SPRINGER BRIEFS IN PHYSICS

Pankaj Sharan

Some Unusual Topics in Quantum Mechanics

Springer

SpringerBriefs in Physics

SpringerBriefs in Physics are a series of slim high-quality publications encompassing the entire spectrum of physics. Manuscripts for SpringerBriefs in Physics will be evaluated by Springer and by members of the Editorial Board. Proposals and other communication should be sent to your Publishing Editors at Springer.

Featuring compact volumes of 50 to 125 pages (approximately 20,000–45,000 words), Briefs are shorter than a conventional book but longer than a journal article. Thus, Briefs serve as timely, concise tools for students, researchers, and professionals. Typical texts for publication might include:

- A snapshot review of the current state of a hot or emerging field
- A concise introduction to core concepts that students must understand in order to make independent contributions
- An extended research report giving more details and discussion than is possible in a conventional journal article
- A manual describing underlying principles and best practices for an experimental technique
- An essay exploring new ideas within physics, related philosophical issues, or broader topics such as science and society

Briefs allow authors to present their ideas and readers to absorb them with minimal time investment. Briefs will be published as part of Springer's eBook collection, with millions of users worldwide. In addition, they will be available, just like other books, for individual print and electronic purchase. Briefs are characterized by fast, global electronic dissemination, straightforward publishing agreements, easy-to-use manuscript preparation and formatting guidelines, and expedited production schedules. We aim for publication 8–12 weeks after acceptance.

More information about this series at http://www.springer.com/series/8902

Pankaj Sharan

Some Unusual Topics in Quantum Mechanics

Springer

Pankaj Sharan
(emeritus)
Department of Physics
Jamia Millia Islamia University
New Delhi, India

ISSN 2191-5423 ISSN 2191-5431 (electronic)
SpringerBriefs in Physics
ISBN 978-3-030-60417-2 ISBN 978-3-030-60418-9 (eBook)
https://doi.org/10.1007/978-3-030-60418-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

Quantum Theory is about 120 years old, and Quantum Mechanics proper, a truly spectacular achievement of twentieth century physics, will soon be a century old. As Abraham Pais once said, quantum theory is 'a uniquely twentieth century mode of thought'.

Quantum mechanics has profound geometric and algebraic features, with probability ideas too playing a crucial role. It has a richness about it, and is still growing, as seen through the developments of weak measurement, the Zeno effect, the geometric phase idea, entanglement and quantum information and computation in recent decades. It is a core subject in the teaching of physics at upper undergraduate, graduate and research levels. Over a long period of time, a very large number of texts have been written, several of them becoming classics of the physics literature. Every teacher of quantum mechanics necessarily has to make some selection of relatively advanced topics to cover, after the basic or irreducible core has been presented. Of course, the successful applications are legion.

Pankaj Sharan is a truly gifted and experienced teacher of the subject having taught it for many decades with passion and enthusiasm. In this book, he has put together a set of concise treatments of special topics that are usually not found in most texts. These include the role of particle position as seen in going from the relativistic to the Galilean domain; a picture of quantum mechanics in the mathematical language of fibre bundles; the basic concepts and calculational methods in scattering theory such as the Moller and S matrices, and their integral equations; the formulation of the subject in the classical phase space language based on the pioneering ideas of Weyl, Wigner and later, Moyal; the motivations for and possibilities of a nonlinear extension of quantum mechanics; the connection between the classical idea of inter-particle interaction potential and the typically quantum idea of particle exchange; and the proof of the Wigner Theorem on representation of symmetries in the form given by Bargmann. The author's pedagogical skills built up over many years are evident in every one of these treatments.

This book provides enrichment material for both students and teachers—as possible projects for individual students to work on and present as special lectures in a regular course, and for teachers to possibly include some interesting ideas in

their classroom treatment of the subject. Thanks and congratulations to the author for this excellent supplementary material every good course on quantum mechanics can draw upon.

Bengaluru, India N. Mukunda

Preface

These topics are unusual only in the sense that they are not on the syllabi of most courses on Quantum Mechanics, and they are usually not treated in textbooks. The book is based on notes for classroom lectures and group seminars on these topics given during four decades of my teaching career.

Three of the seven chapters are on the conceptual difficulties students face: position operators (Chap. 1), plane waves versus real beams of particles (Chap. 3), and the concept of potential in quantum mechanics (Chap. 6).

The other three chapters are on different ways of looking at quantum theory: the fibre bundle approach (Chap. 2), and the intimate relation of quantum mechanics to classical Hamiltonian mechanics. In one case (Chap. 4) one learns how to do quantum mechanics on the phase space, and in the other (Chap. 5), where quantum theory itself is a linear Hamiltonian mechanics, the question is posed whether it is possible to conceive of a non-linear generalization of the theory.

The last chapter is a proof of the Wigner theorem on symmetry transformations. The theorem is used everywhere, but its proof is omitted in most textbooks. I think Bargmann's proof of the theorem presented here is aesthetically appealing and all students must be encouraged to go through it once.

I hope that this short book will be helpful in motivating students of quantum mechanics to explore and acquire a deeper understanding of the subject.

New Delhi, India — Pankaj Sharan

Acknowledgements

I am thankful to Q. N. Usmani and Tabish Qureshi for discussions on various aspects of quantum mechanics.

I am grateful to Vasundhra Choudhry for collaboration on a part of the contents of Chap. 5. I thank Pravabati Chingangbam for collaboration on the matter included in Chap. 2 and for reading and correcting parts of the manuscript.

I am grateful to H. S. Mani and A. K. Kapoor who have helped me in various ways, practically all the time. And I feel privileged that N. Mukunda, from whom I have learned much through his lectures and papers, has written a Foreword to this little book.

I thank B. Ananthanarayan and Lisa Scalone from the editorial board of SpringerBriefs for helpful advice at all stages of preparation of this book.

Finally, it is a pleasure to thank all my colleagues at Jamia Millia Islamia, New Delhi, and particularly, Lekha Nair, M. Sami, Somasri Sen, Asad Niazi, Zubaida Ansari, and Shafeeque Ansari for creating an atmosphere of cheerful devotion to teaching and academic work.

Pankaj Sharan

Contents

Notation

We use both ϕ, ψ, f, g, etc., and $|\phi\rangle, |\psi\rangle, |f\rangle, |g\rangle$, etc., for vectors in a Hilbert space $\mathcal{H}$. The inner product denoted by (ϕ, ψ) or $\langle\phi|\psi\rangle$, is linear in the second argument and anti-linear in the first. The norm $||\phi||$ is the positive square root $\sqrt{(\phi, \phi)}$. We also write the multiplication of a Hilbert space vector $|\phi\rangle$ or ϕ by a complex number c from left or from right as convenient: $c|\phi\rangle = |\phi\rangle c$ or $c\phi = \phi c$ etc.

An operator in a Hilbert space is written with a 'hat': for example $\widehat{x}$, when it has to be distinguished from the corresponding classical quantity x. The hat is omitted when confusion is unlikely. We use the words 'self-adjoint' and 'Hermitian' interchangeably. A dagger (†) denotes a Hermitian adjoint, and a (*) the complex conjugate.

The Minkowski metric in our notation is:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix},$$

$\mu, \nu = 0, 1, 2, 3$.

Three dimensional vectors are written in bold face $\mathbf{p} = (p^1, p^2, p^3)$ or $p^i = p_i, i = 1, 2, 3$.

All the components of momentum 4-vector $p = (p^0, p^1, p^2, p^3)$ of a particle of proper mass m have the physical dimensions of momentum. The 0-component is also denoted by ω at some places, for example: $\omega_{\mathbf{p}} = p^0 = \sqrt{\mathbf{p}^2 + m^2c^2}$. The energy of the particle is denoted by $E_p = c\omega_{\mathbf{p}}$. Similarly, all the components of spacetime 4-vector have dimensions of distance $x = (x^0 = ct, x^1, x^2, x^3)$.

The identity operator or identity matrix will be generically denoted by $\mathbf{1}$.

Derivatives $\partial f/\partial x^i$ are often abbreviated by $\partial_i f$ or $f_{,i}$ when there is no confusion.

Chapter 1
Position Operators of Non-relativistic Quantum Mechanics

The components of the position operator of a particle in non-relativistic quantum mechanics are the non-relativistic limits of generators of the pure Lorentz transformations, or the 'boost operators'. The non-relativistic limit of the relativistic Poincare group is the Galilean group, which reveals interesting complications.

1.1 Introduction

The position operators $\hat{x}_i, i = 1, 2, 3$ of a particle in non-relativistic quantum mechanics are perhaps the most important observables because of their physical meaning. But any non-relativistic theory can only be the limit of a relativistic theory. And a position operator in a relativistic theory is a problem because, in a truly relativistic theory it must be the space part of a 4-vector, whose time-component must be the physical time. But time is not an operator either in relativistic quantum mechanics or in relativistic quantum field theory. Moreover, the single particle picture itself is under threat in relativistic quantum theory. Similarly, there are problems associated with defining a centre of mass for two or more relativistic particles.

These conceptual problems have been thoroughly discussed during the 1960s and 70s. A discussion on the so-called "localization problem" can be found in [1]. A recent account with older references is [2].

Actually, the simplest view is to regard the position operators, up to a constant factor, as limits of the three boost operators of relativistic quantum mechanics. This fact somewhat surprises many practitioners of quantum theory because the boost operators do not commute among themselves!

Non-relativistic quantum mechanics relies heavily on the commutation relations

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{1.1}$$

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_1

which allows a kind of reciprocity between **x** and **p** representations. In the **x**-representation,

$$\hat{p}_i = -i\hbar \frac{\partial}{\partial x_i} \tag{1.2}$$

act as generators of space translations. Conversely, in the momentum space representation,

$$\hat{x}_i = i\hbar \frac{\partial}{\partial p_i} \tag{1.3}$$

act as generators of translations in the momentum space. But while homogeneity of space is a symmetry for a free particle, there is no such symmetry as 'homogeneity of momentum space'. The symmetry transformation which does change the momentum is the change from one Galilean inertial frame of reference to the other moving with a constant velocity with respect to it. In a relativistic theory these changes of inertial frames become 'pure' Lorentz transformations or 'boosts'. It is these that we turn to in the next section.

1.2 Spinless Relativistic Particle

Poincare transformations include the Lorentz transformations (rotations and boosts) as well as pure translations in four-dimensional spacetime. Denote such a transformation by (a, Λ) acting on the spacetime points $x = (x^0, x^1, x^2, x^3)$ as follows:

$$(a, \Lambda)x = \Lambda x + a,$$

where a is the four vector of translation, and Λ is a Lorentz matrix, that is, a 4×4 real matrix such that

$$x \cdot x = \sum_{\mu,\nu} \eta_{\mu\nu} x^\mu x^\nu = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2$$

remains unchanged. If $y = \Lambda x$ then

$$x \cdot x = \sum_{\mu,\nu} \eta_{\mu\nu} x^\mu x^\nu = \sum_{\mu,\nu} \eta_{\mu\nu} y^\mu y^\nu = y \cdot y,$$

and so the Lorentz matrix Λ satisfies

$$\Lambda^T \eta \Lambda = \eta.$$

Matrices such as Λ form a group, $\mathcal{L}$, called the *Lorentz group*. It follows from this condition that $(\det \Lambda)^2 = 1$, and by writing out the 00 element of this equation we see that $\Lambda_{00}^2 \geq 1$. The Lorentz group has a subgroup $\mathcal{L}_+^\uparrow$ of those matrices for which $\det \Lambda = +1$ and $\Lambda_{00} \geq +1$. These are the matrices which are connected to the identity matrix in a continuous fashion. This group has six parameters corresponding to rotations in three dimensional space and 'boosts', which are Lorentz transformations corresponding to frames moving with constant relative velocity but with their spatial axes remaining parallel. Every group element is a product of boosts and rotations. A general element of $\mathcal{L}$ can be obtained by multiplying an element of $\mathcal{L}_+^\uparrow$ by the matrix η (also called time inversion) or $-\eta$, (called space inversion or parity), or both. In this chapter we restrict ourselves to $\mathcal{L}_+^\uparrow$. The so-called discrete symmetries represented by parity and time reversal are treated briefly in the last chapter of this book.

The Poincare group is obtained by adjoining spacetime translations to Lorentz transformations as shown above. The group multiplications law is

$$(a_2, \Lambda_2)(a_1, \Lambda_1) = (a_2 + \Lambda_2 a_1, \Lambda_2 \Lambda_1),$$

the identity element is $(0, \mathbf{1})$ and the inverse is

$$(a, \Lambda)^{-1} = (-\Lambda^{-1} a, \Lambda^{-1}).$$

When we are dealing with only translations $(a, \mathbf{1})$, or only Lorentz transformations $(0, \Lambda)$ it is convenient to write just (a) or (Λ).

