac{1}{2}\int \left(\partial\phi(x)\cdot\partial\phi(x) + m^2\phi(x)^2\right)dx\right) \prod_{x\in\mathbb{R}^{d+1}} d\phi(x) \end{aligned} \tag{12.68}$$

Note that $S(\phi)$ is just the classical action (7.82) (at $\lambda = 0$ and imaginary time). One expects that integrals are dominated by the minima of $S(\phi)$ which come at solutions of $(-\Delta + m^2)\phi = 0$. This is just the Klein–Gordon equation we started with (now at imaginary time). Thus the integral has a leading contribution from the classical solutions, but there are also quantum corrections. This picture manifests itself in other models as well.

12.4 Vacuum correlation functions

We now study correlation functions which are expectation values of products of field operators in some distinguished state. They are of interest because all information about a model can be recovered from them. They are particularly important for nonlinear field theories because they are easier to control than states, fields, or Hamiltonians directly. We explain these points in more detail in the next chapter.

Here we continue with the free scalar field $\phi_0(t,h)$ defined by (12.40). Vacuum correlation functions are defined for $t_i \in \mathbb{R}, h_i \in \mathcal{S}(\mathbb{R}^d)$ by

$$\left(\Omega_0, \phi_0(t_1,h_1)\cdots\phi_0(t_n,h_n)\Omega_0\right) \tag{12.69}$$

Since $\phi_0(t,h) = e^{iH_0t}\phi_0(h)e^{-iH_0t}$ and $e^{-iH_0t}\Omega_0 = \Omega_0$, this can also be written

$$\left(\Omega_0, \phi_0(h_1)e^{-iH_0(t_1-t_2)}\phi_0(h_2)\ldots\phi_0(h_{n-1})e^{-iH_0(t_{n-1}-t_n)}\phi_0(h_n)\Omega_0\right) \tag{12.70}$$

Now suppose we go to imaginary time replacing each t by it. Then we have

$$\begin{aligned} &S(t_1,h_1,\ldots,t_n,h_n) \\ &= \left(\Omega_0, \phi_0(h_1)e^{-(t_2-t_1)H_0}\phi_0(h_2)\ldots\phi_0(h_{n-1})e^{-(t_n-t_{n-1})H_0}\phi_0(h_n)\Omega_0\right) \end{aligned} \tag{12.71}$$

Now we impose the restriction that $t_{i+1} - t_i \geq 0$ so we can deal with bounded operators e^{-tH_0} for $t \geq 0$ rather than unbounded operators for $t < 0$. In fact for complex t with $\mathrm{Re}\, t > 0$ the operators e^{-tH_0} are bounded and analytic (by the spectral theorem and the positivity of H_0). The expression (12.71) is also analytic in

[5] This can be made precise if we approximate $\mathbb{R}^{d+1}$ by a finite lattice, say $\epsilon\mathbb{Z}^{d+1}/L\mathbb{Z}^{d+1}$ with ϵ small and L large.

$\mathrm{Re}(t_{i+1}-t_i) > 0$, and the real time correlation functions (12.70) are boundary values as $\mathrm{Re}(t_{i+1}-t_i) \to 0$. The $S(t_1,h_1,\dots,t_n,h_n)$ are called the *Schwinger functions*.

The result we are after is that the Schwinger functions are the moments for the Gaussian process $\phi(t,h)$, which we have been discussing.

Theorem 12.5 *Let $\phi(t,h)$ be the Gaussian process indexed by $\mathbb{R}\times\mathcal{S}(\mathbb{R}^d)$ with covariance $C(t_1,h_1;t_2,h_2)$ defined in (12.55), (12.56). Then for times $t_1 \le t_2 \cdots \le t_n$ the Schwinger functions for the free scalar field satisfy*

$$S(t_1,h_1,\dots,t_n,h_n) = E\Big(\phi(t_1,h_1)\dots\phi(t_n,h_n)\Big) \tag{12.72}$$

Proof Insert $\phi_0(h_1) = a((2\omega)^{-1/2}\tilde{h}_1) + a^*((2\omega)^{-1/2}\tilde{h}_1)$ in the expression (12.71). Move the creation operator to the left where it becomes an annihilation operator and gives zero on Ω_0. Move the annihilation operator to the right using $a(f)e^{-H_0 s} = e^{-H_0 s}a(e^{-\omega s}f)$ and the commutation relations for a, a^* until it reaches the Ω_0 where it gives zero. This yields the identity

$$\begin{aligned} &S(t_1,h_1,\dots,t_n,h_n) \\ &= \sum_{j=2}^{n}\left(\tilde{h}_1, \frac{e^{-(t_j-t_1)\omega}}{2\omega}\tilde{h}_j\right) S(t_2,h_2,\dots,\widehat{t_j,h_j},\dots,t_n,h_n) \end{aligned} \tag{12.73}$$

The inner product here is identified as $C(t_1,h_1,t_j,h_j)$. Iterating this relation we find that $S(t_1,h_1,\dots,t_n,h_n)$ is zero if n is odd and if n is even is given as a sum over pairings P of $(1,\dots,n)$

$$S(t_1,h_1,\dots,t_n,h_n) = \sum_{P}\prod_{\{i,j\}\in P} C(t_i,h_i;t_j,h_j) \tag{12.74}$$

But this is the same as $E(\phi(t_1,h_1)\dots\phi(t_n,h_n))$ by (12.10). □

Remark These results can be generalized. Working in the Schrödinger representation on a measure space $(\mathcal{M},\Sigma,\mu)$, let $R_1,\dots,R_n$ belong to some subspace of $L^2(\mathcal{M},\mu)$. We want to assert that for $t_1 \le t_2 \cdots \le t_n$

$$\begin{aligned} &E(R_1 e^{-(t_2-t_1)H_0}R_2\cdots R_{n-1}e^{-H_0(t_n-t_{n-1})}R_n) \\ &= E\Big((J_{t_1}R_1)\cdots(J_{t_n}R_n)\Big) \end{aligned} \tag{12.75}$$

If $n = 2$, this is (12.58). If $R_j = \phi(h_j)$, this is the result (12.72) just established. (Recall that $\phi_0(h) = \phi(h)$ and $\Omega_0 = 1$ under the identification of Fock space with $L^2(\mathcal{M},\mu)$.) Since we allow coinciding times, the result also holds for monomials $R = \phi(h_1)\cdots\phi(h_n)$. Since both sides are linear, it then holds for R = polynomial. By approximating with polynomials the result can be extended to $R \in L^\infty(\mathcal{M},\mu)$ or $R \in \cap_{p<\infty}L^p(\mathcal{M},\mu)$, but we do not go into details.

12.5 Thermal correlation functions

We continue with the free scalar field. Suppose now we want to represent not the vacuum correlation functions, but thermal correlation functions in the grand canonical ensemble at chemical potential $\mu = 0$ and inverse temperature $\beta > 0$. These are expectations $< \cdots >_\beta$ of products of field operators of the form

$$< \cdots >_\beta = \frac{Tr([\cdots]e^{-\beta H_0})}{Tr(e^{-\beta H_0})} \tag{12.76}$$

Working on $\mathbb{R}^d$, the operator $e^{-\beta H_0}$ is not trace class, so this must be interpreted as a limit from a sequence of tori $\mathbb{R}^d/L\,\mathbb{Z}^d$ as $L \to \infty$ just as in the non-relativistic case; see section 6.3.

The KMS condition (6.65) for commuting observables can be interpreted as a statement of periodicity in imaginary time. This suggests that we try to represent the imaginary time correlation functions on the cylinder $S_\beta \times \mathbb{R}^d$ rather than $\mathbb{R}^{d+1}$. Here $S_\beta = \mathbb{R}/\beta\mathbb{Z}$ is a circle of circumference β. Indeed let $C_\beta = (-\Delta + m^2)^{-1}$ on $L^2(S_\beta \times \mathbb{R}^d)$. The periodicity in t means we replace the Fourier transform with a Fourier series and so instead of (12.56) we have for real $h_i \in \mathcal{S}(\mathbb{R}^d)$

$$C_\beta(t_1, h_1; t_2, h_2) = \beta^{-1} \sum_{p_0 \in (2\pi/\beta)\mathbb{Z}} \int \overline{\tilde{h}_1(\mathbf{p})} \frac{e^{-ip_0(t_1-t_2)}}{p_0^2 + |\mathbf{p}|^2 + m^2} \tilde{h}_2(\mathbf{p}) d\mathbf{p} \tag{12.77}$$

The path space representation is the following:

Theorem 12.6 *Let $\phi(t,h)$ be the Gaussian process indexed by $S_\beta \times \mathcal{S}(\mathbb{R}^d)$ with covariance $C_\beta(t_1, h_2, t_2, h_2)$. The thermal correlation functions for the free scalar field $< \phi_0(t_1,h_1)\cdots\phi_0(t_n,h_n) >_\beta$ have an analytic continuation to $0 < Im\ t_1 < \cdots < Im\ t_n < \beta$ and at points $t_j = is_j$ with $0 < s_1 < \cdots < s_n < \beta$*

$$\Big[< \phi_0(t_1,h_1)\cdots\phi_0(t_n,h_n) >_\beta \Big]_{t_j = is_j} = E\Big(\phi(s_1,h_1)\cdots\phi(s_n,h_n)\Big) \tag{12.78}$$

Proof We first check it for $n = 2$. To compute the left side we note that for $f, g \in \mathcal{S}(\mathbb{R}^d)$

$$\begin{aligned} < a^*(f)a(g) >_\beta &= \left(g, \frac{e^{-\beta\omega}}{(1-e^{-\beta\omega})} f\right) \\ < a(g)a^*(f) >_\beta &= \left(g, \frac{1}{(1-e^{-\beta\omega})} f\right) \end{aligned} \tag{12.79}$$

and that $< a^*(f)a^*(g) >_\beta = < a(f)a(g) >_\beta = 0$. This computation is the same as in the non-relativistic case lemma 6.1, except that we have $\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}$ instead of $|\mathbf{p}|^2/2m$ and $\mu = 0$ is now allowed.

The field operator from (12.40) then satisfies

$$\begin{aligned} &< \phi_0(t_1,h_1)\phi_0(t_2,h_2) >_\beta \\ &= \left\langle a^*\left(\frac{e^{i\omega t_1}\tilde{h}_1}{2\omega}\right) a\left(\frac{e^{i\omega t_2}\tilde{h}_2}{2\omega}\right) + a\left(\frac{e^{i\omega t_1}\tilde{h}_1}{2\omega}\right) a^*\left(\frac{e^{i\omega t_2}\tilde{h}_2}{2\omega}\right)\right\rangle_\beta \\ &= \left(\tilde{h}_2, \left(\frac{e^{(-\beta - i(t_2-t_1))\omega}}{2\omega(1-e^{-\beta\omega})}\right)\tilde{h}_1\right) + \left(\tilde{h}_1, \left(\frac{e^{i(t_2-t_1)\omega}}{2\omega(1-e^{-\beta\omega})}\right)\tilde{h}_2\right) \end{aligned} \tag{12.80}$$

Since $\omega \geq 0$, this has the analytic continuation and

$$\begin{aligned} &\Big[< \phi_0(t_1,h_1)\phi_0(t_2,h_2) >_\beta \Big]_{t_1 = is_1, t_2 = is_2} \\ &= \left(\tilde{h}_2, \left(\frac{e^{(-\beta + s_2 - s_1)\omega}}{2\omega(1-e^{-\beta\omega})}\right)\tilde{h}_1\right) + \left(\tilde{h}_1, \left(\frac{e^{-(s_2-s_1)\omega}}{2\omega(1-e^{-\beta\omega})}\right)\tilde{h}_2\right) \end{aligned} \tag{12.81}$$

On the other hand the right side of (12.78) is $E(\phi(s_1,h_1)\phi(s_2,h_2)) = C_\beta(s_1,h_1; s_2,h_2)$. But we can relate the covariance C_β on $S_\beta \times \mathbb{R}^d$ to the covariance C on $\mathbb{R}^{d+1}$ defined in (12.55), (12.56) by

$$C_\beta(s_1,h_1;\ s_2,h_2) = \sum_{n\in\mathbb{Z}} C(s_1,h_1;\ s_2+n\beta,h_2) \tag{12.82}$$

To see this is true we first establish the identity for $\omega \neq 0$

$$\beta^{-1} \sum_{p_0 \in (2\pi/\beta)\mathbb{Z}} \frac{e^{ip_0 t}}{p_0^2+\omega^2} = (2\pi)^{-1} \sum_{n\in\mathbb{Z}} \int \frac{e^{ip_0(t+n\beta)}}{p_0^2+\omega^2} dp_0 \tag{12.83}$$

The second expression can be written $\sum_n e^{-\omega|t+n\beta|}/2\omega$, which shows that the sum over n converges. To establish the identity note that both sides are periodic functions with period β and hence define functions on S_β. Both sides satisfy the equation

$$\left(-\frac{d^2}{dt^2} + \omega^2\right) u = \sum_{n\in\mathbb{Z}} \delta(t+n\beta) \tag{12.84}$$

in the sense of distributions. Such solutions are unique, hence the identity. (This is the method of images.) Now in (12.83) let $t = s_1 - s_2$, let $\omega = \omega(\mathbf{p})$, multiply by $\overline{\tilde{h}_1(\mathbf{p})}\tilde{h}_2(\mathbf{p})$, and integrate over $\mathbf{p}$ to get (12.82).

Now we have

$$\begin{aligned} &C_\beta(s_1,h_1;\ s_2,h_2) \\ &= \sum_{n\in\mathbb{Z}} (\tilde{h}_1, \left(\frac{e^{-|s_2-s_1+\beta n|\omega}}{2\omega}\right)\tilde{h}_2) \\ &= \sum_{n=-\infty}^{-1} \left(\tilde{h}_1, \left(\frac{e^{(s_2-s_1)\omega}}{2\omega}\right) e^{n\beta\omega}\tilde{h}_2\right) + \sum_{n=0}^{\infty} \left(\tilde{h}_1, \left(\frac{e^{-(s_2-s_1)\omega}}{2\omega}\right) e^{-n\beta\omega}\tilde{h}_2\right) \\ &= \left(\tilde{h}_1, \left(\frac{e^{(-\beta+s_2-s_1)\omega}}{2\omega(1-e^{-\beta\omega})}\right)\tilde{h}_2\right) + \left(\tilde{h}_1, \left(\frac{e^{-(s_2-s_1)\omega}}{2\omega(1-e^{-\beta\omega})}\right)\tilde{h}_2\right) \end{aligned} \tag{12.85}$$

Taking into account that h real implies $\overline{\tilde{h}_i(\mathbf{p})} = \tilde{h}_i(-\mathbf{p})$ this is the same as (12.81).

For general n the correlation function $< \phi_0(t_1, h_1) \cdots \phi_0(t_n, h_n) >_\beta$ can be expressed as a sum over pairings $\sum_P \prod_{\{i,j\}\in P} < \phi_0(t_i, h_i)\phi_0(t_j, h_j) >_\beta$ by a variation of problem 6.4. Thus it continues to $\sum_P \prod_{\{i,j\}\in P} C_\beta(s_i, h_i; s_j, h_j)$, which is $E\Big(\phi(s_1, h_1) \cdots \phi(s_n, h_n)\Big)$ by (12.10). □

Problem 12.9

1. Check that both sides of (12.83) satisfy (12.84).
2. (uniqueness) Show that if $u \in \mathcal{S}'(\mathbb{R})$ and $(-d^2/dt^2 + \omega^2)u = 0$, then $u = 0$.

Notes on chapter 12: A general reference is Glimm and Jaffe (1987). For Gaussian measures on $\mathcal{S}'(\mathbb{R}^d)$, see Gelfand and Vilenkin (1964). There is an analogue of path integrals for fermions, see for example Salmhofer (1999).

13 A nonlinear field theory

13.1 The model

In this chapter we give an example of a field theory governed by a nonlinear field equation. In its particle aspect the nonlinearity means that particles can interact with each other and can also be created and destroyed.

We take the simplest nontrivial case which is a scalar field on the spacetime $(\mathbb{R}^2, \eta)$ obeying the field equation (7.81)

$$(-\Box + m^2)\phi + 4\lambda\phi^3 = 0 \tag{13.1}$$

Here λ is a positive coupling constant. As a first-order system it has the form

$$\begin{aligned} \frac{d\phi}{dt} &= \pi \\ \frac{d\pi}{dt} &= -(-\Delta + m^2)\phi - 4\lambda\phi^3 \end{aligned} \tag{13.2}$$

This is an infinite-dimensional Hamiltonian system with the Hamiltonian

$$\begin{aligned} H &= H_0 + V \\ H_0(\phi, \pi) &= \frac{1}{2}\int_{-\infty}^{\infty} \Big(\pi(x)^2 + (\nabla\phi(x))^2 + m^2\phi(x)^2\Big)dx \\ V(\phi) &= \lambda\int_{-\infty}^{\infty} \phi(x)^4 dx \end{aligned} \tag{13.3}$$

This is formally positive, which is why we took ϕ^3 in the field equation rather than say ϕ^2. The model is known as the ϕ_2^4 model, the two for dimension $d = 2$. More generally if we replace ϕ^4 by a lower semi-bounded polynomial $P(\phi)$, it is called the $P(\phi)_2$ model.

The problem is to construct operator valued solutions to this equation with initial values $\pi_0(x), \phi_0(x)$, which satisfy the canonical commutation relations $[\phi_0(x), \pi_0(y)] = i\delta(x - y)$. From chapter 8 we already know the solution for $\lambda = 0$. Start with the Fock space $\mathcal{F}^+(\mathcal{H}_0)$ with $\mathcal{H}_0 = L^2(\mathbb{R}^3, d\mathbf{p})$. Define time zero fields ϕ_0, π_0 as in (8.56). If the terms in $H_0 = H_0(\phi_0, \pi_0)$ are Wick-ordered, then the operator H_0 is well-defined as $H_0 = d\Gamma(\omega)$ (problem 8.8) and the

dynamical equations (13.2) are solved by $\phi(t,x) = e^{iH_0 t}\phi_0(x)e^{-iH_0 t}$ and $\pi(t,x) = e^{iH_0 t}\pi_0(x)e^{-iH_0 t}$ interpreted as distributions.

If we try to incorporate V into this picture, we would try to add a term $\lambda\int_{-\infty}^{\infty}\phi_0(x)^4 dx$ to the Hamiltonian. But this is very poorly defined. In the first place we are raising a (operator-valued) distribution to the fourth power, something with no natural meaning. In the second place we are integrating something with no decay over all of $\mathbb{R}$. Our task is to explain how to deal with these problems.

First some comments about the physics. On states with low momentum $\omega(p) = \sqrt{p^2+m^2} \approx m$. Hence in (8.56) for $d = 1$ the field is approximately

$$\phi_0(x) \approx (2m)^{-1/2}(a(x) + a^*(x)) \tag{13.4}$$

where $a(x) = (2\pi)^{-1/2}\int e^{ipx}a(p)$ labels a particle at position x. Hence with Wick ordering

$$\lambda\int :\phi_0(x)^4: dx \approx \frac{3\lambda}{2m^2}\int a^*(x)\,a^*(x)\,a(x)\,a(x)\,dx + \dots \tag{13.5}$$

This leading term preserves particle number and comparing it with (5.79) we see that it describes pairs of particles interacting with a repulsive delta function potential $v(x-y) = 3\lambda/m^2\,\delta(x-y)$. There are also other terms such as $\int a^*(x)^4 dx$ which create or annihilate particles, but these turn out to be less important at low momentum.

13.2 Regularization

We regularize the problem as follows. As mentioned, the first step is to replace $\lambda\int_{-\infty}^{\infty}\phi_0(x)^4 dx$ with a Wick-ordered version

$$V = \lambda\int_{-\infty}^{\infty} :\phi_0(x)^4: dx \tag{13.6}$$

The second step is to restrict the integral to a finite interval by

$$V_L = \lambda\int_{-L}^{L} :\phi_0(x)^4: dx \tag{13.7}$$

The third step is to regularize the field. Let χ be an arbitrary positive function in $\mathcal{C}_0^\infty(\mathbb{R})$ with $\int\chi(x)dx = 1$. Then let

$$\delta_\kappa(x) = \kappa\,\chi(\kappa x) \tag{13.8}$$

This is an approximate delta-function in the sense that any continuous function $\int f(y)\delta_\kappa(x-y)dx \to f(x)$ as $\kappa\to\infty$ (problem 13.1). We define a regularized field by

$$\phi_{0,\kappa}(x) = \phi_0(\delta_k(\cdot - x)) \tag{13.9}$$

and a regularized potential by

$$V_{L,\kappa} = \lambda \int_{-L}^{L} : \phi_{0,\kappa}(x)^4 : dx \tag{13.10}$$

This is a well-defined operator on a dense domain on Fock space.

We now change to the Schrödinger representation as in section 12.2. The field operators $\phi_0(h)$ become a family of Gaussian random variables $\phi(h)$ with covariance $(2\hat{\omega})^{-1}$ on a measure space $(\mathcal{M}, \Sigma, \mu)$. The equivalence is provided by the unitary map V from the Fock space to $L^2(\mathcal{M}, \mu)$ defined in theorem 12.3. The free Hamiltonian is now VH_0V^{-1} also denoted H_0, and (problem 12.6) the interaction becomes $VV_{L,\kappa}V^{-1}$, which is multiplication by

$$V_{L,\kappa}(\phi) = \lambda \int_{-L}^{L} : \phi_\kappa(x)^4 :_{(2\hat{\omega})^{-1}} dx \tag{13.11}$$

where $\phi_\kappa(x)$ is the random variable $\phi_\kappa(x) = \phi(\delta_\kappa(\cdot - x))$. Then $V_{L,\kappa}(\phi)$ is well-defined as a function in $L^p(\mathcal{M}, \mu)$ for all $1 \leq p < \infty$.

Our goal is to remove the regularizations by finding a meaning for the limits $\kappa \to \infty$ and $L \to \infty$. The limit $\kappa \to \infty$ is facilitated by the Wick monomials as we now explain. Note that the random variables $\phi_\kappa(x)$ have a covariance

$$\begin{aligned} c_\kappa(x-y) &\equiv E(\phi_\kappa(x)\phi_\kappa(y)) \\ &= \left\langle \delta_\kappa(\cdot - x), (2\hat{\omega})^{-1}\delta_\kappa(\cdot - y) \right\rangle \\ &= \int e^{ip(x-y)} |\tilde{\chi}(p/\kappa)|^2 (2\omega(p))^{-1} dp \end{aligned} \tag{13.12}$$

Thus we have

$$: \phi_\kappa(x)^4 :_{(2\hat{\omega})^{-1}} = \phi_\kappa(x)^4 - 6c_\kappa(0)\phi_\kappa(x)^2 + 3\big(c_\kappa(0)\big)^2 \tag{13.13}$$

As $\kappa \to \infty$ we have $\tilde{\chi}(p/\kappa) \to \tilde{\chi}(0) = (2\pi)^{-1/2}$ and $c_\kappa(0) \to \infty$, in fact $c_\kappa(0)$ grows like $\log \kappa$. Thus in $\int_{-L}^{L} : \phi_\kappa(x)^4 :_{(2\hat{\omega})^{-1}} dx$ the constant and quadratic terms develop infinite coefficients as $\kappa \to \infty$. The idea is that these should cancel the natural infinities in the quartic term $\int_{-L}^{L} \phi_\kappa(x)^4 dx$. This is an example of renormalization.