A relativistic particle of proper mass m with zero spin is described by momentum space wave functions $\psi(p)$ where $p = (p^0, p^1, p^2, p^3)$ but $p^0 = \omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2 c^2}$. For convenience, we write $\psi(p)$, with the 4-vector p as argument, it being understood that ψ is actually a function only of $\mathbf{p} = (p^1, p^2, p^3)$ and $p^0 = \omega_{\mathbf{p}}$ in what follows.

We can construct a Hilbert space $\mathcal{H}$ by restricting functions ψ to be square integrable with respect to the Lorentz invariant volume element

$$\int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_{\mathbf{p}}} |\psi(p)|^2 < \infty.$$

The inner product in $\mathcal{H}$ is defined by

$$(\psi, \phi) = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_{\mathbf{p}}} \psi^*(p)\phi(p). \tag{1.4}$$

It is convenient to use the Dirac's bracket notation: $\psi(p) = \langle p|\psi\rangle$ where $|p\rangle$ are eigenvectors of momentum such that

$$\langle p|k\rangle = 2\omega_{\mathbf{p}} \delta^3(\mathbf{p} - \mathbf{k}).$$

The action of Poincare group representation is defined on this basis as

$$U(a)|p\rangle = \exp(-ip\cdot a/\hbar)|p\rangle = \exp[(ip^0a^0 - i\mathbf{p}\cdot\mathbf{a})/\hbar])|p\rangle, \tag{1.5}$$
$$U(\Lambda)|p\rangle = |\Lambda p\rangle, \tag{1.6}$$

so that

$$U(a,\Lambda)|p\rangle = U(a)U(\Lambda)|p\rangle = \exp(-i\Lambda p\cdot a/\hbar)|\Lambda p\rangle. \tag{1.7}$$

On momentum space wave functions $\psi(p)$ it acts as follows:

$$(U(a,\Lambda)\psi)(p) = \exp(-ip\cdot a/\hbar)\psi(\Lambda^{-1}p). \tag{1.8}$$

Exercise 1.1 Verify (1.8) and check that operators $U(a,\Lambda)$ are unitary, and that they provide a representation of the Poincare group:

$$U(a_1,\Lambda_1)U(a_2,\Lambda_2) = U(a_1+\Lambda_1 a_2, \Lambda_1\Lambda_2). \tag{1.9}$$

1.3 Lie Algebra of the Poincare Group

An infinitesimal boost in 1-direction by a small α so that $\sinh\alpha \approx \alpha = v/c$ is,

$$\Lambda = \begin{pmatrix} \cosh\alpha & \sinh\alpha & 0 & 0 \\ \sinh\alpha & \cosh\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 1 + \alpha \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \cdots,$$

$$\Lambda^{-1}p = (p^0 - \alpha p^1, p^1 - \alpha p^0, p^2, p^3) + \cdots.$$

Therefore,

$$(U(\Lambda)\psi)(p) = \psi(\Lambda^{-1}p) = [(1 + i\alpha K_1/\hbar + \cdots)\psi](p)$$
$$(K_1\psi)(p) = i\hbar p^0 \frac{\partial\psi}{\partial p^1}, \tag{1.10}$$

which identifies the generator K_1. The two other generators for boosts can be similarly constructed

$$(K_2\psi)(p) = i\hbar p^0 \frac{\partial\psi}{\partial p^2}, \qquad (K_3\psi)(p) = i\hbar p^0 \frac{\partial\psi}{\partial p^3}. \tag{1.11}$$

For an infinitesimal rotation about the 1-axis,

$$\Lambda = \begin{pmatrix} 1\,0 & 0 & 0 \\ 0\,1 & 0 & 0 \\ 0\,0 & \cos\theta & \sin\theta \\ 0\,0 & -\sin\theta & \cos\theta \end{pmatrix} = 1 + \theta \begin{pmatrix} 0\,0 & 0 & 0 \\ 0\,0 & 0 & 0 \\ 0\,0 & 0 & 1 \\ 0\,0 & -1 & 0 \end{pmatrix} + \cdots ,$$
$$\Lambda^{-1} p = (p^0, p^1, p^2 - \theta p^3, p^3 + \theta p^2) + \cdots ,$$

the generator can be identified

$$(U(\Lambda)\psi)(p) = \psi(\Lambda^{-1}p) = [(1 + i\theta J_1/\hbar + \cdots)\psi](p) ,$$
$$(J_1\psi)(p) = -i\hbar \left(p^2 \frac{\partial \psi}{\partial p^3} - p^3 \frac{\partial \psi}{\partial p^2} \right). \tag{1.12}$$

The other two rotation generators are

$$(J_2\psi)(p) = -i\hbar \left(p^3 \frac{\partial \psi}{\partial p^1} - p^1 \frac{\partial \psi}{\partial p^3} \right) , \tag{1.13}$$
$$(J_3\psi)(p) = -i\hbar \left(p^1 \frac{\partial \psi}{\partial p^2} - p^2 \frac{\partial \psi}{\partial p^1} \right) . \tag{1.14}$$

For spacetime translations,

$$a = (a^0, a^1, a^2, a^3), \qquad U(a) = 1 + i(P^0 a^0 - i P^i a^i)/\hbar + \cdots ,$$

the generators are simply

$$(P^\mu)\psi(p) = p^\mu \psi(p). \tag{1.15}$$

Note that the generators J_i, K_i have physical dimensions of angular momentum (or action) and P^μ that of momentum. Note also the relativistic invariant definition of P^μ. The physical, measured energy is cP^0 and not $cP_0 = -cP^0$.

The Lie algebra, or commutation relations, of the generators of the group are the following. As the commutator is antisymmetric, of the 10 generators P^μ, J_i, K_i there are $10 \times 9/2 = 45$ such commutation relations. We write below a list of commutation relations of each type, the other such relations being obtained by renaming of indices. The total number of relations of a given type is written in a square bracket after one typical relation.

$$[P^\mu, P^\nu] = 0, \qquad [6] \tag{1.16}$$
$$[J_1, J_2] = i\hbar J_3, \qquad [3] \tag{1.17}$$
$$\left[J_i, P^0\right] = 0, \qquad [3] \tag{1.18}$$
$$\left[J_1, P^2\right] = i\hbar P^3, \qquad [9] \tag{1.19}$$
$$[J_1, K_2] = i\hbar K_3, \qquad [9] \tag{1.20}$$
$$[K_1, K_2] = -i\hbar J_3, \qquad [3] \tag{1.21}$$

$$\left[K_i, P^0\right] = i\hbar P^i, \qquad [3] \tag{1.22}$$

$$\left[K_i, P^j\right] = i\hbar\delta_i^j P^0. \qquad [9]. \tag{1.23}$$

Exercise 1.2 Prove that the generators J_i and K_i are self adjoint with respect to the inner product defined by (1.4). [The terms produced by differentiating $\omega_{\mathbf{p}}$ in checking for the Hermitian nature of J_i cancel.]

1.4 Position Operators of NRQM

We have written 'NRQM' for 'non-relativistic quantum mechanics'. Define

$$X_i = \frac{1}{mc} K_i, \tag{1.24}$$

then these 'position operators' have the following relations with each other and the momenta P^i.

$$[X_1, X_2] = -i\left(\frac{\hbar J_3}{m^2c^2}\right), \text{ and two more} \tag{1.25}$$

$$\left[X_i, P^j\right] = i\hbar\delta_i^j\left(\frac{P^0}{mc}\right), \tag{1.26}$$

$$\left[X_i, P^0\right] = \frac{i\hbar}{mc} P^i, \tag{1.27}$$

$$\left[P^i, P^j\right] = 0. \tag{1.28}$$

For wave packets of average angular momentum of the order of a few $\hbar$'s, the position operators for different directions fail to commute up to a square of the Compton wavelength ($(\hbar/mc)^2 \sim 10^{-24}\,\mathrm{m}^2$ for electron mass). Similarly, the commutator of position and momentum is the standard one only for wave packets of momenta well below mc so that $P^0/mc \sim 1$. The commutator with energy is the standard non-relativistic one for $P^0 \sim mc + P^i P^i/2mc$.

Under space translations the position operators shift as expected if the average energy is mostly the rest energy ($P^0 \sim mc$):

$$\exp(-i\,P^i a_i) X_j \exp(i\,P^k a_k) = X_j - a_j\left(\frac{P^0}{mc}\right).$$

The non-commuting of different X_j's has a physical interpretation. If we were to write an uncertainty relation for them in a state ψ, then it will look like

$$(\Delta_\psi X_1)(\Delta_\psi X_2) \geq \left(\frac{\hbar\langle J_3\rangle_\psi}{m^2c^2}\right).$$

Therefore, if we try to make a wave-packet too narrow in the 1-direction, then we cannot also squeeze it in the 2-direction too much for a given average value of the orbital angular momentum in the 3-direction. For particles with spin it is even more complicated.

1.5 Projective Representations

There is a deep connection between continuous symmetries and unitary operators in Hilbert space. They are connected by Wigner's theorem. Although we give a proof of the theorem in the last chapter of this book, it is well to recall the statement here.

The state of a physical system in quantum mechanics is described by a unit vector ϕ in a Hilbert space $\mathcal{H}$. But, any other vector which is a multiple of the given vector by a complex number of modulus unity is equally good to describe the same state. This is so because all physically measurable predictions of quantum mechanics depend on the expectation values of observables $\langle A\rangle = (\phi, A\phi)$ or equivalently, on transition probabilities $|(\phi, \psi)|^2$ between two states. These quantities have the same value if ϕ and ψ were to be replaced by $\exp(i\alpha)\phi$ and $\exp(i\beta)\psi$, respectively, where α and β are any real numbers. A number of type $\exp(i\alpha)$ is called a 'phase factor', or simply a 'factor' when the context is understood.

The set $\{\phi\}$ of all such vectors, that is, all multiples by a phase factor of some fixed unit vector ϕ is called a *ray*. A ray represents a physical state of the system in the sense that any member of the ray is equally qualified to represent the state. We say that $\{\phi\}$ is the ray determined by a unit vector ϕ. Of course, we can choose any other vector from the same ray and construct the ray by including all its multiples by phase factors.

A *symmetry transformation* S is a one-to-one invertible *ray mapping* such that the transition probability $|(\phi, \psi)|^2$, is the same as $|(\phi', \psi')|^2$ where ϕ' and ψ' are, respectively, any members of the mapped rays $S\{\phi\}$ and $S\{\psi\}$.

The symmetries obviously form a group, because if two mappings preserve transition probabilities separately, then their composition also does. The identity mapping and the inverse mapping are also symmetries.

The ray mappings are very inconvenient to deal with as rays do not even form a linear space. Therefore it is very useful to have a result like Wigner's theorem. It says that for the ray mapping corresponding to a symmetry, there is an underlying vector mapping, or operator, which is compatible with the ray mapping. This means that under the action of the operator, vectors in a ray are mapped to vectors in the mapped ray. This vector mapping, or operator, can be either unitary or anti-unitary and is determined uniquely except for an unknown phase factor.

Wigner's theorem applies to *one* symmetry. When we have a continuous group of symmetries, like the rotation, or the Poincare group, then we are dealing with an

infinite number of them, and for each symmetry transformation there is a unitary operator with an unknown phase factor.

For continuous group of symmetries there can only be unitary operators, no anti-unitary operators. This is so because the product of two unitary or two anti-unitary operators is always unitary, and every continuous symmetry transformation can always be written as a product of two similar symmetry transformations. For example a rotation by a certain angle about an axis is the same as a product of two rotations by half the angle about the same axis.

For implementing symmetry in quantum mechanics we not only need the operator for each symmetry but these operators should form a *representation* of the group of symmetries.

Let the symmetry group be $\mathcal{G}$ and its elements be denoted by letters r, s, t etc. If two symmetry group elements r and s correspond to U_r and U_s respectively, then the symmetry $rs \in \mathcal{G}$ corresponds to $U_r U_s$. But as each unitary operator is known only up to a factor, we can write

$$U_r U_s = \omega(r, s) U_{rs} \ , \qquad |\omega(r, s)| = 1. \tag{1.29}$$

The question arises: can we not fix these factors or "unknown phases" of each unitary operator in such a way so as to avoid the ambiguity altogether? We can start with the identity e of the group, and since $ee = e$, fix the phase of U_e such that $\omega(e, e) = 1$. Moreover, if there are three group elements then using the associative law

$$\begin{aligned} U_r U_s U_t &= U_r(\omega(s, t) U_{st}) = \omega(s, t)\omega(r, st) U_{rst} \\ &= \omega(r, s) U_{rs} U_t = \omega(r, s)\omega(rs, t) U_{rst}, \end{aligned}$$

we obtain

$$\omega(r, st)\omega(s, t) = \omega(r, s)\omega(rs, t). \tag{1.30}$$

If $f(r)$ are any phase factors defined on $\mathcal{G}$, with $f(e) = 1$, and if we change the unitary operators U_r to $U'_r = f(r) U_r$, the above relation will determine another set of ω's

$$\omega'(r, s) = \omega(r, s) \frac{f(r) f(s)}{f(rs)}.$$

These $\omega'(r, s)$ also satisfy the consistency condition (1.30). Two sets of such factors $\omega(r, s)$ and $\omega'(r, s)$ are called *equivalent*.