Theorem 13.1 *The following limit exists in* $L^2(\mathcal{M}, \mu)$

$$V_L = \lim_{\kappa \to \infty} V_{L,\kappa} \tag{13.14}$$

Proof We first compute the L^2 norm of $V_{L,\kappa}$. We have by (12.18)

$$\begin{aligned} \|V_{L,\kappa}\|_2^2 &= \lambda^2 \int_{-L}^{L} dx \int_{-L}^{L} dy\, E\Big(: \phi_\kappa(x)^4 :_{(2\hat{\omega})^{-1}} : \phi_\kappa(y)^4 :_{(2\hat{\omega})^{-1}} \Big) \\ &= 4!\,\lambda^2 \int_{-L}^{L} dx \int_{-L}^{L} dy\, (c_\kappa(x-y))^4 \\ &= 4!\,\lambda^2 \left\langle \chi_{[-L,L]},\, c_\kappa^4 * \chi_{[-L,L]} \right\rangle \end{aligned} \tag{13.15}$$

Using the Schwarz inequality and the bound $\|f * g\|_2 \leq \|f\|_1 \|g\|_2$ we have

$$\|V_{L,\kappa}\|_2^2 \leq 4!\,\lambda^2 \|\chi_{[-L,L]}\|_2^2 \|c_\kappa^4\|_1 = 48\lambda^2 L \|c_\kappa\|_4^4 \tag{13.16}$$

By the Hausdorff–Young inequality[1] $\|c_\kappa\|_4 \leq \|\tilde{c}_\kappa\|_{4/3}$. Furthermore

$$\tilde{c}_\kappa(p) = (2\pi)^{1/2} |\tilde{\chi}(p/\kappa)|^2 (2\omega(p))^{-1} \leq (2\omega(p))^{-1} \tag{13.17}$$

and so

$$\|\tilde{c}_\kappa\|_{4/3} \leq \left(\int (2\omega(p))^{-4/3} dp \right)^{3/4} < \infty \tag{13.18}$$

Hence $\|c_\kappa\|_4$ is bounded in κ and hence so is $\|V_{L,\kappa}\|_2^2$. This result fails in higher dimensions (the integral (13.18) is infinite), which is why we have taken one space dimension.

Similarly we have with $c_{\kappa,\kappa'}(x-y) = < \delta_\kappa(\cdot - x), (2\hat{\omega})^{-1} \delta_{\kappa'}(\cdot - y) >$

$$\begin{aligned} &\|V_{L,\kappa} - V_{L,\kappa'}\|_2^2 \\ &= 4!\,\lambda^2 \int_{-L}^{L} dx \int_{-L}^{L} dy \Big((c_\kappa(x-y))^4 - 2(c_{\kappa,\kappa'}(x-y))^4 + (c_{\kappa'}(x-y))^4 \Big) \end{aligned} \tag{13.19}$$

Since $\tilde{\chi}(p)$ has a bounded derivative, $|\tilde{\chi}(p) - \tilde{\chi}(p')|$ is bounded by a constant times $|p - p'|$. Since also $|\tilde{\chi}(p)|$ is bounded, we have that $|\tilde{\chi}(p) - \tilde{\chi}(p')|$ is bounded by a constant times $|p - p'|^\epsilon$ for any $0 \leq \epsilon \leq 1$. Hence with $\kappa \wedge \kappa' = \min\{\kappa, \kappa'\}$ we have $|\tilde{\chi}(p/\kappa) - \tilde{\chi}(p/\kappa')| \leq \mathcal{O}((\kappa \wedge \kappa')^{-\epsilon}) |p|^\epsilon$ and therefore

$$|\tilde{c}_\kappa(p) - \tilde{c}_{\kappa,\kappa'}(p)| \leq \mathcal{O}((\kappa \wedge \kappa')^{-\epsilon}) |p|^\epsilon \omega(p)^{-1} \tag{13.20}$$

For ϵ small the extra $|p|^\epsilon$ does not spoil the convergence of our integrals and so

$$\|c_\kappa - c_{\kappa,\kappa'}\|_4 \leq \|\tilde{c}_\kappa - \tilde{c}_{\kappa,\kappa'}\|_{4/3} \leq \mathcal{O}((\kappa \wedge \kappa')^{-\epsilon}) \tag{13.21}$$

Using estimates like this in (13.19) gives that

$$\|V_{L,\kappa} - V_{L,\kappa'}\|_2^2 = \mathcal{O}((\kappa \wedge \kappa')^{-\epsilon}) \tag{13.22}$$

Hence $\|V_{L,\kappa} - V_{L,\kappa'}\|_2^2 \to 0$ as $\kappa, \kappa' \to \infty$ and by the completeness of L^2 there is a limit V_L in L^2. □

Convergence in L^p can also be established. Indeed one can show that for $\epsilon > 0$ and small and any even integer p there is a constant c (depending on ϵ, λ, L) so that

$$\|V_{L,\kappa} - V_{L,\kappa'}\|_p^p \leq c^p (2p)! \, (\kappa \wedge \kappa')^{-\epsilon p} \tag{13.23}$$

The dependence on p can be understood as follows. One can evaluate the Gaussian integrals in this norm as a sum over graphs on p vertices with four legs at each vertex with the restriction that no lines join legs of the same vertex, see (12.26). The number

[1] The Hausdorff–Young inequality says that if a function f is in $L^p(\mathbb{R}^d)$ for $1 \leq p \leq 2$, then the Fourier transform $\tilde{f}$ is in $L^q(\mathbb{R}^d)$ for $p^{-1} + q^{-1} = 1$ and $\|\tilde{f}\|_q \leq (2\pi)^{d/2 - d/p} \|f\| p$. We use it for the inverse transform.

of such graphs is dominated by the number of graphs without the restriction which is $(4p-1)(4p-3)\cdots 3\cdot 1 \le 2^{2p}(2p)!$.

Completing the square in (13.13) gives $:\phi_\kappa(x)^4: \ge -6c_\kappa(0)^2$ and hence

$$V_{L,\kappa} \ge -12\lambda L c_\kappa(0)^2 \ge -b(\log\kappa)^2 + 1 \tag{13.24}$$

for some constant b. This is a sharp bound and so $\exp(-V_{L,\kappa})$ becomes unbounded as $\kappa \to \infty$. Nevertheless e^{-V_L} is integrable for we have:

Theorem 13.2 (Nelson)

$$E\left(e^{-V_L}\right) < \infty \tag{13.25}$$

Proof The idea is to show that although V_L is not bounded below it only becomes very negative on a set with small measure. Using (13.24) have for any κ and any even p

$$\begin{aligned} P\left(e^{-V_L} \ge e^{b(\log\kappa)^2}\right) &= P\left(V_L \le -b(\log\kappa)^2\right) \\ &\le P\Big(|V_L - V_{L,\kappa}| \ge 1\Big) \\ &\le \|V_L - V_{L,\kappa}\|_p^p \\ &\le c^p(2p)!\,\kappa^{-p\epsilon} \end{aligned} \tag{13.26}$$

In the last step we have used (13.23) at $\kappa' = \infty$. Choosing p close to $\kappa^{\epsilon/4}$ and using Stirling's formula for the asymptotics of $(2p)!$ yields for κ sufficiently large

$$P\left(e^{-V_L} \ge e^{b(\log\kappa)^2}\right) \le \exp(-\kappa^{\epsilon/4}) \tag{13.27}$$

or with $t = e^{b(\log\kappa)^2}$

$$P\left(e^{-V_L} \ge t\right) \le \exp\Big(-\exp\Big(\frac{\epsilon}{4}\sqrt{\frac{\log t}{b}}\Big)\Big) \tag{13.28}$$

Since

$$E\left(e^{-V_L}\right) = \int_0^\infty P\left(e^{-V_L} \ge t\right) dt \tag{13.29}$$

this is sufficient to establish the integrability. □

The Hamiltonian for the model is now $H_L = H_0 + V_L$ on a dense domain in $L^2(\mathcal{M},\mu)$. The potential V_L is not a Kato perturbation of H_0. Nevertheless by a more difficult proof[2] which uses the result (13.25):

Theorem 13.3 *H_L is essentially self-adjoint on $D(H_0)\cap D(V_L)$ and H_L is bounded below.*

This theorem and the next can actually be circumvented as we explain later.

[2] See Glimm and Jaffe (1970), or Reed and Simon (1975: 267).

Theorem 13.4 *Let $E_L = \inf \sigma(H_L)$ be the bottom of the spectrum for H_L. Then $E_L < 0$ is a simple eigenvalue with eigenvector Ω_L which is a strictly positive function.*

This is a generalization of theorem 11.5 to an infinite-dimensional measure space.[3]

The state Ω_L is the *vacuum* for the model and can be thought of as the distortion of the free vacuum Ω_0. Well outside of the interval $[-L, L]$ we have $H_L \approx H_0$ and so we expect Ω_L to be close to Ω_0, which has no particles. However inside the interval $[-L, L]$ the vacuum Ω_L is filled with particles. The vacuum is not empty. (These remarks are best visualized in the Fock representation.)

Problem 13.1 Let $\chi \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ with $\int \chi = 1$ and let $\delta_\kappa(x) = \kappa^d \chi(\kappa x)$. Show that δ_κ is a family of approximate delta functions in the sense that for any function f continuous on a neighborhood of the origin in $\mathbb{R}^d$

$$\lim_{\kappa \to \infty} \langle \delta_\kappa, f \rangle = < \delta, f > \equiv f(0) \tag{13.30}$$

Problem 13.2 Fill in the details in the proof of theorem 13.2.

13.3 Infinite volume

13.3.1 Wightman functions

Now we would like to take the limit $L \to \infty$. The operators H_L and vacuum vectors Ω_L actually have no limit in Fock space. But we use them to define some correlation functions which do have limits. Then from these infinite volume correlation functions we will give an abstract construction of a new Hilbert space, new field operators, and a new vacuum vector, which reproduce the correlation functions.

The correlation functions are vacuum expectation values of products of field operators. The field operator is $\phi_L(t, x) = \exp(iH_L t)\phi(x)\exp(-iH_L t)$ or smeared with $f \in \mathcal{S}(\mathbb{R}^2)$

$$\phi_L(f) = \int \phi_L(t, x) f(t, x) dx dt = \int e^{iH_L t} \phi(f(t, \cdot)) e^{-iH_L t} dt \tag{13.31}$$

One can show that $\phi_L(f)$ is a well-defined operator mapping the dense domain $\mathcal{C}^\infty(H_L) = \bigcap_{n=1}^\infty D(H_L^n)$ to itself. If f is real, it is symmetric, but we can extend the definition to complex f by linearity and then $\phi_L(\bar{f}) \subset \phi_L(f)^*$. Since Ω_L is a vector in $\mathcal{C}^\infty(H_L)$, we can define vacuum correlation functions or *Wightman functions* by

[3] See Glimm and Jaffe (1970) or Reed and Simon (1978: p.208).

$$W_{n,L}(f_1,\dots,f_n) = (\Omega_L, \phi_L(f_1)\dots\phi_L(f_n)\Omega_L) \tag{13.32}$$

These Wightman functions are bounded multilinear functionals on $\mathcal{S}(\mathbb{R}^2)$ uniformly in L. All these results follow from the "ϕ-bound"[4], which says that there is a constant C so for real $h \in \mathcal{S}(\mathbb{R}^1)$ and all L

$$\pm\,\phi(h) \le C\|h\|_1(H_L + 1) \tag{13.33}$$

We note some properties of these finite volume Wightman functions. By the kernel theorem (see appendix C) there exist distributions $W_{n,L} \in \mathcal{S}'(\mathbb{R}^{2n})$ such that

$$W_{n,L}(f_1 \otimes \cdots \otimes f_n) = W_{n,L}(f_1,\dots,f_n) \tag{13.34}$$

Lemma 13.1 *The finite volume Wightman functions $W_{n,L}(F)$ for F in complex $\mathcal{S}(\mathbb{R}^{2n})$ satisfy the following properties:*

1. *Let $F^*(x_1,\dots,x_n) = \overline{F(x_n,\dots,x_1)}$, then*

$$\overline{W_{n,L}(F)} = W_{n,L}(F^*) \tag{13.35}$$

2. *Let $F_0, F_1, F_2, \dots$ be a finite sequence with $F_0 \in \mathbb{C}$ and $F_n \in \mathcal{S}(\mathbb{R}^{2n})$. Then with $W_{0,L} = 1$*

$$\sum_{i,j} W_{i+j,L}(F_i^* \otimes F_j) \ge 0 \tag{13.36}$$

3. *Let $F_t(x_1^0, x_1^1, \dots, x_n^0, x_n^1) = F(x_1^0 - t, x_1^1, \dots, x_n^0 - t, x_n^1)$ be the time translate of F. Then for any $t \in \mathbb{R}$*

$$W_{n,L}(F_t) = W_{n,L}(F) \tag{13.37}$$

Proof For an identity like (13.35) it suffices to prove the result for $F = f_1 \otimes \cdots \otimes f_n$ in which case it follows from $\phi_L(\bar{f}) \subset \phi_L(f)^*$.

For the second point define

$$\Phi_{n,L}(f_1,\dots,f_n) = \phi_L(f_1)\cdots\phi_L(f_n)\Omega_L \tag{13.38}$$

This is a continuous (vector-valued) multilinear functional on $\mathcal{S}(\mathbb{R}^2)$ and so by a (vector-valued) kernel theorem there is a unique extension to a (vector-valued) linear function $\Phi_{n,L}(F)$ on $\mathcal{S}(\mathbb{R}^{2n})$ such that

$$\Phi_{n,L}(f_1 \otimes \cdots \otimes f_n) = \Phi_{n,L}(f_1,\dots,f_n) \tag{13.39}$$

Then for $F \in \mathcal{S}(\mathbb{R}^{2i})$ and $G \in \mathcal{S}(\mathbb{R}^{2j})$ we have

$$W_{i+j,L}(F^* \otimes G) = (\Phi_{i,L}(F), \Phi_{j,L}(G)) \tag{13.40}$$

again by $\phi_L(\bar{f}) \subset \phi_L(f)^*$. Then for sequences $F_0, F_1, F_2, \dots$

$$\sum_{i,j} W_{i+j,L}(F_i^* \otimes F_j) = \left(\sum_i \Phi_{i,L}(F_i), \sum_j \Phi_{j,L}(F_j)\right) \ge 0 \tag{13.41}$$

The third point follows by $\phi_L(f_t) = e^{iH_Lt}\phi_L(f)e^{-iH_Lt}$ and $e^{-iH_Lt}\Omega_L = \Omega_L$. □

[4] Glimm and Jaffe (1972).

We would like to show that $W_n = \lim_{L\to\infty} W_{n,L}$ exists in $\mathcal{S}'(\mathbb{R}^{2n})$. This is possible, but difficult. We will have more to say about it. For the moment suppose that the limit does exist. The properties (13.35), (13.36), (13.37) established in the previous lemma then carry over to the limit. These are sufficient to reconstruct a field theory as we now explain.

13.3.2 Reconstruction

The reconstruction theorem is quite general and we state the result for a scalar field in a d-dimensional spacetime. There are also versions for other spins.

Theorem 13.5 (Wightman reconstruction theorem) *Let $W_n \in \mathcal{S}'(\mathbb{R}^{nd})$ be a family of distributions with $W_0 = 1$. Suppose they satisfy*

$$\overline{W_n(F)} = W_n(F^*) \tag{13.42}$$

and for any finite sequence $F_0, F_1, F_2, \ldots$ with $F_n \in \mathcal{S}(\mathbb{R}^{nd})$

$$\sum_{i,j} W_{i+j}(F_i^* \otimes F_j) \geq 0 \tag{13.43}$$

Then there exists a Hilbert space $\mathcal{H}$, a dense domain $\mathcal{D} \subset \mathcal{H}$, a family of field operators $\phi(f) : \mathcal{D} \to \mathcal{D}$ for $f \in \mathcal{S}(\mathbb{R}^d)$, and a vector $\Omega \in \mathcal{D}$ such that

$$W_n(f_1 \otimes \cdots \otimes f_n) = (\Omega, \phi(f_1) \ldots \phi(f_n)\Omega) \tag{13.44}$$

Furthermore $\phi(\bar{f}) \subset \phi(f)^*$. In addition if the W_n are time translation invariant in the sense that

$$W_n(F_t) = W_n(F) \tag{13.45}$$

then there is a self-adjoint operator H such that e^{-iHt} preserves $\mathcal{D}$ and satisfies

$$e^{-iHt}\Omega = \Omega \qquad\qquad e^{iHt}\phi(f)e^{-iHt} = \phi(f_t) \tag{13.46}$$

Proof Let $\mathcal{E}$ be the space of sequences $F = (F_0, F_1, F_2, \ldots)$ with $F_0 \in \mathbb{C}$ and $F_n \in \mathcal{S}(\mathbb{R}^{nd})$. On $\mathcal{E}$ define

$$(F, G) = \sum_{i,j} W_{i+j}(F_i^* \otimes G_j) \tag{13.47}$$

By (13.42)

$$\overline{(F, G)} = \sum_{i,j} \overline{W_{i+j}(F_i^* \otimes G_j)} = \sum_{i,j} W_{i+j}(G_j^* \otimes F_i) = (G, F) \tag{13.48}$$

and by (13.43) $(F, F) \geq 0$. Thus (F, G) is an inner product except that it is not positive definite. Let $\mathcal{N}$ be the subspace of all F with $\|F\|^2 = (F, F) = 0$ and form the factor space $\mathcal{E}/\mathcal{N}$ consisting of equivalence classes $[F]$. Using the Schwarz inequality one shows that the inner product is well-defined on $\mathcal{E}/\mathcal{N}$ by $([E], [F]) = (E, F)$. On this

space it is positive definite. Thus $\mathcal{D} \equiv \mathcal{E}/\mathcal{N}$ is a pre-Hilbert space and we let $\mathcal{H}$ be the completion. Ω is the equivalence class of $(1, 0, 0, \dots)$.

The field operator is defined on $\mathcal{E}$ by

$$\phi(f)F = (0, f \otimes F_0, f \otimes F_1, f \otimes F_2, \dots) \tag{13.49}$$

Then

$$\begin{aligned}(F, \phi(f)G) &= \sum_{i,j} W_{i+j+1}(F_i^* \otimes f \otimes G_j) \\ &= \sum_{i,j} W_{i+j+1}((\bar{f} \otimes F_i)^* \otimes G_j) = (\phi(\bar{f})F, G)\end{aligned} \tag{13.50}$$

By the Schwarz inequality

$$\|\phi(f)F\|^2 = (\phi(f)F, \phi(f)F) = (F, \phi(\bar{f})\phi(f)F) \leq \|F\|\|\phi(\bar{f})\phi(f)F\| \tag{13.51}$$

Hence if $F \in \mathcal{N}$, then $\phi(f)F \in \mathcal{N}$ and so $\phi(f)$ is defined on $\mathcal{E}/\mathcal{N}$ and $\phi(f)[F] = [f \otimes F]$. The identity (13.44) holds since $\phi(f_1) \dots \phi(f_n)\Omega$ is the equivalence class of $(0, 0, \dots, f_1 \otimes \cdots \otimes f_n, 0, \dots)$ and so

$$\begin{aligned}&W_n(f_1 \otimes \cdots \otimes f_n) \\ &= \Big((1, 0, 0, \cdots), (0, 0, \dots, f_1 \otimes \cdots \otimes f_n, 0, \dots) \Big) \\ &= (\Omega, \phi(f_1) \dots \phi(f_n)\Omega)\end{aligned} \tag{13.52}$$

For the second point define $F_t = (F_0, F_{1,t}, F_{2,t}, \dots)$. Then $t \to F_t$ is a representation of the group $\mathbb{R}$ on $\mathcal{E}$ and it preserves the inner product by (13.45). Hence it preserves $\mathcal{N}$ and defines a representation of $\mathbb{R}$ on $\mathcal{E}/\mathcal{N}$ by $U(t)[F] = [F_t]$, which also preserves the inner product. This extends to a one-parameter unitary group $U(t)$ on $\mathcal{H}$, which is continuous since translations are continuous on the Schwartz space. The Hamiltonian H is defined to be the generator $U(t) = e^{iHt}$ and the identities (13.46) are easily checked. □

We have given a bare bones version of the reconstruction theorem. The full theorem[5] has more features some of which we develop in the following problems.

Problem 13.3 (Poincaré invariance) The proper Poincaré group on $(\mathbb{R}^d, \eta)$ acts on $\mathcal{S}(\mathbb{R}^d)$ by $f_{a,\Lambda}(x) = f(\Lambda^{-1}(x - a))$. Show that if the Wightman functions in theorem 13.5 are invariant in the sense that

$$W_n((f_1)_{a,\Lambda} \otimes \cdots \otimes (f_n)_{a,\Lambda}) = W_n(f_1 \otimes \cdots \otimes f_n) \tag{13.53}$$

then there is a unitary representation $U(a, \Lambda)$ of the Poincaré group on $\mathcal{H}$ such that

$$U(a, \Lambda)\Omega = \Omega \qquad U(a, \Lambda)\phi(f)U(a, \Lambda)^{-1} = \phi(f_{a,\Lambda}) \tag{13.54}$$

[5] See Streater and Wightman (1964).

Problem 13.4 (locality) Show that if the Wightman functions in theorem 13.5 satisfy

$$W_n(f_1 \otimes \cdots \otimes f_j \otimes f_{j+1} \otimes \cdots \otimes f_n) = W_n(f_1 \otimes \cdots \otimes f_{j+1} \otimes f_j \otimes \cdots \otimes f_n) \quad (13.55)$$

whenever f_j and f_{j+1} have spacelike separated supports, then

$$[\phi(f), \phi(g)] = 0 \quad (13.56)$$

whenever f and g have spacelike separated supports.

13.3.3 Interpretation

Following these general considerations we return to the ϕ_2^4 model. Assuming that the infinite volume Wightman functions exist, the triple $(\mathcal{H}, \Omega, \phi(f))$ of the reconstruction theorem forms the basic model. The physical picture is that the Hilbert space $\mathcal{H}$ has a distinguished vacuum vector Ω which is invariant under time evolution. Since it is in some sense the limit as $L \to \infty$ of the Fock vacua Ω_L localized in $[-L, L]$, we can think of it as a sea of Fock particles (also called bare particles) filling all space. States $\phi(f_1) \dots \phi(f_n)\Omega$ represent local disturbances in the vacuum. These states span a dense set so the entire Hilbert space can be thought of as local distortions of the vacuum.

Furthermore suppose that we could establish the Poincaré invariance. This would not be straightforward since the finite volume Wightman functions are not invariant. Nevertheless once it is known, we would have a representation of the Poincaré group. Physical particles could be identified by finding irreducible subspaces for this representation. A physical particle can be thought of as a cloud of bare particles.

13.4 Path integrals – interacting fields

Continuing with the ϕ_2^4 model we seek to represent the imaginary time dynamics as an integral over paths just as for the free field. That is we seek another version of the Feynman–Kac formula as in theorem 12.4.

Again let $\phi(h)$ be a Gaussian process indexed by $h \in \mathcal{S}(\mathbb{R})$ with covariance $< h_1, (2\hat{\omega})^{-1} h_2 >$ on a measure space $(\mathcal{M}, \Sigma, \mu)$ and let $\phi(t, h)$ be a Gaussian process indexed by $(t, h) \in \mathbb{R} \times \mathcal{S}(\mathbb{R})$ with covariance $C(t_1, h_1; t_2, h_2)$ (defined in (12.55), (12.56)) on a measure space $(\mathcal{M}', \Sigma', \mu')$. We now define a potential on fields $\phi(t, h)$ by

$$V_{[0,T]\times[-L,L]} = \lambda \int_0^T \int_{-L}^{L} : \phi(t, x)^4 :_C \, dtdx \quad (13.57)$$

This can be defined as a limit in $L^2(\mathcal{M}', \mu')$ of regularized potentials with $:\phi(t,x)^4:_C$ replaced by $:\phi(t,\delta_\kappa(\cdot - x))^4:_C$. This is entirely similar to the treatment in theorem 13.1 of the potential V_L in $L^2(\mathcal{M},\mu)$. In fact we have the identity

$$V_{[0,T]\times[-L,L]} = \int_0^T J_t V_L dt \tag{13.58}$$

where $J_t : L^2(\mathcal{M},\mu) \to L^2(\mathcal{M}',\mu')$ is defined in theorem 12.4. This holds with the regularizations on both sides since

$$:\phi(t,\delta_\kappa(\cdot - x))^4:_C = J_t :\phi(\delta_\kappa(\cdot - x))^4:_{(2\hat{\omega})^{-1}} \tag{13.59}$$

and hence in the limit $\kappa \to \infty$.