A representation in which the operators carry phase factors which cannot be chosen to make $\omega(r, s) = 1$ for all group elements is called a *projective representation* or a ray representation.

Exercise 1.3 Verify that the condition (1.30) is satisfied by $\omega'(r, s)$.

A unitary representation where we can choose all the factors $\omega(r, s)$ equal to 1, is called a *true representation*. It is a matter of satisfaction that for Poincare group the ambiguity in the unknown phases in $U(a, \Lambda)$ can be brought down to ± 1 using continuity arguments, and can be further eliminated completely if instead of $\mathcal{L}_+^\uparrow$ we use its *covering group* $SL(2, C)$ which is simply connected. This is just as well, because spin half-odd-integer particles are described by $SL(2, C)$ representations and not by those of $\mathcal{L}_+^\uparrow$. The representation we have defined for a relativistic spin zero particle above is a true unitary representation, without involving any unknown phases. But its non-relativistic limit, a representation of the Galilean group, turns out to be a *projective representation*!

1.6 Non-relativistic Limit

We discuss the non-relativistic limit of the relativistic representation. If $|\mathbf{p}| << mc$, the inner product can be re-written as

$$\int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_{\mathbf{p}}}\psi^*(p)\phi(p) \approx \frac{1}{2mc}\int \mathrm{d}^3\mathbf{p}\ \psi^*(\mathbf{p})\phi(\mathbf{p}).$$

We can redefine the inner-product and factor out the constant $1/2mc$ and simply write the inner product as

$$(\psi, \phi) = \int \mathrm{d}^3\mathbf{p}\ \psi^*(\mathbf{p})\phi(\mathbf{p}).$$

This will affect the self-adjoint property of the generators and we will come to that later. Moreover, as

$$p^0 \approx mc + \frac{|\mathbf{p}|^2}{2mc},$$

we define

$$E_G\psi(\mathbf{p}) = \frac{|\mathbf{p}|^2}{2m}\psi(\mathbf{p}).$$

The momentum generators remain the same

$$P_G^i\psi(\mathbf{p}) = p^i\psi(\mathbf{p}),$$

and we define

$$K_i^G = \frac{1}{c}K_i \approx im\hbar\frac{\partial}{\partial p^i}.$$

The other generators J_i retain their form.

The Lie algebra of Poincare group, *in the representation for a particle of mass m*, is reduced to that of the representation of the Galilean group whose typical commutators are:

$$\left[P_G^i, P_G^j\right] = 0, \tag{1.31}$$
$$\left[P_G^i, E_G\right] = 0, \tag{1.32}$$
$$\left[J_1^G, J_2^G\right] = i\hbar J_3^G, \tag{1.33}$$
$$\left[J_i^G, E_G\right] = 0, \tag{1.34}$$
$$\left[J_1^G, P_G^2\right] = i\hbar P_G^3, \tag{1.35}$$
$$\left[J_1^G, K_2^G\right] = i\hbar K_3^G, \tag{1.36}$$
$$\left[K_i^G, K_j^G\right] = 0, \tag{1.37}$$
$$\left[K_i^G, P_G^j\right] = im\hbar\delta_i^j. \tag{1.38}$$
$$\left[K_i^G, E_G\right] = i\hbar P_G^i. \tag{1.39}$$

This Lie algebra representation is called a *contraction* of the Poincare Lie algebra in the limit of $c \to \infty$. It depends on mass m just as the Poincare representation did. But whereas the relativistic representation was a true representation, this one gives a *projective representation* as we see next.

1.7 Phase Factors of the Galilean Group

The Galilean group, like the Poincare group, consists of space and time translations, rotations and Galilean boosts. If $(t, \mathbf{x})$ and $(t', \mathbf{x}')$ are the spacetime coordinates in two frames, then a group element $(\tau, \mathbf{a}, \mathbf{v}, R)$ where R is a rotation, $\mathbf{v}$ the relative velocity, $\mathbf{a}$ the space translation and τ the time translation acts as

$$t' = t + \tau, \tag{1.40}$$
$$\mathbf{x}' = R\mathbf{x} + \mathbf{v}t + \mathbf{a}. \tag{1.41}$$

The group multiplication can be obtained by writing two successive transformations:

$$(\tau_1, \mathbf{a}_1, \mathbf{v}_1, R_1)(\tau_2, \mathbf{a}_2, \mathbf{v}_2, R_2) = (\tau_3, \mathbf{a}_3, \mathbf{v}_3, R_3) \tag{1.42}$$

with

$$\begin{aligned}
\tau_3 &= \tau_1 + \tau_2, \\
\mathbf{a}_3 &= \mathbf{a}_1 + R_1\mathbf{a}_2 + \mathbf{v}_1\tau_2, \\
\mathbf{v}_3 &= \mathbf{v}_1 + R_1\mathbf{v}_2, \\
R_3 &= R_1R_2.
\end{aligned} \tag{1.43}$$

For later use, note the inverse of the element

$$(\tau, \mathbf{a}, \mathbf{v}, R)^{-1} = (-\tau, -R^{-1}\mathbf{a} + R^{-1}\mathbf{v}\tau, -R^{-1}\mathbf{v}, R^{-1}). \tag{1.44}$$

In analogy with the relativistic case, we define Galilean transformations on non-relativistic momentum eigenstates $|\mathbf{p}\rangle$, normalized as $\langle\mathbf{p}|\mathbf{p}'\rangle = \delta^3(\mathbf{p} - \mathbf{p}')$, as

$$U(t)|\mathbf{p}\rangle = \exp(i\mathbf{p}^2/2m\hbar)|\mathbf{p}\rangle, \tag{1.45}$$

$$U(\mathbf{a})|\mathbf{p}\rangle = \exp(-i\mathbf{p}\cdot\mathbf{a}/\hbar)|\mathbf{p}\rangle, \tag{1.46}$$

$$U(\mathbf{v})|\mathbf{p}\rangle = |\mathbf{p} + m\mathbf{v}\rangle, \tag{1.47}$$

$$U(R)|\mathbf{p}\rangle = |R\mathbf{p}\rangle. \tag{1.48}$$

Therefore,

$$\begin{aligned}
U(\tau, \mathbf{a}, \mathbf{v}, R)|\mathbf{p}\rangle &= \exp\left[i(\mathbf{p}')^2\tau/2m\hbar - i\mathbf{p}'\cdot\mathbf{a}/\hbar\right]|\mathbf{p}'\rangle, \\
\mathbf{p}' &= R\mathbf{p} + m\mathbf{v}.
\end{aligned} \tag{1.49}$$

In order to check whether this provides a representation of the Galilean group, we apply the two transformations $U_2 = (\tau_2, \mathbf{a}_2, \mathbf{v}_2, R_2)$ and $U_1 = (\tau_1, \mathbf{a}_1, \mathbf{v}_1, R_1)$ successively on $|\mathbf{p}\rangle$.

$$U_2|\mathbf{p}\rangle = \exp\left[\frac{i}{\hbar}\left(\frac{(\mathbf{p}')^2}{2m}\,\tau_2 - \mathbf{p}'\cdot\mathbf{a}_2\right)\right]|\mathbf{p}'\rangle, \qquad \mathbf{p}' = R_2\mathbf{p} + m\mathbf{v}_2,$$

and

$$\begin{aligned}
U_1U_2|\mathbf{p}\rangle = \exp&\left[\frac{i}{\hbar}\left(\frac{(\mathbf{p}')^2}{2m}\,\tau_2 - \mathbf{p}'\cdot\mathbf{a}_2\right)\right]\times \\
&\exp\left[\frac{i}{\hbar}\left(\frac{(\mathbf{p}'')^2}{2m}\,\tau_1 - \mathbf{p}''\cdot\mathbf{a}_1\right)\right]|\mathbf{p}''\rangle, \\
\mathbf{p}'' = R_1\mathbf{p}' + m\mathbf{v}_1 &= R_1R_2\mathbf{p} + m(R_1\mathbf{v}_2 + \mathbf{v}_1) = R_3\mathbf{p} + m\mathbf{v}_3.
\end{aligned}$$

To put it into standard form for U_3, we notice that

$$\mathbf{p}' = R_2\mathbf{p} + m\mathbf{v}_2 = R_1^{-1}(R_1R_2\mathbf{p} + mR_1\mathbf{v}_2) = R_1^{-1}(\mathbf{p}'' - m\mathbf{v}_1).$$

Since R_1^{-1} is a rotation,

$$(\mathbf{p}')^2 = (\mathbf{p}'' - m\mathbf{v}_1)^2 = (\mathbf{p}'')^2 + m^2\mathbf{v}_1^2 - 2m\mathbf{p}'' \cdot \mathbf{v}_1.$$

Similarly,

$$\mathbf{p}' \cdot \mathbf{a}_2 = R_1^{-1}(\mathbf{p}'' - m\mathbf{v}_1) \cdot \mathbf{a}_2 = (\mathbf{p}'' - m\mathbf{v}_1) \cdot R_1\mathbf{a}_2.$$

Substituting these expressions for $\mathbf{p}'$ and collecting terms, we obtain, using the group multiplication relations (1.43)

$$U_1U_2|\mathbf{p}\rangle = \exp\left[\frac{i}{\hbar}\left(\frac{m\mathbf{v}_1^2}{2}\,\tau_2 + m\mathbf{v}_1 \cdot R_1\mathbf{a}_2\right)\right] \times \tag{1.50}$$

$$\exp\left[\frac{i}{\hbar}\left(\frac{(\mathbf{p}'')^2}{2m}\,\tau_3 - \mathbf{p}'' \cdot \mathbf{a}_3\right)\right]|\mathbf{p}''\rangle, \tag{1.51}$$

$$= \exp\left[\frac{i}{\hbar}\left(\frac{m\mathbf{v}_1^2}{2}\,\tau_2 + m\mathbf{v}_1 \cdot R_1\mathbf{a}_2\right)\right]U_3|\mathbf{p}\rangle. \tag{1.52}$$

This clearly shows that our representation is not a true representation, but only up to the phase factor

$$\omega(1,2) = \exp[im\mathbf{v}_1^2\tau_2/2\hbar + im\mathbf{v}_1 \cdot R_1\mathbf{a}_2/\hbar]. \tag{1.53}$$

Exercise 1.4 Verify that the Galilean representation phase (1.53) satisfies the associativity condition (1.30):

$$\omega(1,23)\omega(2,3) = \omega(1,2)\omega(12,3).$$

What is surprising about this phase factor is that according to the group law (1.42) of the Galilean group the boosts and space translations commute

$$(0,\mathbf{a},0,\mathbf{1})(0,0,\mathbf{v},1) = (0,\mathbf{a},\mathbf{v},\mathbf{1}) = (0,0,\mathbf{v},1)(0,\mathbf{a},0,\mathbf{1}),$$

but in our representation, they do not:

$$U(\mathbf{a})U(\mathbf{v}) \neq U(\mathbf{v})U(\mathbf{a}).$$

It can be shown that these factors ω cannot be chosen to become equal to one by any choice of changing to $\omega'(1,2) = \omega(1,2)f_1f_2/f_{12}$.

However there do exist true representations of the Galilean group. But they are not physically relevant. See the note in Sect. 1.9.3 below.