Theorem 13.6

1. $\exp\left(-V_{[0,T]\times[-L,L]}\right)$ *is in* $L^p(\mathcal{M}',\mu')$ *for all* $p < \infty$.
2. *Let* $\psi, \chi \in L^2(\mathcal{M},\mu)$ *be polynomials and let* $H_L = H_0 + V_L$. *Then*

$$(\psi, e^{-TH_L}\chi) = E\left(\overline{J_0\psi}\ \exp\left(-V_{[0,T]\times[-L,L]}\right) J_T\chi\right) \tag{13.60}$$

Remark The proof of the first part is entirely similar to the proof of theorem 13.2 and is omitted. The second part is a Feynman–Kac formula. The proof given below is analogous to the proof of theorem 11.3, but here it is just heuristic. It could be made rigorous but this would probably not be the most efficient way to obtain the result.[6]

Proof We compute with $t_j = jT/n$

$$\begin{aligned}
(\psi, e^{-TH_L}\chi) &= \lim_{n\to\infty}\left(\psi, \left(e^{-TH_0/n}e^{-TV_L/n}\right)^n \chi\right) \\
&= \lim_{n\to\infty} E\left(\overline{J_0\psi}\left(\prod_{j=1}^n J_{t_j} e^{-TV_L/n}\right) J_T\chi\right) \\
&= \lim_{n\to\infty} E\left(\overline{J_0\psi}\exp\left(-\frac{T}{n}\sum_{j=1}^n J_{t_j}V_L\right) J_T\chi\right) \\
&= E\left(\overline{J_0\psi}\exp\left(-\int_0^T J_t V_L\, dt\right) J_T\chi\right)
\end{aligned} \tag{13.61}$$

Here the first step is the Trotter product formula. The second step is the free field result (12.75). The third step uses $J_t \exp(-V_L) = \exp(-J_t V_L)$. In the last step we take the limit inside the integral and identify a Riemann sum. Finally use (13.58) to complete the proof. □

A variation of this result for the imaginary time correlation functions is the following. For

$$-T < t_1 < \cdots < t_n < T \tag{13.62}$$

[6] See for example Simon (1975: 163).

the bare vacuum correlation functions are (with $\Omega_0 = 1$ in our Schrödinger representation)

$$\begin{aligned}&\left(\Omega_0, e^{-(T+t_1)H_L}\phi(h_1)e^{-(t_2-t_1)H_L}\phi(h_2)\dots\phi(h_n)e^{-(T-t_n)H_L}\Omega_0\right)\\&= E\Big(\phi(t_1,h_1)\dots\phi(t_n,h_n)\exp\left(-V_{[-T,T]\times[-L,L]}\right)\Big)\end{aligned} \tag{13.63}$$

Formally this follows as in (13.61). In any case the expression on the left is well-defined since the ϕ bound (13.33) implies $(H_L+1)^{-1/2}\phi(h)(H_L+1)^{-1/2}$ is bounded and since $(H_L+1)^{1/2}e^{-tH_L}(H_L+1)^{1/2}$ is bounded for $t>0$. Next divide by

$$\|e^{-TH_L}\Omega_0\|^2 = E\left(\exp\left(-V_{[-T,T]\times[-L,L]}\right)\right) \tag{13.64}$$

By theorem 13.4 the lowest eigenvalue of H_L is simple with eigenvector Ω_L and so (cf. problem 11.7)

$$\lim_{T\to\infty}\frac{e^{-TH_L}\Omega_0}{\|e^{-TH_L}\Omega_0\|} = \Omega_L \tag{13.65}$$

Then we have for correlation functions with the physical vacuum

$$\begin{aligned}&\left(\Omega_L, \phi(h_1)e^{-(t_2-t_1)H_L}\phi(h_2)\dots\phi(h_{n-1})e^{-(t_n-t_{n-1})H_L}\phi(h_n)\Omega_L\right)\\&= \lim_{T\to\infty}\frac{E\left(\phi(t_1,h_1)\dots\phi(t_n,h_n)\exp\left(-V_{[-T,T]\times[-L,L]}\right)\right)}{E\left(\exp\left(-V_{[-T,T]\times[-L,L]}\right)\right)}\end{aligned} \tag{13.66}$$

An advantage of this representation is that the L dependence is in a place where we can get our hands on it. It is indeed possible to take the limit $L\to\infty$ in this form, see the notes for references. If the fields $\phi(t,x)$ were independent random variables, this would be easy since one could cancel the large distance contributions from $V_{[-T,T]\times[-L,L]}$ in the numerator and denominator. They are not independent but they are approximately independent as points separate since the covariance C is exponentially decaying. This is the basic mechanism behind the result.

Now we can give an indication of why the Wightman functions have an infinite volume limit as well.[7] By the ϕ-bound the functions

$$(\Omega_L, \phi(h_1)e^{-(\tau_2-\tau_1)H_L}\phi(h_2)\dots\phi(h_{n-1})e^{-(\tau_n-\tau_{n-1})H_L}\phi(h_n)\Omega_L) \tag{13.67}$$

are analytic and bounded uniformly in L on compact subsets of the complex region $Re(\tau_i-\tau_{i-1})>0$. Since the functions converge as $L\to\infty$ when the τ_i are real, it follows by the Vitali convergence theorem[8] that these functions have a limit for τ_i in the entire region. Then for $\epsilon>0$ and $f_i\in\mathcal{S}(\mathbb{R}^2)$

$$\int dt_1\dots dt_n\,(\Omega_L, \phi(f_1(t_1,\cdot))e^{-i(t_2-t_1-i\epsilon)H_L}\phi(f_2(t_2,\cdot))\cdots\phi(f_n(t_n,\cdot))\Omega_L) \tag{13.68}$$

[7] Glimm Jaffe and Spencer (1974).
[8] See for example Titchmarsh (1939).

also converges as $L \to \infty$. However as $\epsilon \to 0$ this expression converges to the Wightman function $W_{n,L}(f_1, \ldots, f_n)$ and by the ϕ-bound again one can show that the limit is uniform in L. Hence the Wightman functions converge as $L \to \infty$.

13.5 A reformulation

The approach we have sketched for the ϕ_2^4 model is the most intuitive, but it is not the most efficient. We now explain a variation which involves constructing the Hilbert space directly at imaginary time.

Start with the equivalent Gaussian random process $\phi(f)$ indexed by real $f \in \mathcal{S}(\mathbb{R}^2)$ with mean zero and covariance $C = (-\Delta + m^2)^{-1}$. Define the unnormalized finite volume Schwinger functions

$$S^0_{n,T,L}(f_1, \ldots, f_n) = E\Big(\phi(f_1) \ldots \phi(f_n) \exp\left(-V_{[-T,T]\times[-L,L]}\right)\Big) \tag{13.69}$$

and extend to complex test functions by linearity. By the kernel theorem this defines a linear functional $S^0_{n,T,L}$ on $\mathcal{S}(\mathbb{R}^{2n})$ such that

$$S^0_{n,T,L}(f_1 \otimes \cdots \otimes f_n) = S^0_{n,T,L}(f_1, \ldots, f_n) \tag{13.70}$$

It will be convenient to restrict attention to the algebraic tensor product $\otimes^n \mathcal{S}(\mathbb{R}^2) = \mathcal{S}(\mathbb{R}^2) \otimes \cdots \otimes \mathcal{S}(\mathbb{R}^2)$ which is the subspace of $\mathcal{S}(\mathbb{R}^{2n})$ consisting of finite combinations of the $f_1 \otimes \cdots \otimes f_n$.

If $f_1 \otimes \cdots \otimes f_n$ has support in the region (13.62), then we have the identity

$$\begin{aligned} &S^0_{n,T,L}(f_1 \otimes \cdots \otimes f_n) \\ &= \int \big(\Omega_0, e^{-(T+t_1)H_L} \phi(f_1(t_1, \cdot)) e^{-(t_2-t_1)H_L} \\ &\qquad \phi(f_2(t_2, \cdot)) \cdots \phi(f_n(t_n, \cdot)) e^{-(T-t_n)H_L} \Omega_0\big) dt_1 \cdots dt_n \end{aligned} \tag{13.71}$$

Indeed the field $\phi(f)$ can be written in terms of the sharp time field $\phi(t, h)$ as $\phi(f) = \int \phi(t, f(t, \cdot))dt$ and the result follows from (13.63).

Lemma 13.2 *The finite volume Schwinger functions have the following properties:*

1. *Let $F \in \otimes^n \mathcal{S}(\mathbb{R}^2)$ have support in $-T < t_1 < \cdots < t_n < T$ and define*

$$(\Theta F)(t_1, x_1, \ldots, t_n x_n) = \overline{F(-t_n, x_n, \ldots, -t_1, x_1)} \tag{13.72}$$

Then

$$\overline{S^0_{n,T,L}(F)} = S^0_{n,T,L}(\Theta F) \tag{13.73}$$

2. *(reflection positivity) Let $F_0, F_1, F_2, \ldots$ be a finite sequence with $F_0 \in \mathbb{C}$ and $F_n \in \otimes^n \mathcal{S}(\mathbb{R}^2)$ with support in $0 < t_1 < \cdots < t_n < T$. Then with $S^0_{0,T,L} = 1$*

$$\sum_{i,j} S^0_{i+j,T,L}(\Theta F_i \otimes F_j) \geq 0 \tag{13.74}$$

Proof It suffices to check the first identity for $F = f_1 \otimes \cdots \otimes f_n$. Then it follows from the representation (13.71) and the facts that e^{-Ht} is self-adjoint and $\phi(\bar{h}) \subset \phi(h)^*$.

For the second point let $f_1, \ldots, f_n$ have supports in $0 < t_1 < \cdots < t_n < T$ and define

$$\begin{aligned} &\Psi_{n,T,L}(f_1, \ldots, f_n) = \int dt_1 \cdots dt_n \\ &e^{-t_1 H_L} \phi(f(t_1, \cdot)) e^{-(t_2 - t_1)H_L} \phi(f(t_2, \cdot)) \ldots \phi(f(t_n, \cdot)) e^{-(T - t_n)H_L} \Omega_0 \end{aligned} \tag{13.75}$$

As a multilinear functional this defines a linear function $\Psi_{n,T,L}(F)$ on the subspace of $\otimes^n \mathcal{S}(\mathbb{R}^2)$ of restricted supports satisfying

$$\Psi_{n,T,L}(f_1 \otimes \cdots \otimes f_n) = \Psi_{n,T,L}(f_1, \ldots, f_n) \tag{13.76}$$

Now for $F \in \otimes^i \mathcal{S}(\mathbb{R}^2)$ and $G \in \otimes^j \mathcal{S}(\mathbb{R}^2)$ with restricted supports we have

$$S^0_{i+j,T,L}(\Theta F \otimes G) = \Big(\Psi_{i,T,L}(F), \Psi_{j,T,L}(G) \Big) \tag{13.77}$$

Again this follows from the representation (13.71) and adjoint relations. Then for sequences $F_0, F_1, F_2, \ldots$

$$\sum_{i,j} S^0_{i+j,T,L}(\Theta F_i \otimes F_j) = \Big(\sum_i \Psi_{i,T,L}(F_i), \sum_j \Psi_{j,T,L}(F_j) \Big) \geq 0 \tag{13.78}$$

□

Now consider the normalized Schwinger functions:

$$S_{n,T,L}(f_1, \ldots, f_n) = \frac{E\Big(\phi(f_1) \ldots \phi(f_n) \exp\big(-V_{[-T,T]\times[-L,L]}\big) \Big)}{E\Big(\exp\big(-V_{[-T,T]\times[-L,L]}\big) \Big)} \tag{13.79}$$

As noted earlier for these one can establish an infinite volume limit

$$S_n(f_1, \ldots, f_n) = \lim_{T,L \to \infty} S_{n,T,L}(f_1, \ldots, f_n) \tag{13.80}$$

These infinite volume Schwinger functions again satisfy $\overline{S_n(F)} = S_n(\Theta F)$ as well as the reflection positivity condition

$$\sum_{i,j} S_{i+j}(\Theta F_i \otimes F_j) \geq 0 \tag{13.81}$$

Now we can sketch a reconstruction theorem; the details are a bit too much to go into here. One uses (13.81) to define an inner product on sequences. Factoring out the null space and completing in the resulting norm gives a Hilbert space. On

the Hilbert space we can reconstruct field operators which reproduce the Schwinger functions S_n. The infinite volume Schwinger functions S_n will be time translation invariant (although the finite volume were not) and one can use this to construct a semi-group e^{-tH} which generates the time translations. All this is analogous to the Wightman reconstruction theorem. Then using the semi-group one can make an analytic continuation to real time and get a family of distributions W_n satisfying the Wightman axioms and hence a full field theory. This reconstruction theorem exists in various forms and was originally due to Osterwalder and Schrader.

This approach generalizes to other models and has a number of advantages:

1. It turns out to be relatively easy to prove the reflection positivity (13.81) directly in path space without establishing the connection with Fock space as in (13.71). Thus the whole Fock space construction can be dispensed with.
2. Properties of the Wightman functions can be deduced from simpler properties of the Schwinger functions. In particular Poincaré invariance of the Wightman functions can be deduced from the invariance of the Schwinger functions under the Euclidean group (translations, rotations). Also locality for the Wightman functions can be deduced from the symmetry of the Schwinger functions.
3. Expressions such as (13.79) for the finite volume Schwinger functions can be thought of as the correlation functions for a problem in classical statistical mechanics. The phase space is all field configurations on a Euclidean space. Hence in studying the Schwinger functions one can sometimes use techniques developed for classical statistical mechanics, for example in proving the existence of the infinite volume limit. (This works the other way also: field theory techniques have proved to be useful in statistical mechanics problems, both classical and quantum.)

The ϕ^4_2 model and more generally the $P(\phi)_2$ models have been completely constructed along the lines we have been discussing. Other nonlinear models have also been treated with varying degrees of success. In higher dimensions the renormalization problems become much more severe (for us Wick ordering was sufficient). The model ϕ^4_3 has been constructed, but ϕ^4_4 probably does not exist. There is as yet no model completely constructed in $d = 4$.

Once a model is constructed the next task is to find what particles are present. Then one looks for states whose long time behavior consists of a finite number of particles moving in separate trajectories. This is analogous to the construction of wave operators in section 4.5 and is the content of the Haag–Ruelle scattering theory. From these asymptotic states, one forms scattering amplitudes which can in principle be compared with the results of scattering experiments (in $d = 4$).

Notes on chapter 13: For the $P(\phi)_2$ model see Glimm and Jaffe (1970), Nelson (1973), Simon (1975), Glimm and Jaffe (1987).

For an axiomatic treatment of quantum field theory see Streater and Wightman (1964) or Bogolubov Logunov and Todorov (1975).

There is also a treatment of relativistic quantum physics in which fields are displaced as the primary objects and replaced by C^* algebras with a local structure. This has certain advantages and some believe is a more fundamental approach. For this algebraic version of quantum field theory see Haag (1992).

There are also books which attempt to explain quantum field theory as practiced by theoretical physicists to a mathematical audience, for example Folland (2008).

A Appendix A **Normed spaces**

A.1 Banach spaces

We review some basic facts. A *Banach space* X is a complete normed vector space. The vector space can be real or complex, but is complex unless specified otherwise. The *norm* is a real-valued function on X sending $x \in X$ to $\|x\|$ which satisfies:

1. $\|cx\| = |c|\|x\|$ for $c \in \mathbb{C}$
2. $\|x + y\| \leq \|x\| + \|y\|$
3. $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$.

The norm makes X into a metric space with distance function

$$d(x, y) = \|x - y\| \tag{A.1}$$

As such it is a topological space and we have all the usual notions of open sets, closed sets, dense sets, connected sets, compact sets, etc. To say X is *complete* means that every Cauchy sequence in X has a limit in X. That is if $\|x_n - x_m\| \to 0$ as $n, m \to \infty$, then there exists a (unique) $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$.

Examples:

1. For $1 \leq p < \infty$ let ℓ^p be the space of infinite sequences of complex numbers $x = (x_1, x_2, \dots)$ such that

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} \tag{A.2}$$

   is finite. Then ℓ^p is a Banach space with this norm.
2. For $1 \leq p < \infty$ let $L^p(\mathbb{R}^n)$ be the space of all complex measurable functions u on $\mathbb{R}^n$ such that the integral with respect to Lebesgue measure

$$\|u\|_p = \left(\int |u(x)|^p dx \right)^{1/p} \tag{A.3}$$

   is finite. If we identify functions which are equal almost everywhere, then $L^p(\mathbb{R}^n)$ is a Banach space with this norm.

3. Let $L^\infty(\mathbb{R}^n)$ be the space of all complex measurable functions u on $\mathbb{R}^n$ which are bounded almost everywhere. Identifying functions which are equal almost everywhere, this is a Banach space with the essential supremum norm

$$\|u\|_\infty = \text{ess sup}_x |u(x)| \tag{A.4}$$

4. Any closed subspace of a Banach space is a Banach space

Theorem A.1 *Any normed vector space X_0 can be identified as dense subspace of a Banach space $X = \overline{X_0}$ called the* completion *of X_0.*

We sketch the construction. The space X is equivalence classes of Cauchy sequences in X_0 with $\{x_i\} \sim \{x_i'\}$ if $\|x_i - x_i'\| \to 0$ as $i \to \infty$. Then X is naturally a vector space. One shows that $\lim_{i\to\infty} \|x_i\|$ exists and depends only on the equivalence class. This gives a norm on X and one shows that X is complete. X_0 is identified as the subspace of (equivalence classes of) constant sequences.

A.2 Hilbert spaces

A *Hilbert space* $\mathcal{H}$ is a complete inner product space. An inner product on a vector space $\mathcal{H}$ is a map from pairs $u, v \in \mathcal{H}$ to $(u, v) \in \mathbb{C}$ such that

1. (u, v) is linear in v and anti-linear in u.
2. $\overline{(u, v)} = (v, u)$
3. $(u, u) \geq 0$ and $(u, u) = 0$ iff $u = 0$

The inner product defines a norm by $\|u\| = \sqrt{(u, u)}$. Thus an inner product space is a normed space. Complete means complete as a normed space. Thus a Hilbert space is a Banach space.

The inner product can be recovered from the norm by the polarization identity

$$(u, v) = \frac{1}{4}\Big(\|u + v\|^2 - \|u - v\|^2 - i\|u + iv\|^2 + i\|u - iv\|^2\Big) \tag{A.5}$$

We also have the Schwarz inequality

$$|(u, v)| \leq \|u\|\|v\| \tag{A.6}$$

Examples:

1. ℓ^2 is a Hilbert space with $(x, y) = \sum_{i=1}^\infty \bar{x}_i y_i$.
2. $L^2(\mathbb{R}^n)$ is a Hilbert space with $(u, v) = \int \overline{u(x)} v(x) dx$.
3. Any closed subspace of a Hilbert space is a Hilbert space.
4. Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces then $\mathcal{H}_1 \times \mathcal{H}_2$ (= all pairs $< u_1, u_2 >$ with $u_1 \in \mathcal{H}_1$ and $u_2 \in \mathcal{H}_2$) is a Hilbert space when supplied with the inner product

$$(< u_1, u_2 >, < v_1, v_2 >) = (u_1, v_1) + (u_2, v_2) \tag{A.7}$$

This Hilbert space is denoted $\mathcal{H}_1 \oplus \mathcal{H}_2$.

If $\mathcal{S}$ is any subspace of a Hilbert space $\mathcal{H}$, not necessarily closed, then

$$\mathcal{S}^{\perp} = \{u \in \mathcal{H} : (u, v) = 0 \text{ for all } v \in \mathcal{S}\} \tag{A.8}$$

is the orthogonal subspace. Then $\mathcal{S}^{\perp}$ is closed and $(\mathcal{S}^{\perp})^{\perp} = \bar{S}$, the closure of $\mathcal{S}$.

Theorem A.2 (Projection theorem) *Given a closed subspace $\mathcal{M}$ of $\mathcal{H}$, any $u \in \mathcal{H}$ can be uniquely written as a sum $u = u_1 + u_2$ where $u_1 \in \mathcal{M}$ and $u_2 \in \mathcal{M}^{\perp}$.*

Then the map $u \rightarrow < u_1, u_2 >$ gives a natural isomorphism between $\mathcal{H}$ and $\mathcal{M} \oplus \mathcal{M}^{\perp}$ and we write

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp} \tag{A.9}$$

A linear functional L on $\mathcal{H}$ is a linear function from $\mathcal{H}$ to $\mathbb{C}$. We say L is bounded if $|L(u)| \leq C\|u\|$ for some constant C. A bounded linear functional is a continuous linear functional, and the converse is also true. The space of all bounded linear functionals denoted $\mathcal{H}'$ is a normed space with

$$\|L\| = \sup_{u \neq 0} \frac{|L(u)|}{\|u\|} = \sup_{\|u\|=1} |L(u)| \tag{A.10}$$

Then $\|L(u)\| \leq \|L\|\|u\|$. The space $\mathcal{H}'$ is also a Hilbert space called the dual space of $\mathcal{H}$. The dual space can be identified with $\mathcal{H}$ because:

Theorem A.3 (Riesz representation theorem) *Let L be a bounded linear functional on a Hilbert space $\mathcal{H}$. Then there is a unique $v \in \mathcal{H}$ such that $L(u) = (v, u)$. Furthermore $\|L\| = \|v\|$.*

An *orthonormal set* in $\mathcal{H}$ is a sequence of vectors $\{\phi_i\}$ with $i = 1, 2, \cdots$ such that

$$(\phi_i, \phi_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{A.11}$$

Theorem A.4 *The following conditions on an orthonormal set $\{\phi_i\}$ are equivalent*

1. $(\phi_i, f) = 0$ *for all* i *implies* $f = 0$.
2. *The subspace of finite linear combinations of the $\{\phi_i\}$ is dense.*
3. $f = \sum_{i=1}^{\infty} (\phi_i, f)\phi_i$ *for* $f \in \mathcal{H}$.
4. $\|f\|^2 = \sum_{i=1}^{\infty} |(\phi_i, f)|^2$ *for* $f \in \mathcal{H}$.
5. $(f, g) = \sum_{i=1}^{\infty} (f, \phi_i)(\phi_i, g)$ *for* $f, g \in \mathcal{H}$.

If one and hence all of these conditions hold, the $\{\phi_i\}$ are said to be complete and constitute an *orthonormal basis*.

B Appendix B Tensor product

We give a definition of the tensor product of two Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$. Given $u \in \mathcal{H}_1$ and $v \in \mathcal{H}_2$ we define an anti-linear functional $u \otimes v$ on $\mathcal{H}_1 \times \mathcal{H}_2$ by

$$(u \otimes v) < w_1, w_2 >= (w_1, u)(w_2, v) \tag{B.1}$$

We have

$$\begin{aligned} u_1 \otimes v + u_2 \otimes v &= (u_1 + u_2) \otimes v \\ (\alpha u \otimes v) &= \alpha(u \otimes v) = (u \otimes \alpha v) \end{aligned} \tag{B.2}$$

for $\alpha \in \mathbb{C}$. The *algebraic tensor product* $\mathcal{H}_1 \check{\otimes} \mathcal{H}_2$ is defined to be the space of functionals which are finite combinations of the $u \otimes v$, that is all functionals of the form $\sum_j u_j \otimes v_j$. A particular functional may have more than one representation of this form.

We want to define an inner product on the algebraic tensor product so that

$$\left(\sum_j u_j \otimes v_j, \sum_k u'_k \otimes v'_k \right) = \sum_{j,k} (u_j, u'_k)(v_j, v'_k) \tag{B.3}$$

Note that this entails

$$\|u \otimes v\| = \|u\| \|v\| \tag{B.4}$$

Lemma B.1 *Equation (B.3) defines an inner product on* $\mathcal{H}_1 \check{\otimes} \mathcal{H}_2$.

Proof First we must check that it is well-defined, that is independent of the representation. It suffices to show that if $\sum_j u_j \otimes v_j = 0$, then the inner product with any other element is zero. But this follows since

$$\left(\sum_k u'_k \otimes v'_k, \sum_j u_j \otimes v_j \right) = \sum_k \left(\sum_j u_j \otimes v_j \right) < u'_k, v'_k >= 0 \tag{B.5}$$

It is straightforward to check that the inner product is bilinear. We must check that it is positive definite. Given $\Psi \in \mathcal{H}_1 \check{\otimes} \mathcal{H}_2$ we can pick orthonormal bases $\{\phi_a\}$ for $\mathcal{H}_1$ and $\{\psi_b\}$ for $\mathcal{H}_2$ such that

$$\Psi = \sum_{ab} c_{ab}\, \phi_a \otimes \psi_b \tag{B.6}$$

with only a finite number of the c_{ab} not equal to zero. Then

$$(\Psi, \Psi) = \sum_{aba'b'} \bar{c}_{ab} c_{a'b'} (\phi_a \otimes \psi_b, \phi_{a'} \otimes \psi_{b'}) = \sum_{ab} |c_{ab}|^2 \geq 0 \tag{B.7}$$

If the inner product is zero, then all $c_{ab} = 0$ and hence $\Psi = 0$. □

Now we define the *tensor product* $\mathcal{H}_1 \otimes \mathcal{H}_2$ to be the Hilbert space which is the completion of $\mathcal{H}_1 \check{\otimes} \mathcal{H}_2$ in the inner product (B.3).