1.8 Problems

Problem 1.1 Consider relativistic spinless particle of mass m whose basis states in momentum space $|p\rangle$ are defined in Sect. 1.2. Construct states

$$|x\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_{\mathbf{p}}} e^{-ip\cdot x/\hbar} |p\rangle$$

for any spacetime point $x = (x^0, x^1, x^2, x^3)$. Show that

1. $U(a, \Lambda)|x\rangle = |\Lambda x + a\rangle$,
2. Calculate the 'equal time' product: $\langle x^0, \mathbf{x}|x^0, \mathbf{y}\rangle$. Can these states $|x^0, \mathbf{x}\rangle$ be eigenstates of some self-adjoint 'position operator'?
3. Show that the set of states $\{|x^0, \mathbf{x}\rangle\}$ for constant x^0 form a non-orthogonal but complete set. (Hint: express $|p\rangle$ in terms of these states.)
4. If $|\phi\rangle$ and $|\psi\rangle$ are two vectors in the Hilbert space and $\phi(x) = \langle x|\phi\rangle$ and $\psi(x) = \langle x|\psi\rangle$ then

$$\left[\frac{\partial^2}{\partial x^{0\,2}} - \frac{\partial^2}{\partial \mathbf{x}^2} + m^2\right]\phi(x) = 0, \quad \text{and}$$

$$\langle\phi|\psi\rangle = i \int_{x^0=\text{const.}} \mathrm{d}^3\mathbf{x} \left[\phi^*(x)\frac{\partial\psi(x)}{\partial x^0} - \frac{\partial\phi^*(x)}{\partial x^0}\psi(x)\right].$$

Problem 1.2 Express Poincare group elements (a, Λ) as 5×5 matrices like

$$\begin{pmatrix} \Lambda_{00} & \Lambda_{01} & \Lambda_{02} & \Lambda_{03} & a^0 \\ \Lambda_{10} & \Lambda_{11} & \Lambda_{12} & \Lambda_{13} & a^1 \\ \Lambda_{20} & \Lambda_{21} & \Lambda_{22} & \Lambda_{23} & a^2 \\ \Lambda_{30} & \Lambda_{31} & \Lambda_{32} & \Lambda_{33} & a^3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

acting on column vectors $(x^0, x^1, x^2, x^3, 1)$. Work out the Poincare group Lie algebra from infinitesimal forms (near identity) of these matrices.

Problem 1.3 In order to take the limit of Poincare group to Galilean group as $c \to \infty$, the factor c in $x^0 = ct$ should be separated and transformations expressed as acting on $(t, x^1, x^2, x^3, 1)$ and not on $(x^0, x^1, x^2, x^3, 1)$. Modify the 5×5 Poincare transformation matrices above accordingly showing c-dependence. Obtain the Galilean Lie algebra when $c \to \infty$. That will be the Lie algebra of the Galilean group, with $[K_i^G, P_G^j] = 0$, and not of the representation.

Hint: Only the boost matrices contain c, and a^0 will be replaced by τ without the factor of c to give $t \to t + \tau$.

1.9 Notes and References

1.9.1 Time Translation and Time Evolution

The Poincare group (and the Galilean group) includes space and *time* translations. One can ask the question: if time translations are already present, what is the need for equations of motion? As far as group elements are concerned, there is a unitary operator for both a group element and its inverse. But time evolution has a meaning only for forward progress of time. How are time translations related to the Schrödinger picture or the Heisenberg picture?

The symmetry transformations of the Poincare group represent relations between two inertial frames of reference. We can think of time evolution as a continuous change of frame. A frame S'_τ lies in the *future* of a frame S by an amount $\tau > 0$ if the coordinate t' of S'_τ is related to the time coordinate of frame S by $t' = t - \tau$. Therefore a state ψ seen by S at time $t = 0$, say, will be seen in the continuously changing frames S'_τ as $\psi(\tau) = U(-\tau)\psi$. This is the time evolution in the Schrödinger picture.

The Heisenberg's way of representing time evolution is through unchanging states but changing observables. So, if we change the frames S'_τ continuously by changing τ, the state ψ of the system can be chosen to remain constant but the observables A evolve with τ so that their expectation values, (which are experimentally measured), change in the same manner as in the Schrödinger picture:

$$(\psi, A(\tau)\psi) = (\psi(-\tau), A\psi(-\tau)).$$

Since this happens for every state ψ, we get $A(\tau) = U(\tau)AU(-\tau)$. If $U(\tau) = \exp(iH\tau/\hbar)$ then the differential form of this relation is

$$i\hbar\frac{dA(\tau)}{d\tau} = [A(\tau), H].$$

In the next chapter we will generalize this idea of continuously changing frames.

Irreducible unitary representations of Poincare group are treated in the classic paper of Wigner [3].

1.9.2 The Newton–Wigner Position Operator

One can ask the question, for a relativistic particle, why cannot we just define a three vector $\hat{X}_i = i\hbar\partial/\partial p^i$ as position operators in one frame reference? But these are unsuitable as observables because they are not self-adjoint. In fact, by the definition of an adjoint,

$$(\phi, \hat{X}_i^\dagger \psi) = (\hat{X}_i \phi, \psi) = -i\hbar \int \left(\frac{d^3\mathbf{p}}{2p^0}\right) \frac{\partial \phi^*}{\partial p^i} \psi(p)$$

After an integration by part, we see that

$$\hat{X}_i^\dagger = \hat{X}_i - \frac{i\hbar p^i}{(p^0)^2}.$$

Therefore, we can define a self-adjoint operator, called the Newton–Wigner operator as

$$X^i_{NW} = \frac{\hat{X}_i + \hat{X}_i^\dagger}{2} = i\hbar \frac{\partial}{\partial p^i} - \frac{i\hbar p^i}{2(p^0)^2}.$$

This has the advantage that different components of the position operator commute, and can be simultaneously diagonalized. But a common eigenstate of X^i_{NW} (that is, a state "localized" at a given point), when space translated by $U(\mathbf{a})$ is not orthogonal to the original state.

One can read more about it in the original paper [4], or Sect. 3c of the classic book by Schweber [5].

1.9.3 Poincare and Galilean Group Representations

We have discussed only the Poincare group representation corresponding to a particle of mass m and spin zero. General irreducible unitary representations of the Poincare group are characterized or labelled by two quantities: the value of $w = P^\mu P_\mu$ and a number s which can be an integer or a half-odd integer or a real number corresponding to spin. If w is negative, equal to $-m^2$ we get the representation for a particle of finite mass m. If $w = 0$, we get zero mass representations. In this case s can be an integer or half-odd integer and is called helicity. In $w = 0$, case, s can also be an arbitrary real number as well, but this 'continuous spin' quantum number is clearly unphysical. Also unphysical are those representations for which $w > 0$, or, as we say, P^μ is space-like.

The Galilean group representation which is derived from the representation of the Poincare group for finite mass m shows the dependence on m in its Lie algebra commutator between boost and translation. The Galilean group, of course, requires them to commute. The true representations of the Galilean group do not seem to correspond to anything physical.

A very detailed introduction to Lie groups, Lie algebras, and particularly the Galilean group can be found in the classic book by Sudarshan and Mukunda [6]. Inonu and Wigner [7] found that the true representations of Galilean group are physically not relevant, whereas of the physically relevant representations, there is only one type, and that carries phase factors which cannot be wished away. This paper also

introduces the idea of *group contraction*: how one group G_1 reduces to another group G_2 when a parameter on which the group elements of G_1 depend, goes to a limiting value. In the present case of Poincare to Galilean group, it is the velocity of light $c \to \infty$. Galilean Lie algebra is touched upon briefly in the Sect. 2.4 of the book by Weinberg [8] from a physicist's point of view.

A rigorous treatment of the question of true and ray representations is in Bargmann [9] which discusses the Galilean group in Sect. 6 of the paper.

References

1. A.J. Kalnay, The localization problem, in *Problems in the Foundations of Physics*. Studies in the Foundations, Methodology and Philosophy of Science, vol. 4, ed. by M. Bunge (Springer, Berlin, 1971)
2. P. Aguilar, C. Chryssomalakos, H. Hernandez Coronado, E. Okon, Position operators and center of mass: new perspectives, arXiv:1306.0504v1
3. E.P. Wigner, Ann. Math. **40**, 149 (1939). Reprinted in F.J. Dyson, *Symmetry Groups in Nuclear and Particle Physics* (W. A. Benjamin, New York, 1966)
4. T.D. Newton, E.P. Wigner, Localized states for elementary systems. Rev. Mod. Phys. **21**, 400 (1949)
5. S.S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1961)
6. E.C.G. Sudarshan, N. Mukunda, *Classical Dynamics, A Modern Perspective* (World Scientific, Singapore, 2016)
7. E. Inonu, E.P. Wigner, Il Nuovo Cimento **IX**(8), 705 (1952)
8. S. Weinberg, *The Quantum Theory of Fields*, vol. I (Cambridge, 1995)
9. V. Bargmann, Ann. Math. **59**, 1 (1954)

Chapter 2
A Bundle Picture of Quantum Mechanics

The vector that represents the state of a quantum mechanical system depends, apart from the system itself, on the frame of reference in which the description is being made. There are an infinite number of frames and there is a vector in a Hilbert space for each of them. This suggests a vector bundle picture for describing the dynamics of the system. This way of looking at dynamics is especially useful where change of frames is involved.

Note: In this chapter we use Einstein's convention of summing over repeated indices to simplify formulas.

2.1 The Bundle Picture

The concept of state of a physical system in classical or quantum mechanics always depends on the frame of reference in which the description is being made. Usually, the frame of reference is fixed and therefore not explicitly mentioned.

A quantum state is described by a vector in a Hilbert space $\mathcal{H}$ characteristic of the physical system. There is thus a Hilbert space for each frame of reference to which the state as seen in that frame belongs. All these spaces, called *fibers*, are of the same type because they refer to the same physical system. But quantum mechanical description has this special feature: a state ψ and the set of observables $\{A_i\}$, taken together, describe the same physics as do states $U\psi$ and observables $\{UA_iU^{-1}\}$, where U is any unitary operator.

So we have the following 'bundle' picture. There is a set X of frames of reference, and for each frame P in X, there is a fiber of a Hilbert space $\mathcal{H}_P$ which contains the vectors as observed by P. Let us assume that the frames are labelled by a number of parameters or 'coordinates' $x = (x^1, \ldots, x^n)$ etc. Then we can ask the question: how is the state $\psi \in \mathcal{H}_P$ of a physical system as seen by the observer in frame P

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_2

(with coordinates x) related to the state as seen by the observer in a neighboring frame Q (with coordinates $x + \Delta x$) considering Δx to be small?

In general, it is not possible to compare vectors in two different vector spaces $\mathcal{H}_P$ and $\mathcal{H}_Q$ because the states and observables in each frame can be changed by a unitary operator.

We need two rules: one for identifying the different individual Hilbert spaces $\mathcal{H}_P$, $\mathcal{H}_Q$ etc., and second, a way of comparing vectors in neighboring spaces. The first of these rules is called 'local trivialization' and the second a law of 'parallel transport'.

A trivialization allows us to use the cartesian product $X \times \mathcal{H}$ for describing the vectors of all frames in a common Hilbert space, i.e., as elements $(x.\psi(x))$, where x represents the frame, and $\psi(x) \in \mathcal{H}$, as the state seen by that frame. Two different trivializations will differ by an x-dependent unitary operator $U(x)$ changing states (and observables): $(x, \psi(x))$ being replaced by $(x, U(x)\psi(x))$.

We use the term 'trivialization' here is a limited sense where coordinates of the base manifold X are not changed, but only the identification of fibers with $\mathcal{H}$ changes. The 'bundle picture' presented here has the same mathematical features as in general relativity or Yang–Mills gauge theories. Of course, the specific base manifold and the fiber in all these cases are different.

2.1.1 Sections

A *section* is a smooth function which assigns a vector in $\mathcal{H}_P$ for all frames $P \in X$. If we have chosen a trivialization then $(x, \psi(x)) \in X \times \mathcal{H}$ is a section. We say that a section is smooth if $\psi(x)$ changes continuously and smoothly with x. A section is the mapping $x \to \psi(x)$ after we have taken some fixed trivialization.

An orthonormal (o.n.) basis of sections $x \to \phi_n(x)$, $n = 1, 2, \ldots$ is such that for each x

$$(\phi_n(x), \phi_m(x)) = \delta_{nm}.$$

Given an o.n. basis of sections every section $\psi(x)$ can be be replaced by its components functions, that is, a sequence of complex functions $c_n(x) = (\phi_n(x), \psi(x))$.

2.1.2 Parallel Transport

Let P and Q be two neighboring points in X with coordinates x and $x + \Delta x$ respectively. We treat Δx^i as infinitesimally small.

Given a vector in $\mathcal{H}_x$ at a point P with coordinates x we can compare it with a vector at Q (with coordinates $x + \Delta x$) if we provide a rule by which these vectors

can be brought together. If $\psi \in \mathcal{H}_{x+\Delta x}$ at Q is brought to P *without change* we expect the transported vector $\psi^{\parallel}_{Q\to P}$

1. to depend linearly under transport: that is, if we bring a linear combination $a\phi + b\psi$ from Q to P then the vectors brought at P will be the same linear combination $a\phi^{\parallel}_{Q\to P} + b\psi^{\parallel}_{Q\to P}$,
2. to remain unchanged if $\Delta x^i \to 0$,
3. to preserve inner product that is $(\phi^{\parallel}_{Q\to P}, \psi^{\parallel}_{Q\to P}) = (\phi, \psi)$.