Lemma B.2

1. *If D_1 is a dense subspace of $\mathcal{H}_1$ and D_2 is a dense subspace of $\mathcal{H}_2$, then the algebraic tensor product $D_1 \check{\otimes} D_2$ is a dense subspace of $\mathcal{H}_1 \otimes \mathcal{H}_2$.*
2. *If $\{\phi_a\}$ is an orthonormal basis for $\mathcal{H}_1$ and $\{\psi_b\}$ is an orthonormal basis for $\mathcal{H}_2$, then $\phi_a \otimes \psi_b$ is an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$.*

Proof Given $\Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ and $\epsilon > 0$ choose $\sum_{k=1}^{N} u_k \otimes v_k$ so that

$$\|\Psi - \sum_{k=1}^{N} u_k \otimes v_k\| < \frac{\epsilon}{3} \tag{B.8}$$

Now let $M = \sup_k \{\|u_k\|, \|v_k\|\}$ and choose $u'_k \in D_1$ and $v'_k \in D_2$ so that $\|u_k - u'_k\| \leq \epsilon/3MN$ and $\|v_k - v'_k\| \leq \epsilon/3MN$. Then we have

$$\begin{aligned} &\|\sum_{k=1}^{N} u_k \otimes v_k - \sum_{k=1}^{N} u'_k \otimes v'_k\| \\ &\leq \sum_{k=1}^{N} \|(u_k - u'_k) \otimes v_k\| + \sum_{k=1}^{N} \|u'_k \otimes (v_k - v'_k)\| \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} \end{aligned} \tag{B.9}$$

Combining the above gives

$$\|\Psi - \sum_{k=1}^{N} u'_k \otimes v'_k\| < \epsilon \tag{B.10}$$

and proves the first point.

For the second point let D_1 be the finite span of the first basis which is dense in $\mathcal{H}_1$ and let D_2 be the finite span of the second basis which is dense in $\mathcal{H}_2$. Then $D_1 \otimes D_2$ is the finite span of $\{\phi_a \otimes \psi_b\}$ which is therefore dense. □

If the Hilbert spaces are L^2 spaces, then we can also identify the tensor product as an L^2 space:

Theorem B.1 *There is a unitary operator U from $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}^m)$ to $L^2(\mathbb{R}^{n+m})$ such that*

$$\Big(U(u \otimes v)\Big)(x, y) = u(x)v(y) \tag{B.11}$$

Proof Let $\{\phi_a(x)\}$ be an orthonormal basis for $L^2(\mathbb{R}^n)$ and let $\{\psi_b(y)\}$ be an orthonormal basis for $L^2(\mathbb{R}^m)$. Then $\{\phi_a(x)\psi_b(y)\}$ is an orthonormal set in $L^2(\mathbb{R}^{n+m})$. We define U on the finite span of the $\phi_a \otimes \psi_b$ by

$$\Big(U\Big(\sum_{ab} c_{ab}\, \phi_a \otimes \psi_b\Big)\Big)(x,y) = \sum_{ab} c_{ab}\, \phi_a(x)\psi_b(y) \tag{B.12}$$

This is densely defined and norm preserving and so extends to an isometry. To show U is unitary we have to show that the range is dense, that is that $\{\phi_a(x)\psi_b(y)\}$ are complete in $L^2(\mathbb{R}^{n+m})$.

Suppose that $f(x,y) \in L^2(\mathbb{R}^{n+m})$ is orthogonal to all $\{\phi_a(x)\psi_b(y)\}$. By Fubini's theorem

$$\int \left(\int f(x,y)\phi_a(x)dx\right) \psi_b(y)dy = 0 \tag{B.13}$$

Since the ψ_b are dense, this implies $\int f(x,y)\phi_a(x)dx = 0$ for almost every y. Since the ϕ_a are dense, this implies $f(x,y) = 0$ for almost every (x,y), hence $f = 0$ in $L^2(\mathbb{R}^{n+m})$ as required.

It is now straightforward to check (B.11) to complete the proof. □

If S is a bounded operator on $\mathcal{H}_1$ and T is a bounded operator on $\mathcal{H}_2$, then we define an operator $S \otimes T$ on $\mathcal{H}_1 \check{\otimes} \mathcal{H}_2$ by

$$(S \otimes T)\left(\sum_k u_k \otimes v_k\right) = \sum_k Su_k \otimes Tv_k \tag{B.14}$$

This is well-defined since if $\sum_k u_k \otimes v_k = 0$, then

$$\left(\sum_k Su_k \otimes Tv_k\right) < w_1, w_2 >= \left(\sum_k u_k \otimes v_k\right) < S^*w_1, T^*w_2 >= 0 \tag{B.15}$$

Lemma B.3 *$S \otimes T$ is bounded and extends to a bounded operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$ satisfying*

$$\|S \otimes T\| = \|S\|\|T\| \tag{B.16}$$

Proof First consider the operator $S \otimes I$. Let $\{\phi_a\}$ be an orthonormal basis for $\mathcal{H}_1$ and let $\{\psi_b\}$ be an orthonormal basis for $\mathcal{H}_2$. Consider vectors which are finite sums of basis vectors of the form

$$\Psi = \sum_{ab} c_{ab}\, \phi_a \otimes \psi_b \tag{B.17}$$

Then we have

$$\begin{aligned}\|(S\otimes I)\Psi\|^2 &= \|\sum_{ab} c_{ab}\, S\phi_a\otimes\psi_b\|^2 = \|\sum_b\left(\sum_a c_{ab}\, S\phi_a\right)\otimes\psi_b\|^2\\ &= \sum_b\|\sum_a c_{ab}\, S\phi_a\|^2 \le \|S\|^2\sum_b\|\sum_a c_{ab}\,\phi_a\|^2 \qquad \text{(B.18)}\\ &= \|S\|^2\sum_{ab}|c_{ab}|^2 = \|S\|^2\|\Psi\|^2\end{aligned}$$

Thus $S\otimes I$ is bounded. Similarly $I\otimes T$ is bounded. It follows that $S\otimes T$ is bounded since $S\otimes T = (S\otimes I)(I\otimes T)$. Hence it extends to a bounded operator on $\mathcal{H}_1\otimes\mathcal{H}_2$ and it is straightforward to check that the extension satisfies (B.14). We have also

$$\|S\otimes T\| \le \|S\|\|T\| \qquad \text{(B.19)}$$

We omit the proof that this is actually an equality. □

Reference: Reed and Simon (1980).

C Appendix C **Distributions**

Distributions are a generalization of functions. In this appendix we consider a special class of distributions called tempered distributions.

Recall from chapter 1 that the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is the space of all complex-valued infinitely differentiable functions on $\mathbb{R}^d$ such that for any choice of multi-indices α, β we have $\|x^\beta D^\alpha f\|_\infty < \infty$. A *tempered distribution* T is an element of the dual space $\mathcal{S}'(\mathbb{R}^d)$, that is it is a continuous linear functional from $\mathcal{S}(\mathbb{R}^d)$ to $\mathbb{C}$. To each $f \in \mathcal{S}(\mathbb{R}^d)$ it assigns a complex number denoted $T(f)$ or $< T, f >$.

To complete the definition we have to specify what "continuous" means in this situation and this means specifying a topology for $\mathcal{S}(\mathbb{R}^d)$. Since $\mathcal{S}(\mathbb{R}^d)$ is not a Banach space, we do not have a norm to help us. Instead the topology is specified by the family of semi-norms $\|x^\beta D^\alpha f\|_\infty$. (A semi-norm $\rho(f)$ has the properties of a norm except that $\rho(f) = 0$ need not imply $f = 0$.) Skipping the exact definition of the topology we say that a linear functional T on $\mathcal{S}(\mathbb{R}^d)$ is continuous if there exists a semi-norm

$$\|f\|_{nm} = \sup_{|\alpha| \le n, |\beta| \le m} \|x^\beta D^\alpha f\|_\infty \tag{C.1}$$

and a constant C such that

$$|T(f)| \le C \|f\|_{nm} \tag{C.2}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$.

Examples:

1. Let g be a polynomially bounded measurable function on $\mathbb{R}^d$. Then there is an associated distribution T_g defined by

$$< T_g, f >= \int g(x) f(x) dx \tag{C.3}$$

Polynomially bounded means that $h(x) = (1 + |x|^2)^{-N} g(x)$ is in $L^1(\mathbb{R}^d)$ for N sufficiently large and so we have

$$|T_g(f)| \le \|h\|_1 \sup_x |(1 + |x|^2)^N f(x)| \tag{C.4}$$

This can be dominated by a constant times a norm $\|f\|_{nm}$ and hence T_g is a tempered distribution. Usually we would write $< g, f >$ instead of $< T_g, f >$.

2. Let μ be a measure on $\mathbb{R}^d$ such that for some N

$$\int (1+|x|^2)^{-N} d\mu(x) < \infty \tag{C.5}$$

Then we can define a tempered distribution by

$$T_\mu(f) = \int f(x) d\mu(x) \tag{C.6}$$

The proof is similar to the previous example.

3. The delta function δ_{x_0} at $x_0 \in \mathbb{R}^d$ is the tempered distribution defined by

$$\delta_{x_0}(f) =< \delta_{x_0}, f >= f(x_0) \tag{C.7}$$

This is a special case of the previous example. The measure is the point measure at x.

4. Given complex numbers $c_1, \ldots, c_n$, multi-indices $\alpha_1, \ldots, \alpha_n$, and points $x_1, \ldots, x_n$ in $\mathbb{R}^d$ there is a tempered distribution defined by

$$T(f) = \sum_i c_i (D^{\alpha_i} f)(x_i) \tag{C.8}$$

These are generally not given as functions or measures.

In the first example we have seen that polynomially bounded functions determine distributions. In fact it turns out that the map $g \to T_g$ is injective in the sense that if $T_g = T_h$, then $g = h$ almost everywhere.[1] Thus we can identify such functions as a subspace of $\mathcal{S}'(\mathbb{R}^d)$. Accordingly it is appropriate to refer to the tempered distributions as *generalized functions*.

Even for distributions which are not functions it is sometimes convenient to write $< T, f >$ as if it were a function by $< T, f >= \int T(x) f(x) dx$ with some suggestive symbol $T(x)$. In particular we write:

$$< \delta_{x_0}, f >= \int \delta(x - x_0) f(x) dx \tag{C.9}$$

Let $\mathcal{O} \subset \mathbb{R}^d$ be open and let $\mathcal{C}_0^\infty(\mathcal{O})$ be the infinitely differentiable functions with compact support in $\mathcal{O}$. A distribution T is said to vanish on $\mathcal{O}$ if $< T, f >= 0$ for all $f \in \mathcal{C}_0^\infty(\mathcal{O})$. The *support* of T is the smallest closed set such that T vanishes on the complement. For example the support of δ_{x_0} is the single point x_0.

Next we define some operations on tempered distributions.

1. (Multiplication by a smooth function) Suppose that h is a smooth polynomially bounded function and $T \in \mathcal{S}'(\mathbb{R}^d)$. Then we can define $hT \in \mathcal{S}'(\mathbb{R}^d)$ by

$$< hT, f >=< T, hf > \tag{C.10}$$

[1] If g, h are in $L^2(\mathbb{R}^d)$ this is immediate since $T_g = T_h$ implies that $(\bar{g} - \bar{h}, f) = 0$ for all f in the dense set $\mathcal{S}(\mathbb{R}^d)$, hence for all $f \in L^2$, hence $g - h = 0$ in L^2.

This makes sense since if $f \in \mathcal{S}(\mathbb{R}^d)$, then $hf \in \mathcal{S}(\mathbb{R}^d)$. The definition satisfies $hT_g = T_{hg}$.

2. (Derivatives) Any tempered distribution T has a partial derivative $\partial_\mu T = \partial T/\partial x^\mu$, which is the tempered distribution defined by

$$< \partial_\mu T, f >= - < T, \partial_\mu f > \tag{C.11}$$

If the distribution is a differentiable function, say T_g with $g \in \mathcal{S}(\mathbb{R}^d)$, then integrating by parts this is computed as

$$- < T_g, \partial_\mu f >= - \int g(x)\partial_\mu f(x)dx = \int \partial_\mu g(x) f(x)dx =< T_{\partial_\mu g}, f > \tag{C.12}$$

Hence $\partial_\mu T_g = T_{\partial_\mu g}$ and the definition extends the definition on smooth functions.