To find the form of such a rule, it is convenient to use an o.n. basis of sections $\phi_n(x)$. From the first point above if we can define the parallel transport for a basis, then we can define for any vector, using the linearity.

With this in mind, we write the rule for $\phi_r(x + \Delta x)$ at Q brought to the point P with coordinates x as (we drop the suffix $Q \to P$ in what follows)

$$\phi^{\parallel}_r = \phi_r(x) + \Delta x^i \phi_s(x)\Gamma_{isr}(x), \tag{2.1}$$

where $\Gamma_{isr}(x)$ are complex numbers expressing the transported vector in the basis $\phi_r(x)$. To check if it preserves the inner product, we calculate (keeping to first order of smallness in Δx),

$$\delta_{rs} = (\phi^{\parallel}_r, \phi^{\parallel}_s) = \delta_{rs} + \Delta x^i \left[\delta_{rt}\Gamma_{its}(x) + \Gamma^*_{itr}(x)\delta_{ts}\right],$$

or,

$$\Gamma_{irs}(x) + \Gamma^*_{isr}(x) = 0,$$

which shows that Γ_{isr} is an *anti-Hermitian* matrix of an operator on $\mathcal{H}$ in the basis $\{\phi_r\}$. We can define a Hermitian operator $\hat{P}_i(x)$ as

$$\Gamma_{irs}(x) = i(\phi_r(x), \hat{P}_i(x)\phi_s(x)), \tag{2.2}$$

and call it the *connection operator* with connection components Γ_{isr} in this basis.

2.1.3 Change of Basis or of Trivialization

A change in trivialization results in a unitary operator $\hat{U}(x)$ acting on each of the spaces $\mathcal{H}_x$. The basis of sections $\phi_r(x)$ now becomes $\tilde{\phi}_r = \hat{U}(x)\phi_r(x) = \phi_s(x)U_{sr}$ where $U_{sr} = (\phi_s(x), \hat{U}\phi_r(x))$. Under parallel transport,

$$\tilde{\phi}^{\parallel}_r = \tilde{\phi}_r(x) + \Delta x^i \tilde{\phi}_s(x)\tilde{\Gamma}_{isr}(x),$$

where $\tilde{\Gamma}$ are the new connection components. On the other hand, when $\tilde{\phi}_r(x + \Delta x) = \phi_s(x + \Delta x)U_{sr}(x + \Delta x)$ is brought to x, the coefficients $U_{sr}(x + \Delta x)$, being

complex numbers, are brought as such, and contribute an additional term: $\Delta x^i \partial_i U_{sr}$. Thus, omitting the arguments x of functions, and keeping to first order in Δx,

$$\begin{aligned}\overset{\sim}{\phi}{}^{\parallel}_r &= \phi_s U_{sr} + \Delta x^i \phi_s \partial_i U_{sr} + \Delta x^i \phi_t \Gamma_{its} U_{sr} \\ &= \overset{\sim}{\phi}_r + \Delta x^i \overset{\sim}{\phi}_t U^{-1}_{ts} \partial_i U_{sr} + \Delta x^i \overset{\sim}{\phi}_u U^{-1}_{ut} \Gamma_{its} U_{sr} \\ &= \overset{\sim}{\phi}_r + \Delta x^i [\overset{\sim}{\phi}_t U^{-1}_{ts} \partial_i U_{sr} + \overset{\sim}{\phi}_u U^{-1}_{ut} \Gamma_{its} U_{sr}].\end{aligned}$$

Comparison shows, written as a matrix,

$$\overset{\sim}{\Gamma}_i = U^{-1} \Gamma_i U + U^{-1} \partial_i U. \tag{2.3}$$

If the connection components are all zero, $\Gamma_{ius} = 0$, then in some other trivialization, the connection components are "pure gauge":

$$\overset{\sim}{\Gamma}_{itr} = U^{-1}_{ts} \partial_i U_{sr},$$

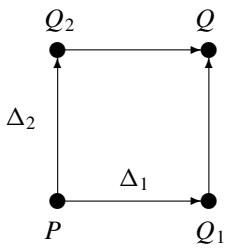
with $U_{sr}(x)$ a unitary matrix dependent on x.

2.1.4 Transport Round a Loop and Curvature

We have discussed the form of a vector $\phi_r(Q)$ parallel transported from a point $Q = \{x + \Delta x\}$ to $P = \{x\}$. One can ask a question, if a vector is parallel transported around a closed loop and brought back to the starting point, will it agree with the initial vector, or not?

We discuss first for an infinitesimally small loop. And, instead of going round the loop we can equivalently take the vector along two different infinitesimal paths, having the same starting and ending points, and compare the transported vectors. Let us take four points with coordinates as shown:

$$Q = \{x^i + \Delta_1 x^i + \Delta_2 x^i\}, \quad Q_1 = \{x^i + \Delta_1 x^i\}, \quad Q_2 = \{x^i + \Delta_2 x^i\}, \quad P = \{x^i\}.$$



The vector $\phi_r(Q)$ when brought to Q_1 is

$$\phi^{\parallel}_{Q\to Q_1 r} = \phi_r(x+\Delta_1 x) + \Delta_2 x^i \phi_s(x+\Delta_1 x)\Gamma_{isr}(x+\Delta_1 x).$$

Going through another step and transporting this to P, and keeping to second order in Δx's,

$$\begin{aligned}\phi^{\parallel}_{Q\to Q_1\to Pr} &= \phi_r(x) + \Delta_1 x^i \phi_s(x)\Gamma_{isr}(x)\\ &\quad + \Delta_2 x^i[\phi_s(x) + \Delta_1 x^j \phi_t(x)\Gamma_{jts}(x)][\Gamma_{isr}(x) + \Delta_1 x^j \partial_j \Gamma_{isr}]\\ &= \phi_r(x) + \Delta_1 x^i \phi_s(x)\Gamma_{isr}(x) + \Delta_2 x^i \phi_s(x)\Gamma_{isr}(x)\\ &\quad + \Delta_2 x^i \Delta_1 x^j \phi_t(x)\big[\partial_j \Gamma_{itr} + \Gamma_{jts}(x)\Gamma_{isr}(x)\big].\end{aligned}$$

The other route for transport $Q \to Q_2 \to P$ will give similarly, with the role of Δ_1 and Δ_2 interchanged:

$$\begin{aligned}\phi^{\parallel}_{Q\to Q_2\to Pr} &= \phi_r(x) + \Delta_2 x^i \phi_s(x)\Gamma_{isr}(x) + \Delta_1 x^i \phi_s(x)\Gamma_{isr}(x)\\ &\quad + \Delta_1 x^i \Delta_2 x^j \phi_t(x)\big[\partial_j \Gamma_{itr} + \Gamma_{jts}(x)\Gamma_{isr}(x)\big].\end{aligned}$$

The difference is therefore,

$$\phi^{\parallel}_{Q\to Q_1\to Pr} - \phi^{\parallel}_{Q\to Q_2\to Pr} = \Delta_2 x^i \Delta_1 x^j \phi_t(x) R_{trij}$$

where

$$R_{trij} = \partial_j \Gamma_{itr}(x) - \partial_i \Gamma_{jtr}(x) + \Gamma_{jts}(x)\Gamma_{isr}(x) - \Gamma_{its}(x)\Gamma_{jsr}(x) \tag{2.4}$$

is called the *curvature tensor*. It is well to remember that indices t, r belong to the o.n. basis in Hilbert space $\mathcal{H}$, whereas i, j run over the variables of the base X. If the components of the curvature tensor are zero then parallel transport of a vector can be taken along any path without dependance on the path.

Exercise 2.1 Show that the change in the curvature tensor components, under a change in the basis sections by $\phi_r(x) \to \widetilde{\phi}_r = \phi_s(x)U_{sr}$ is, as a matrix,

$$\widetilde{R}_{ji} = U^{-1} R_{ji} U. \tag{2.5}$$

2.2 Covariant Derivative

2.2.1 Covariant Derivative of a Section

The parallel transport law gives us a way to define derivative of sections in any direction. Given a section $x \to \psi(x) = c_r(x)\phi_r(x)$ we can bring $\psi(x + \Delta x)$ at Q to P using (2.1), and find the difference from $\psi(x)$,

$$\begin{aligned}\psi^{\parallel}_{Q\to P} - \psi(x) &= c_r(x + \Delta x)[\phi_r(x) + \Delta x^i \Gamma_{isr}(x)] - c_r(x)\phi_r(x)\\ &= \Delta x^i\, \phi_s(x)\, [\partial_i c_s(x) + \Gamma_{ist} c_t(x)].\end{aligned}$$

Dividing by the displacement and taking the limit:

$$D_i\psi(x) = \phi_s(x)[\partial_i c_s(x) + \Gamma_{ist} c_t(x)]. \tag{2.6}$$

If we were to change the trivialization then $\widetilde{\phi}_r\ (x) = \phi_s(x)U_{sr}$ and express the vector ψ in both the bases

$$\psi = c_s\phi_s = \widetilde{c}_r\widetilde{\phi}_r = \phi_s U_{sr}\ \widetilde{c}_r,$$

then, suppressing the indices, and using matrix notation,

$$c = U\ \widetilde{c}, \qquad \widetilde{c} = U^{-1}c.$$

Using (2.3) and (2.6) we write the covariant derivative in the new trivialization

$$\begin{aligned}\widetilde{D}_i\ \psi(x) &= \widetilde{\phi}\ [\partial_i\ \widetilde{c} + \widetilde{\Gamma}_i \widetilde{c}],\\ &= \phi\, U\, [\partial_i(U^{-1}c) + (U^{-1}\Gamma_i U + U^{-1}\partial_i U)U^{-1}c],\\ &= \phi\, [U(\partial_i U^{-1})c + \partial_i c + \Gamma_i c + (\partial_i U)U^{-1}c].\end{aligned}$$

The first and last term in the last line can be added $U(\partial_i U^{-1}) + (\partial_i U)U^{-1} = \partial_i(UU^{-1}) = 0$, Therefore $\widetilde{D}_i\ \psi(x) = D_i\psi(x)$ and the covariant derivative is independent of the trivialization chosen to calculate it. This justifies the name.

2.2.2 Covariant Derivative of an Operator

Let $\hat{A}$ be an operator defined in $\mathcal{H}_Q$, $Q = \{x + \Delta x\}$ and let

$$A_{rs} = (\phi_r(x + \Delta x), \hat{A}\phi_s(x + \Delta x))$$

be its matrix elements. To define its parallel transport to $P = \{x\}$ we use the fact that transport is a linear process which preserves the inner product. Therefore $A^{\parallel}_{Q\to P}$ has the same matrix elements in the transported basis. Keeping to first order,

$$\begin{aligned} A_{sr} &= (\phi^{\parallel}_{s(Q\to P)}, A^{\parallel}_{Q\to P}\phi^{\parallel}_{r(Q\to P)}) \\ &= (\phi_s(x), A^{\parallel}_{Q\to P}\phi_r(x)) + \Delta x^i \Gamma^*_{its}(x)(\phi_t(x), A^{\parallel}_{Q\to P}\phi_r(x)) \\ &\quad + \Delta x^i \Gamma_{itr}(x)(\phi_s(x), A^{\parallel}_{Q\to P}\phi_t(x)). \end{aligned}$$

We realize that $A_{sr}(x + \Delta x)$ and $A^{\parallel}_{sr}(x) \equiv (\phi_s(x), A^{\parallel}_{Q\to P}\phi_r(x))$ differ by an order of magnitude of Δx and become equal as $\Delta x \to 0$. Therefore, using the anti-Hermitian nature of Γ,

$$\begin{aligned} A^{\parallel}_{sr}(x) &= A_{sr} - \Delta x^i \Gamma^*_{its}(x) A_{tr} - \Delta x^i \Gamma_{itr}(x) A_{st} \\ &= A_{sr} + \Delta x^i (\Gamma_{ist}(x) A_{tr} - A_{st}\Gamma_{itr}(x)). \end{aligned}$$

Thus the numbers A_{sr} (which are the matrix elements of $\hat{A}$ at $x + \Delta x$) define, by parallel transport a new operator $A^{\parallel}$ through its matrix elements with respect to the same basis of sections.

If we had a *section* of operators $x \to \hat{A}(x)$ then A_{sr} in the equation above would be $A_{sr}(x + \Delta x)$, and the difference of transported operator's matrix elements with those of $A_{sr}(x)$ will be

$$A^{\parallel}_{sr}(x) - A_{sr}(x) = A_{sr}(x + \Delta x) - A_{sr}(x) + \Delta x^i (\Gamma_{ist}(x) A_{tr} - A_{st}\Gamma_{itr}(x)).$$

This allows us to define the covariant derivative of an operator matrix elements, (writing Γ_i as a matrix)

$$D_i A_{sr}(x) = \frac{\partial A_{sr}}{\partial x^i} + [\Gamma_i, A]_{sr}. \tag{2.7}$$

2.3 The Base X of Galilean Frames

So far we have focussed our attention on the fibers of the Hilbert space bundle. Now we look at the structure of the base X of frames of reference. Our starting point was that frames of reference are related to each other through symmetry operations such as translation in space and time, or rotations and boosts. These transformations are not accidental, but reflect the invariance of the metric or the line element of spacetime. The infinitesimal transformations, determined by these symmetries or 'isometries' of spacetime, show up as Killing vector fields. It can be shown that in an n-dimensional Riemannian space there can be only $n(n + 1)/2$ such independent Killing vector fields.

Starting from a standard frame we can label all other frames by the symmetry transformation parameters which connect them to the standard frame. Thus there are as many frames as the group elements of the symmetry transformations, namely, the ten parameter group of Galilean or Poincare transformations.

We have seen in Chap. 1 that all the Galilean frames can be obtained by applying rotation, boost, and translations in space and time, to some arbitrarily chosen frame S_0:

$$S_x = (\tau, \mathbf{a}, \mathbf{v}, R)S_0, \qquad x = (\tau, \mathbf{a}, \mathbf{v}, R)$$

which corresponds to the transformation

$$\begin{aligned} t' &= t + \tau, \\ \mathbf{x}' &= R\mathbf{x} + \mathbf{v}t + \mathbf{a}\,. \end{aligned}$$

We parametrize the rotation R through the Euler angles (rotation by angle ϕ about 3-axis, a rotation by θ about 1-axis, followed by a rotation of ψ about the 3-axis)

$$R(\psi, \theta, \psi) = R_3(\psi)R_1(\theta)R_3(\phi). \tag{2.8}$$

There are thus ten parameters in $x = (\tau, \mathbf{a}, \mathbf{v}, R)$.

2.3.1 The Hypothesis of Parallel Transport

We have seen in Chap. 1 that unitary representations of the Galilean or the Poincare group reflect the equivalence of all frames of reference in the sense that physical measurements will not distinguish one frame from another. In the language of differential geometry this means that state vectors in various frames related by the unitary operators of the group element are carried from one frame to another 'without change', or by parallel transport. Therefore, if $x \to \psi(x) = c_r(x)\phi_r(x)$ is the section determined by the physical system's state vector as seen in various frames, the equation for determining the change from frame at x to $x + \Delta x$ is

$$D_i\psi = \phi_s(x)\left(\frac{\partial c_s}{\partial x^i} + \Gamma_{isr}(x)c_r(x)\right) = 0. \tag{2.9}$$

For a trivialization which uses *a constant o.n. basis of sections*, so that there is no dependence on x from the basis, we can write directly the Eq. (2.9) with operators $\hat{\Gamma}_i$

$$\Gamma_{isr}(x) = i(\phi_s, \hat{\Gamma}_i\phi_r) \tag{2.10}$$

as the Schrödinger equation

$$i\frac{\partial\psi(x)}{\partial x^i} = \hat{\Gamma}_i(x)\psi(x). \tag{2.11}$$

2.3.2 Calculation of $\hat{\mathbf{\Gamma}}_i$

The connection operators Γ_i can be calculated by putting all the Δx^j, $j \neq i$ zero and looking at the coefficient of Δx^i. The state Ψ as seen in the frame $S_x, x = (\tau, \mathbf{a}, \mathbf{v}, R)$ is obtained by the successive unitary operators in reverse order on state Ψ_0 in S_0:

$$\Psi_x = U(R^{-1})U(-\mathbf{v})U(-\mathbf{a})U(-\tau)\Psi_0.$$

All we need to do is to differentiate this with each of the parameters x^i and obtain the corresponding $\hat{\Gamma}_i$ using the Schrödinger equation (2.11).

The unitary operators are

$$\begin{aligned}
U(-\tau) &= \exp(-i\hat{\mathbf{P}}^2\tau/2m\hbar),\\
U(-\mathbf{a}) &= \exp(i\hat{\mathbf{P}}\cdot\mathbf{a}/\hbar),\\
U(-\mathbf{v}) &= \exp(im\hat{\mathbf{X}}\cdot\mathbf{v}/\hbar),\\
U(R^{-1}) &= \exp(i\phi J_3/\hbar)\exp(i\theta J_2/\hbar)\exp(i\psi J_3/\hbar).
\end{aligned}$$

The ten connection operators at different points $x = (\tau, \mathbf{a}, \mathbf{v}, \psi, \theta, \phi)$ are determined by the parallel transport hypothesis are:

$$\hat{\Gamma}_\tau(x) = \frac{1}{\hbar}\frac{(R^{-1}\hat{\mathbf{P}} - m\mathbf{v})^2}{2m} = \frac{1}{\hbar}\frac{(\hat{\mathbf{P}} - mR\mathbf{v})^2}{2m}, \tag{2.12}$$

$$\hat{\Gamma}_{a_i}(x) = -\frac{1}{\hbar}(R^{-1}\hat{\mathbf{P}} - m\mathbf{v})_i, \tag{2.13}$$

$$\hat{\Gamma}_{v_i}(x) = -\frac{1}{\hbar}m(R^{-1}\hat{\mathbf{X}})_i, \tag{2.14}$$

$$\hat{\Gamma}_\psi(x) = -\frac{1}{\hbar}(J_3\cos\theta + J_2\sin\theta\cos\phi + J_1\sin\theta\sin\phi), \tag{2.15}$$

$$\hat{\Gamma}_\theta(x) = -\frac{1}{\hbar}(J_1\cos\phi - J_2\sin\phi), \tag{2.16}$$

$$\hat{\Gamma}_\phi(x) = -\frac{1}{\hbar}J_3\,. \tag{2.17}$$

We have chosen a particular order of operators to label frames of reference $x = (\tau, \mathbf{a}, \mathbf{v}, R)$. The same frames can be labelled in other ways too, leading to connection operators different from the ones above. But since that would amount to a change by

unitary operators, the two connection operator sets will be related by the connection transformation law given in Sect. 2.1.3 above.

2.3.3 Galilean Bundle Curvature is Zero

It is worthwhile to check that all the 45 components of curvature

$$R_{ji} = \frac{\partial}{\partial x^j}\hat{\Gamma}_i - \frac{\partial}{\partial x^i}\hat{\Gamma}_j + i[\hat{\Gamma}_j, \hat{\Gamma}_i], \qquad \{x^i\} = (\tau, \mathbf{a}, \mathbf{v}, \phi, \theta, \psi),$$

for these 10 connection operators are zero. It should not cause surprise that the curvature components are zero because they were obtained by the process $dU = (dUU^{-1})U$ which is a 'pure gauge' term of the transformation formula for the connection.

But it is still worthwhile to remark that the space translation determined by $\hat{\Gamma}_{a_i}$ and boost $\hat{\Gamma}_{v_i}$ give zero curvature even if they do not commute and give rise to the non-removable phase factor for the representation as discussed in Sect. 1.7. A simpler example is the non-commutativity of the angular momentum operators. The curvature obtained is zero because in going round a loop, the connection operators in different frames are different, and exactly compensate for the changes due to non-commutativity.

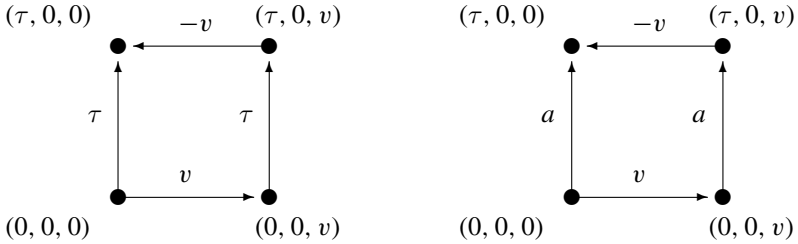
For the sake of simplicity we take two cases restricted to one dimension to illustrate the point. There are only three variables (τ, a, v) with the connection operators

$$\begin{aligned}
\Gamma_\tau(\tau, a, v) &= \frac{1}{\hbar}\frac{(\hat{P} - mv)^2}{2m} \\
\Gamma_a(\tau, a, v) &= -\frac{1}{\hbar}(\hat{P} - mv) \\
\Gamma_v(\tau, a, v) &= -m\frac{1}{\hbar}\hat{X}.
\end{aligned}$$

In the first case a frame at $(0, 0, 0)$ is changed to a moving frame $(0, 0, v)$ then time translated by $\Gamma_\tau(0, 0, v)$ and then brought back to $(\tau, 0, 0)$ by $\Gamma_v(\tau, 0, v)$ by reverse velocity $-v$. During this process the state ψ_0 in frame $(0, 0, 0)$ is seen to change to

$$e^{-imv\hat{X}/\hbar}e^{-i\tau(\hat{P}-mv)^2/2m\hbar}e^{imv\hat{X}/\hbar}\psi_0 = e^{-i\tau\hat{P}^2/2m\hbar}\psi_0$$

This is the same as the direct change from $(0, 0, 0)$ to $(\tau, 0, 0)$ by $\Gamma_\tau(0, 0, 0)$.



Similarly if we choose space translation in place of time translation we get the same result. We do not get the irremovable phase factor in the representation of the Galilean group because the connection Γ_a at $(0, 0, 0)$ is $-\hat{P}$ and at $(0, 0, v)$ is $-(\hat{P} - mv)$:

$$e^{-imv\hat{X}/\hbar}e^{i\tau(\hat{P}-mv)a/\hbar}e^{imv\hat{X}/\hbar}\psi_0 = e^{i\tau\hat{P}a/\hbar}\psi_0.$$

2.4 Application to Accelerated Frames

In the previous sections we have developed a geometric picture. The dynamics of a system can be reduced to changing frames with respect to time translations as has already been discussed in Sect. 1.9.1.

If all observers are equivalent, a physical system can be described by state and observables in any of the frames. However, dynamics requires *evolution curve* which connects different frames, just as, in the simplest of cases, time translated frames determine the dynamics.

2.5 Non-inertial Frames

2.5.1 A Linearly Accelerated Frame

Choose a standard frame S_0, with a state of the system represented by $|\Psi_0\rangle$ and observables $\hat{\mathbf{X}}$, $\hat{\mathbf{P}}$ etc.

Let S be a frame whose origin lies a time τ in the future of the origin of S_0. Then the vector

$$|\Psi\rangle = \exp(-i\hat{\mathbf{P}}^2\tau/2m\hbar)|\Psi_0\rangle$$

represents the same physical state of the system in frame S. The average value of position $\langle\hat{\mathbf{X}}\rangle \equiv \langle\Psi|\hat{\mathbf{X}}|\Psi\rangle$ is related to $\langle\hat{\mathbf{X}}\rangle_0 \equiv \langle\Psi_0|\hat{\mathbf{X}}|\Psi_0\rangle$ by

$$\langle \hat{\mathbf{X}} \rangle = \langle \hat{\mathbf{X}} \rangle_0 + \frac{\langle \hat{\mathbf{P}} \rangle_0}{m} \tau.$$

Similarly, if S is displaced by a distance $\mathbf{a}$ with respect to S_0, then

$$|\Psi\rangle = \exp(i\hat{\mathbf{P}} \cdot \mathbf{a}/\hbar)|\Psi_0\rangle$$

is a state with the property that

$$\langle \hat{\mathbf{X}} \rangle = \langle \hat{\mathbf{X}} \rangle_0 - \mathbf{a}.$$

Therefore a wave packet located at some distance $\mathbf{x}$ in S_0 is seen at $\mathbf{x}' = \mathbf{x} - \mathbf{a}$ in S. Moreover,

$$|\Psi\rangle = \exp(-im\hat{\mathbf{X}} \cdot \mathbf{v}/\hbar)|\Psi_0\rangle$$

represents a wave packet with no change in location, but the average momentum changes by

$$\langle \hat{\mathbf{P}} \rangle = \langle \hat{\mathbf{P}} \rangle_0 - m\mathbf{v}.$$