It follows that any linear differential operator can be applied to a distribution, even operators with smooth variable coefficients.

Examples: In $d = 1$ consider the distribution defined by the Heaviside function θ which is the characteristic function of $[0, \infty)$. Then

$$< \frac{d\theta}{dx}, f >= - < \theta, \frac{df}{dx} >= - \int_0^\infty \frac{df}{dx} = f(0) \tag{C.13}$$

and thus

$$\frac{d\theta}{dx} = \delta_0 \tag{C.14}$$

3. (Fourier transform) Since the Fourier transform maps $\mathcal{S}(\mathbb{R}^d)$ to itself (see chapter 1), we can define the Fourier transform on a distribution T by

$$< \mathcal{F}T, f >=< T, \mathcal{F}f > \tag{C.15}$$

This agrees with the definition on functions, that is if $g \in L^2(\mathbb{R}^d)$, then $\mathcal{F}T_g = T_{\mathcal{F}g}$. To see this use the fact the $\mathcal{F}$ is unitary on $L^2(\mathbb{R}^d)$ to compute for $f \in \mathcal{S}(\mathbb{R}^d)$

$$\begin{aligned} &< \mathcal{F}T_g, f >=< T_g, \mathcal{F}f >= (\bar{g}, \mathcal{F}f) \\ &= (\mathcal{F}^{-1}\bar{g}, f) = (\overline{\mathcal{F}g}, f) =< T_{\mathcal{F}g}, f > \end{aligned} \tag{C.16}$$

Also $\mathcal{F}$ is a bijection on $\mathcal{S}'(\mathbb{R}^d)$; this follows from the fact that $\mathcal{F}$ is a bijection on $\mathcal{S}(\mathbb{R}^d)$.

Examples: We have $(\mathcal{F}\delta_x)(p) = (2\pi)^{-d/2} e^{-ipx}$ since

$$< \mathcal{F}\delta_x, f >=< \delta_x, \mathcal{F}f >= \mathcal{F}f(x) = (2\pi)^{-d/2} \int e^{-ipx} f(p)dp \tag{C.17}$$

Other examples are

$$\mathcal{F}(1) = (2\pi)^{d/2}\delta_0 \qquad \mathcal{F}(\partial_\mu \delta_0) = ip_\mu (2\pi)^{-d/2} \tag{C.18}$$

Next we quote the *kernel theorem* which says that a multilinear functional on $\mathcal{S}$ has a kernel in $\mathcal{S}'$.

Theorem C.1 *Let $T(f_1, \ldots, f_n)$ be a multilinear functional continuous in each $f_i \in \mathcal{S}(\mathbb{R}^d)$. Then there is a unique $T \in \mathcal{S}'(\mathbb{R}^{nd})$ such that*

$$T(f_1, \ldots, f_n) = T(f_1 \otimes \cdots \otimes f_n) \tag{C.19}$$

where $f_1 \otimes \cdots \otimes f_n \in \mathcal{S}(\mathbb{R}^{nd})$ is defined by

$$(f_1 \otimes \cdots \otimes f_n)(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n) \tag{C.20}$$

Examples: Let T be a bounded operator on $L^2(\mathbb{R}^d)$. Then $(\bar{f}_1, Tf_2)$ is a continuous bilinear function on $\mathcal{S}(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d)$ and so by the theorem there is a $T \in \mathcal{S}'(\mathbb{R}^{2d})$ such that $(\bar{f}_1, Tf_2) = T(f_1 \otimes f_2)$. Replacing f_1 by $\bar{f}_1$ and denoting the distribution T by $T(x, y)$ this can also be written

$$(f_1, Tf_2) = \int \bar{f}_1(x) T(x, y) f_2(y) dx dy \tag{C.21}$$

Then $T(x, y)$ is called the kernel of the operator. (Not to be confused with the null space, which is also called the kernel.) For example the kernel of the identity is $\delta(x - y)$.

For completeness we quote the general definition of distributions. These are defined in open sets $\mathcal{O} \subset \mathbb{R}^d$. Let $\mathcal{D}(\mathcal{O}) = \mathcal{C}_0^\infty(\mathcal{O})$, the infinitely differentiable functions on $\mathcal{O}$ with compact support. A *distribution* T is an element of the dual space $\mathcal{D}'(\mathcal{O})$, that is it is a linear functional $f \to T(f)$ on $\mathcal{D}(\mathcal{O})$ which is continuous in the sense that for every compact subset $K \subset \mathcal{O}$ there are constants k, C such that

$$|T(f)| \leq C \sup_{|\alpha| \leq k, x \in K} |(D^\alpha f)(x)| \tag{C.22}$$

for all $f \in \mathcal{D}(\mathcal{O})$ with $\operatorname{supp} f \subset K$.

Tempered distributions are distributions in this sense: $\mathcal{S}'(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d)$.

References: Yosida (1966), Reed and Simon (1980), or Taylor (1996).

References

Abraham, R. and Marsden, J. 1978. *Foundations of Mechanics*. Benjamin.

Albeverio, S., Gesztesy, G., Hoegh-Krohn, R., and Holden, H. 1988. *Solvable Models in Quantum Mechanics*. Springer.

Albeverio, S., Hoegh-Krohn, R., and Mazzucchi, S. 2008. *Mathematical Theory of Feynman Path Integrals: An Introduction*. Springer.

Bachelot, A. 1999. The Hawking effect. *Ann. Inst. Henri Poincaré*, **70**, 41–99.

Baez, J., Segal, I., and Zhou, Z. 1992. *Introduction to Algebraic and Constructive Quantum Field Theory*. Princeton University Press.

Bär, C., Ginoux, N., and Pfäffle, F. 2007. *Wave Equations on Lorentzian Manifolds and Quantization*. European Mathematical Society Publishing House.

Billingsley, P. 1979. *Probability and Measure*. Wiley.

Bogolubov, N.N., Logunov, A.A., and Todorov, I.T. 1975. *Introduction to Axiomatic Quantum Field Theory*. Benjamin.

Bratteli, O. and Robinson, D. 1981. *Operator Algebras and Quantum Statistical Mechanics II*. Springer.

Brunetti, R., Fredenhagen, K., and Verch, R. 2003. The generally covariant locality principle - a new paradigm for local quantum field theory. *Commun. Math. Phys.*, **237**, 31–68.

Choquet-Bruhat, Y., DeWitt-Morette, C., and Dillard-Bleick, M. 1977. *Analysis, Manifolds, and Physics I*. North Holland.

Choquet-Bruhat, Y. and DeWitt-Morette, C. 1989. *Analysis, Manifolds, and Physics II*. North Holland.

Cycon, H., Froese, R., Kirsch, W., and Simon, B. 1987. *Schrödinger Operators*. Springer.

Dimock, J. 1980. Algebras of local observables on a manifold. *Commun. Math. Phys.*, **77**, 219–228.

Drechsler, W. and Mayer, M. 1977. *Fiber Bundle Techniques in Gauge Theory*. Springer.

Durrett, R. 1996. *Stochastic Calculus*. CRC press.

Folland, G. 2008. *Quantum Field Theory: A Tourist Guide for Mathematicians*. American Mathematical Society.

Frankel, T. 2004. *The Geometry of Physics*. Cambridge University Press.

Gallavotti, G. 1983. *The Elements of Mechanics*. Springer.

Gelfand, I.M. and Vilenkin, N. 1964. *Generalized Functions IV*. Academic Press.

Glimm, J. and Jaffe, A. 1970. Quantum field theory models, pages 1–108 of: DeWitt, C. and Stora, R. (eds), *Statistical Mechanics and Quantum Field Theory*. Gordon & Breach.

Glimm, J. and Jaffe, A. 1972. The $(\lambda\phi^4)_2$ quantum field theory without cutoffs IV. Perturbations of the Hamiltonian. *J. Math. Phys.*, **13**, 1568–1584.

Glimm, J. and Jaffe, A. 1987. *Quantum Physics: A Functional Integral Point of View*. Springer.

Glimm, J., Jaffe, A., and Spencer, T. 1974. The Wightman axioms and particle structure in the $P(\phi)_2$ quantum field model. *Ann. Math.*, **100**, 585–632.

Gustafson, S. and Sigal, I.M. 2003. *Mathematical Concepts of Quantum Mechanics*. Springer.

Haag, R. 1992. *Local Quantum Physics*. Springer.

Hall, B. 2003. *Lie Groups, Lie Algebras, and Representations*. Springer.

Huang, K. 2009. *Introduction to Statistical Physics*. CRC Press.

Isham, C.J. 1995. *Lectures on Quantum Theory: Mathematical and Structural Foundations*. Imperial College Press.

Itzykson, C. and Zuber, J. 1980. *Quantum Field Theory*. McGraw-Hill.

Jauch, J. 1968. *Foundations of Quantum Mechanics*. Addison-Wesley.

Kato, T. 1966. *Perturbation Theory for Linear Operators*. Springer.

Kay, B. 1979. A uniqueness result in the Segal–Weinless approach to linear Bose fields. *J. Math. Phys.*, **20**, 1712–1713.

Landau, L. and Lifschitz, E. 1969. *Statistical Physics*. Addison-Wesley.

Lieb, E. and Seiringer, R. 2010. *The Stability of Matter in Quantum Mechanics*. Cambridge University Press.

Marsden, J. and Ratiu, T. 1994. *Introduction to Mechanics and Symmetry*. Springer.

Miller, W. 1972. *Symmetry Groups and Their Applications*. Academic Press.

Misner, C., Thorne, K., and Wheeler, J. 1973. *Gravitation*. W.H. Freeman.

Nelson, E. 1973. Probability theory and Euclidean field theory, pages 94–124 of: Velo, G. and Wightman, A. (eds), *Constructive Quantum Field Theory*. Springer.

von Neumann, J. 1955. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press.

Newton, T. and Wigner, E. 1949. Localized states for elementary systems. *Rev. Mod. Phys.*, **21**, 400–406.

Ohnuki, Y. 1988. *Unitary Representations of the Poincaré Group and Relativistic Wave Equations*. World Scientific.

Peskin, M. and Schroeder, D. 1995. *An Introduction to Quantum Field Theory*. Perseus.

Reed, M. and Simon, B. 1980. *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis*. Academic Press.

Reed, M. and Simon, B. 1975. *Methods of Modern Mathematical Physics, Vol. II: Fourier Analysis, Self-adjointness*. Academic Press.

Reed, M. and Simon, B. 1979. *Methods of Modern Mathematical Physics, Vol. III: Scattering Theory*. Academic Press.

Reed, M. and Simon, B. 1978. *Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators*. Academic Press.

Roman, P. 1969. *Introduction to Quantum Field Theory*. Wiley.

Ruelle, D. 1969. *Statistical Mechanics*. Benjamin.

Sachs, R. and Wu, H.-H. 1977. *General Relativity for Mathematicians*. Springer.

Salmhofer, M. 1999. *Renormalization: An Introduction*. Springer.

Schweber, S. 1962. *An Introduction to Relativistic Quantum Field Theory*. Harper & Row.

Segal, I.E. 1956. Tensor algebras over Hilbert spaces. *Trans. Am. Math. Soc.*, **81**, 106–134.

Simon, B. 1975. *The $P(\phi)_2$ Euclidean Field Theory*. Princeton University Press.

Simon, B. 1979. *Functional Integration and Quantum Physics*. Academic Press.

Streater, R. and Wightman, A. 1964. *PCT, Spin-Statistics, and All That*. Benjamin.

Strocchi, F. and Wightman, A. 1974. Proof of the charge superselection rule in local relativistic quantum field theory. *J. Math. Phys.*, **15**, 2198–2224.

Takhtajan, L. 2008. *Quantum Mechanics for Mathematicians*. American Mathematical Society.

Taylor, M. 1996. *Partial Differential Equations I: Basic Theory*. Springer.

Titchmarsh, E. 1939. *The Theory of Functions*. Oxford University Press.

Wald, R. 1994. *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. University of Chicago Press.

Weinberg, S. 1995. *The Quantum Theory of Fields*. Cambridge University Press.

Yosida, K. 1966. *Functional Analysis*. Springer.

Index