A frame S_τ initially coincident with S_0 and moving with constant acceleration $\mathbf{g}$ sees the state Ψ_0 as

$$|\Psi\rangle_\tau = \exp(-im\hat{\mathbf{X}} \cdot \mathbf{v}/\hbar) \exp(i\hat{\mathbf{P}} \cdot \mathbf{a}/\hbar) \exp(-i\hat{\mathbf{P}}^2\tau/2m\hbar)|\Psi_0\rangle, \qquad (2.18)$$
$$\mathbf{a} = \mathbf{g}\tau^2/2, \qquad \mathbf{v} = \mathbf{g}\tau,$$

which determines the average values

$$\langle \hat{\mathbf{X}} \rangle_\tau = \langle \hat{\mathbf{X}} \rangle_0 + \frac{\langle \hat{\mathbf{P}} \rangle_0}{m} \tau - \frac{1}{2}\mathbf{g}\tau^2,$$
$$\langle \hat{\mathbf{P}} \rangle_\tau = \langle \hat{\mathbf{P}} \rangle_0 - m\mathbf{g}\tau.$$

Differentiating the first of these equations by τ twice we get the Newtonian equations for the inertial force, (or equivalence principle, so to speak!)

$$\frac{d}{d\tau} \langle \hat{\mathbf{X}} \rangle_\tau = \frac{1}{m} \langle \hat{\mathbf{P}} \rangle_\tau \qquad (2.19)$$
$$\frac{d^2}{d\tau^2} \langle \hat{\mathbf{X}} \rangle_\tau = -\mathbf{g}. \qquad (2.20)$$

This is what we expect from Ehrenfest's theorem. It is instructive to find the Hamiltonian in the accelerating frame S_τ. By a simple calculation

$$i\hbar\frac{d}{d\tau}|\Psi\rangle_\tau = \left[mg\hat{\mathbf{X}} - (\hat{\mathbf{P}} + m\mathbf{g}\tau)\cdot\mathbf{g}\tau + \frac{1}{2m}(\hat{\mathbf{P}} + m\mathbf{g}\tau)^2\right]|\Psi\rangle_\tau$$
$$= \left[\frac{1}{2m}\hat{\mathbf{P}}^2 + m\mathbf{g}\cdot\hat{\mathbf{X}} + \frac{1}{2}m\mathbf{g}^2\tau^2\right]|\Psi\rangle_\tau, \tag{2.21}$$

which shows the presence of the 'gravitational' potential as well as the extra c-number term which generates the phase factor

$$\exp\left(-i(m/2\hbar)\int_0^\tau \mathbf{g}^2\tau^2 d\tau\right) = \exp(-im\mathbf{g}^2\tau^3/6\hbar),$$

identified by Eliezer and Leach [1] for the Schrödinger equation of a linearly accelerated particle.

2.5.2 Rotating Frame: Centrifugal and Coriolis Forces

Let the frame S be chosen with the same origin as S_0 rotating about the 3-axis with constant angular velocity ω. The rotation R as function of the parameter τ is

$$x'^1 = x^1\cos\omega\tau + x^2\sin\omega\tau,$$
$$x'^2 = x^2\cos\omega\tau - x^1\sin\omega\tau.$$

For the free particle the frame

$$S = (\tau)(R)S_0$$

which requires the state Ψ_0 seen as

$$\Psi(\tau) = U(R)^{-1}U(-\tau)\Psi_0,$$

where

$$U(R)^{-1} = \exp(i\omega\tau\hat{J}_3/\hbar), \qquad \hat{J}_3 = \hat{X}_1\hat{P}_2 - \hat{X}_2\hat{P}_1,$$
$$U(-\tau) = \exp(-i\hat{\mathbf{P}}^2\tau/2m\hbar).$$

The evolution takes place as

$$i\hbar\frac{d\Psi}{d\tau} = \left(\frac{\hat{\mathbf{P}}^2}{2m} - \omega\hat{J}_3\right)\Psi(\tau)$$

One can wonder, how is this related to the centrifugal and Coriolis forces? A classical particle has position and velocity simultaneously, and therefore the centrifugal force

(which depends on position) and the Coriolis force (which depends on velocity) both determine the trajectory. For a quantum description we can only see the evolution of a wave packet through Ehrenfest's theorem. Thus, as first derivatives of average position we find

$$\begin{aligned}\frac{d}{d\tau}\langle\hat{X}_1\rangle &= \frac{d}{d\tau}\langle\Psi(\tau)|\hat{X}_1|\Psi(\tau)\rangle \\ &= \frac{1}{i\hbar}\langle\Psi(\tau)|\left[\hat{X}_1, \frac{\hat{\mathbf{P}}^2}{2m} - \omega\hat{J}_3\right]|\Psi(\tau)\rangle \\ &= \frac{1}{m}\langle\hat{P}_1\rangle + \omega\langle\hat{X}_2\rangle. \end{aligned} \tag{2.22}$$

Similarly,

$$\frac{d}{d\tau}\langle\hat{X}_2\rangle = \frac{1}{m}\langle\hat{P}_2\rangle - \omega\langle\hat{X}_1\rangle, \tag{2.23}$$

$$\frac{d}{d\tau}\langle\hat{X}_3\rangle = \frac{1}{m}\langle\hat{P}_3\rangle. \tag{2.24}$$

Using these the second derivatives can be calculated

$$\begin{aligned} m\frac{d^2}{d\tau^2}\langle\hat{X}_1\rangle &= 2\omega\langle\hat{P}_2\rangle - m\omega^2\langle\hat{X}_1\rangle, \\ m\frac{d^2}{d\tau^2}\langle\hat{X}_2\rangle &= -2\omega\langle\hat{P}_1\rangle - m\omega^2\langle\hat{X}_2\rangle, \\ m\frac{d^2}{d\tau^2}\langle\hat{X}_3\rangle &= 0. \end{aligned}$$

Substituting for $\langle\hat{P}_2\rangle$ and $\langle\hat{P}_1\rangle$ from (2.23) and (2.22) respectively, the Coriolis and centrifugal forces become manifest:

$$m\frac{d^2}{d\tau^2}\langle\hat{X}_1\rangle = 2\omega\frac{d}{d\tau}\langle\hat{X}_2\rangle + m\omega^2\langle\hat{X}_1\rangle, \tag{2.25}$$

$$m\frac{d^2}{d\tau^2}\langle\hat{X}_2\rangle = -2\omega\frac{d}{d\tau}\langle\hat{X}_1\rangle + m\omega^2\langle\hat{X}_2\rangle, \tag{2.26}$$

$$m\frac{d^2}{d\tau^2}\langle\hat{X}_3\rangle = 0. \tag{2.27}$$

2.6 Problems

Problem 2.1 Let $\phi(x,t)$ be a solution of the Schrödinger equation for a (non-relativistic) free particle:

$$i\hbar\frac{\partial\phi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\phi}{\partial x^2}.$$

Show that in an accelerated frame $t' = t$, $x' = x - gt^2/2$, the wave function

$$\psi(x', t') = \phi'(x', t')\exp(iS(x', t')/\hbar), \qquad \phi'(x', t') = \phi(x = x' + gt'^2/2, t = t'),$$

satisfies the Schrödinger equation in the accelerated frame with gravitational potential mgx':

$$i\hbar\frac{\partial\psi}{\partial t'} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x', t')}{\partial x'^2} + mgx'\psi(x', t'),$$

provided that S is chosen as

$$S(x', t') = -mgt'x' - \frac{1}{6}mg^2t'^3.$$

This is the argument used in references in [1].

Problem 2.2 Consider a frame (x', y', t) rotating (with angular velocity ω) in a circle of radius R about the origin of an inertial frame (x, y, t) such that the origin of the rotating frame is always on the circle and its x' axis always in the radial direction.

Find the Hamiltonian for a free particle of mass m in the rotating frame.

2.7 Notes and References

2.7.1 Bundle Picture

The equation $idU = HdtU$ or $idUU^{-1} = Hdt$ occurs so often in quantum mechanics that it is surprising that the identification of Hdt as a connection 1-form has been late in coming, although the ideas of differential geometry were being applied to quantum field theory for quite some time. One reason may be that discussions in quantum mechanics are predominantly about time evolution and one dimensional base space is not interesting enough. The relevance of ideas of differential geometry came with discovery of the geometric phase where the Hamiltonian depends on several changing parameters.

Despite that, as far as we know, the first explicit mention of time evolution as parallel transport occurs in Asorey et al. [2]. The idea was developed by Graudenz [3], Iliev [4] as a mathematical formalism. It was Chingangbam [5] who applied the bundle picture to actual problems like the quantum mechanics of accelerated frames. Experimental confirmation of quantum effects of Einstein's equivalence principle in non-relativistic as well as relativistic domain is too vast a subject to be included

here. The reader can follow the recent activity beginning from the paper by Zych and Brukner [6] and the chapter by Mashhoon [7].

2.7.2 Geometric Phase

Geometric phase also uses the language of a vector bundle, connection and curvature, where the base is not that of frames of reference, but of parameters on which the Hamiltonian of the system depends. The literature on geometric phase, or the Berry phase, is vast. It was indeed the 'phase that launched a thousand scripts'! The collection of original papers can be found in Shapere and Wilczek [8]. The book by Bohm et al. [9] is a good introduction. The review by Mukunda and Simon [10] is devoted to applications of the geometric phase to group representations.

References

1. C.J. Eliezer, P.G. Leach, Am. J. Phys. **45**, 1218 (1977), See also M. Nauenberg, Am. J. Phys. **84**, 879 (2016)
2. M. Asorey, J.F. Carinena, M. Paramio, J. Math. Phys. **23**, 1451 (1982)
3. D. Graudenz, arXiv:gr-qc/9412013
4. B.Z. Iliev, Int. J. Mod. Phys. A **17**, 245 (2002)
5. P. Chingangbam, Connection and curvature in the fibre bundle formulation of quantum theory. Ph.D. thesis, Jamia Millia Islamia (2002) (Unpublished), P. Chingangbam, P. Sharan, Phys. Rev. A **64**, 042107 (2001)
6. M. Zych, C. Brukner, Nat. Phys. **14**, 1027 (2018)
7. B. Mashhoon, Lect. Notes Phys. **702**, 112 (2006)
8. A. Shapere, F. Wilczek (eds.), *Geometric Phases in Physics* (World Scientific, Singapore, 2006)
9. A. Bohm, A. Mostafazadeh, H. Koizumi, Q. Niu, J. Zwanziger, *The Geometric Phase in Quantum Systems* (Springer, Berlin, 2003)
10. N. Mukunda, R. Simon, Ann. Phys. **228**, 205 (1993)

Chapter 3
A Beam of Particles = A Plane Wave?

How can a plane wave represent a beam of particles? And what exactly is 'probability per unit time'? These questions bother every student of quantum mechanics when they encounter them in scattering theory. This chapter takes the mystery out of these issues, apart from offering a very concise introduction to formal scattering theory.

3.1 A Coherent Bunch

In this section we prove a result which is quite general, but which has a special importance in scattering theory.

A quantum system with translational degrees of freedom has a complete basis of eigenstates of total momentum labeled as

$$|\mathbf{p}, \beta\rangle, \qquad \langle \mathbf{p}'\beta'|\mathbf{p}\beta\rangle = 2\omega_{\mathbf{p}}\delta^3(\mathbf{p}' - \mathbf{p})\delta_{\beta'\beta}, \quad \omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2c^2},$$

where β's are collective name of eigenvalues of other observables needed for a complete set of basis states. And, although the result applies equally well to non-relativistic case, (apart from a normalization factor), we use the relativistic normalization to be specific. The eigenvalues β may be discrete, or continuous, but we use the notation for discrete values for simplicity. We also refer to the system as a 'particle'.

Let $|\phi\beta_0\rangle$ be a state of the system with respect to an observer. We assume the state to have total momentum close to some fixed value $\mathbf{k}$:

$$\langle \mathbf{p}\beta|\phi\beta_0\rangle = f(\mathbf{p})\delta_{\beta\beta_0}$$

with $f(\mathbf{p})$ being sharply peaked around $\mathbf{p} = \mathbf{k}$. Normalization of $|\phi\beta_0\rangle$ requires

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_3

$$\langle \phi\beta_0 | \phi\beta_0 \rangle = \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_\mathbf{p}} |f(\mathbf{p})|^2 = 1.$$

Consider the same system as seen by another observer located with respect to the first observer at a distance $-\mathbf{r}$. If there is translational symmetry, this state is represented by a unitary translation operator $U(-\mathbf{r})$ acting on $|\phi\beta_0\rangle$:

$$U(-\mathbf{r})|\phi\beta_0\rangle = \exp(i\mathbf{r}\cdot\hat{\mathbf{P}}/\hbar)|\phi\beta_0\rangle,$$

where $\hat{\mathbf{P}}$ is the operator for total momentum. This state is the same as would be seen by the original observer had the system *as a whole* been displaced by a distance $\mathbf{r}$.

Note that we do not assume that the system with state $|\phi\beta_0\rangle$ is localized at any point in space. In fact, as we know from non-relativistic quantum mechanics, where position operators make sense, that a state with linear momentum wave function $f(\mathbf{p})$ more or less sharply defined, will have a very large wave packet in spatial dimensions.

Now, imagine a very large number, N, of copies of the system all prepared in the same identical state $|\phi\beta_0\rangle$ and then, spatially translated from the place of their preparation to positions $\mathbf{r}_i$, $i = 1, \ldots, N$, the points being distributed with a uniform, constant density ρ in space in a very large volume. These systems are then in states $|\phi_i\rangle$ whose momentum wave-functions are given by

$$\langle \mathbf{p}\beta|\phi_i\rangle = f(\mathbf{p})\delta_{\beta\beta_0}\exp(i\mathbf{r}_i.\mathbf{p}/\hbar). \tag{3.1}$$

We call such a collection of particles a *coherent bunch* of particles.

Suppose we want to make measurements on an observable B on all these states. The average of the expectation values of B in these states will be

$$\frac{1}{N}\sum_i \langle\phi_i|B|\phi_i\rangle = \frac{1}{N}\sum_i \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_\mathbf{p}} \int \frac{\mathrm{d}^3\mathbf{p}'}{2\omega_{\mathbf{p}'}} \exp[i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}_i/\hbar] \times f(\mathbf{p}')^* f(\mathbf{p}) \langle \mathbf{p}'\beta_0|B|\mathbf{p}\beta_0\rangle.$$

If the points $\mathbf{r}_i$ are closely spaced, and the volume very large, we can replace the sum over points $\mathbf{r}_i$ by $\rho\int \mathrm{d}^3\mathbf{r}$ because ρ is constant:

$$\sum_i [\ldots] = \int \rho \mathrm{d}^3\mathbf{r}[\ldots] = \rho\int \mathrm{d}^3\mathbf{r}[\ldots]. \tag{3.2}$$

Carrying out the integration over $\mathbf{r}$ we get a factor $(2\pi\hbar)^3\delta^3(\mathbf{p}-\mathbf{p}')$ which can be integrated out as well:

$$\frac{1}{N}\sum_i \langle\phi_i|B|\phi_i\rangle = \langle B\rangle = \frac{\rho(2\pi\hbar)^3}{N}\int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_\mathbf{p}} \frac{|f(\mathbf{p})|^2}{2\omega_\mathbf{p}} \langle \mathbf{p}\beta_0|B|\mathbf{p}\beta_0\rangle.$$

If the observable B is such that $\langle \mathbf{p}\beta_0|B|\mathbf{p}\beta_0\rangle$ varies smoothly near $\mathbf{p} = \mathbf{k}$, then the matrix element can be pulled out at value $\mathbf{k}$ because that is where the main contribution to the integral comes from, and the expectation value, to a good approximation, is

$$\begin{aligned}\langle B\rangle &= \frac{\rho(2\pi\hbar)^3}{N}\frac{1}{2\omega_{\mathbf{k}}}\langle \mathbf{k}\beta_0|B|\mathbf{k}\beta_0\rangle \int \frac{\mathrm{d}^3\mathbf{p}}{2\omega_{\mathbf{p}}}|f(\mathbf{p})|^2 \\ &= \frac{\rho(2\pi\hbar)^3}{N}\frac{1}{2\omega_{\mathbf{k}}}\langle \mathbf{k}\beta_0|B|\mathbf{k}\beta_0\rangle, \qquad \text{(Relativistic)} \end{aligned} \tag{3.3}$$

because the last factor is equal to one. Dramatically, *the dependence on details of the momentum profile of the state* $|\phi\beta_0\rangle$ *disappears completely!* All that mattered was that the momentum was sharply peaked around $\mathbf{k}$. We call the formula (3.3) the (relativistic) "bunch formula". In fact, if the bunch had been made from a mixture of several states, all of them with different momentum profiles (but the same β_0), and with a sharp peak at the same value of $\mathbf{k}$, we would have obtained the same result.

We get the bunch formula for the non-relativistic case if we take the non-relativistic normalization for the momentum states:

$$\langle B\rangle = \frac{\rho(2\pi\hbar)^3}{N}\langle \mathbf{k}\beta_0|B|\mathbf{k}\beta_0\rangle. \qquad \text{(Non-relativistic)} \tag{3.4}$$

There is another way to look at this result. If these N particles were inside a large volume L^3 then $\rho = N/L^3$. Let $\mathrm{d}^3\mathbf{k}$ be a small volume element in the momentum space. The above result (for the non-relativistic case, for example) can then be written as

$$\langle B\rangle = \frac{\langle \mathbf{k}\beta_0|B|\mathbf{k}\beta_0\rangle \mathrm{d}^3\mathbf{k}}{n},$$

where n is the number of 'cells' of size h^3 in the occupied phase space,

$$n = \frac{L^3\mathrm{d}^3\mathbf{k}}{h^3}, \qquad h = 2\pi\hbar = \text{Planck's constant}\,.$$

3.2 Scattering Theory

3.2.1 Scattering States and the Moller Operator

We work with Hamiltonian H of the system, and also with the free Hamiltonian H_0 when the interaction is 'switched off'. Both the operators are defined in a common Hilbert space $\mathcal{H}$. We use the Schrödinger picture where all time dependence is carried by the states.

Let $\Psi(t_0)$ be the state of the system at some fixed time, say, $t = t_0$. Then the state at time t is

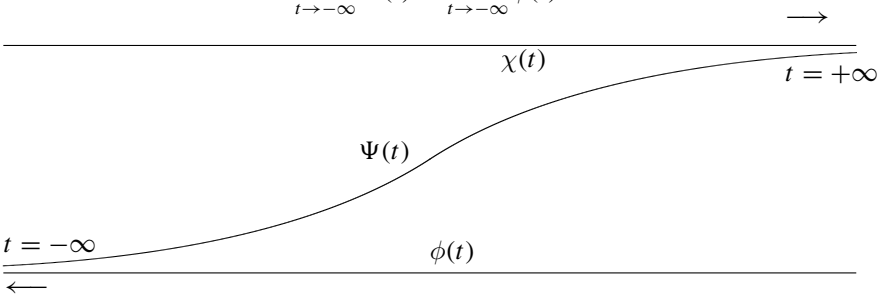
$$\Psi(t) = U(t - t_0)\Psi(t_0), \qquad U(t - t_0) = \exp(-iH(t - t_0)/\hbar).$$

A vector $\phi(t_0) \in \mathcal{H}$ would represent the state of the system when free from interaction at $t = t_0$ if, at any other time t, it is given by

$$\phi(t) = U_0(t - t_0)\phi(t_0), \qquad U_0(t - t_0) = \exp(-iH_0(t - t_0)/\hbar).$$

Now suppose that $\Psi(t)$ is such that it becomes *indistinguishable* from some free state $\phi(t)$ for large negative t, then we can say that the system *behaves* like a free system in remote past, that is, if

$$\lim_{t\to-\infty} \Psi(t) = \lim_{t\to-\infty} \phi(t).$$



A *scattering state* $\Psi(t_0)$ at time $t = t_0$ is a state such that its 'ancient history' $\Psi(t) = U(t - t_0)\Psi(t_0)$ for large negative times $t \to -\infty$ coincides with that of *some* free state $\phi(t) \equiv U_0(t - t_0)\phi(t_0)$:

$$\lim_{t\to-\infty} \|\Psi(t) - \phi(t)\| = 0. \tag{3.5}$$

Such a scattering state *looks like a free state in remote future too*. That is, there is a free state $\chi(t_0)$ such that

$$\lim_{t\to\infty} \|\Psi(t) - \chi(t)\| = 0, \tag{3.6}$$

where $\chi(t) = U_0(t - t_0)\chi(t_0)$. The existence of the first limit is equivalent to

$$\begin{aligned}\lim_{t\to-\infty} \|\Psi(t) - \phi(t)\| &= \lim_{t\to-\infty} \|U(t - t_0)\Psi(t_0) - U_0(t - t_0)\phi(t_0)\| \\ &= \lim_{t\to-\infty} \|\Psi(t_0) - U^{-1}(t - t_0)U_0(t - t_0)\phi(t_0)\| \to 0,\end{aligned}$$

and defines the Moller wave operator $\Omega^{(+)}$:

$$\Omega^{(+)} \equiv \lim_{t\to-\infty} U(t-t_0)^{-1}U_0(t-t_0). \tag{3.7}$$

As a result, at $t = t_0$

$$\Psi(t_0) = \Omega^{(+)}\phi(t_0). \tag{3.8}$$

It is not necessary to start at time $t = t_0$. We could have started at any finite time t_1 and obtained

$$\Psi(t_1) = \Omega^{(+)}\phi(t_1). \tag{3.9}$$

Thus, $\Omega^{(+)}$ is actually independent of time. Similarly, for $t \to \infty$ we would get

$$\Omega^{(-)} \equiv \lim_{t\to+\infty} U(t-t_0)^{-1}U_0(t-t_0), \tag{3.10}$$

and, for any time t,

$$\Psi(t) = \Omega^{(-)}\chi(t). \tag{3.11}$$

Not every state can be a scattering state. If H admits bound states, then those certainly do not go to free states in remote past or future. There are confining Hamiltonians H which have only bound states, and no scattering states.

3.2.2 Scattering Matrix

A system starting out with a free state $\phi(t)$ in remote past becomes a free state $\chi(t)$ in remote future. The two free states, final and initial, are related (using ${\Omega^{(-)}}^\dagger\Omega^{(-)} = 1$, proved later), by

$$\chi(t) = {\Omega^{(-)}}^\dagger\Psi(t) = {\Omega^{(-)}}^\dagger\Omega^{(+)}\phi(t) \equiv S\phi(t). \tag{3.12}$$

The operator S is called the *scattering matrix* or S-matrix. As we shall see, $\Omega^{(+)}$ and S contain all information about scattering.

3.2.3 Properties of $\Omega^{(\pm)}$ and S

We now list a number of properties of the Moller operators $\Omega^{(\pm)}$, and the scattering matrix S with proofs detailed in a later section.

(1) $\Omega^{(\pm)}$ are *norm preserving* operators

$$\Omega^{(+)^\dagger}\Omega^{(+)} = 1, \qquad \Omega^{(-)^\dagger}\Omega^{(-)} = 1. \tag{3.13}$$

But $\Omega^{(+)}\Omega^{(+)^\dagger}$ (or $\Omega^{(-)}\Omega^{(-)^\dagger}$) may not be equal to 1. That means, $\Omega^{(\pm)}$ may not be unitary if the scattering states do not span the whole Hilbert space $\mathcal{H}$.

(2) $\Omega^{(\pm)}$ convert $U_0(t)$ into $U(t)$:

$$\Omega^{(\pm)}U_0(t) = U(t)\Omega^{(\pm)}, \tag{3.14}$$

$$\Omega^{(\pm)}H_0 = H\Omega^{(\pm)}, \tag{3.15}$$

$$\Omega^{(\pm)^\dagger}H = H_0\Omega^{(\pm)^\dagger}. \tag{3.16}$$

(3) S is unitary.

$$SS^\dagger = S^\dagger S = 1. \tag{3.17}$$

(4) S commutes with the *free* hamiltonian H_0, although it may not commute with the total Hamiltonian.

$$SH_0 = H_0S. \tag{3.18}$$

(5) There is an integral equation for the Moller operator $\Omega^{(+)}$. Denote $V \equiv H - H_0$, then

$$\Omega^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^{t_0} U_0(t-t_0)^{-1}V\Omega^{(+)}U_0(t-t_0)dt. \tag{3.19}$$

(6) Similarly, there is an integral equation for the S-matrix:

$$S = 1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} U_0(t)^{-1}V\Omega^{(+)}U_0(t)dt. \tag{3.20}$$

3.3 Transition Rate

Let ϕ and ξ be two free states at time $t = t_0$ orthogonal to each other, $(\phi, \xi) = 0$.

Let us begin with a large number N of *identically prepared* systems in the free state $\phi(-T_1) = U_0(-T_1)\phi$, in remote past ($T_1$ a large positive number). In scattering theory we are interested in a question of the following type: how many of these N systems will be found in a free state $\xi(+T_2) = U_0(+T_2)\xi$ in some remote future $t = +T_2$ (T_2 being a large positive number)?

If there was no interaction and evolution took place only with the free Hamiltonian H_0, then the answer would be zero, because for all values of time t the state $\phi(t)$ would be orthogonal to $\xi(t)$, both states evolving with H_0 so that $(\phi(T_2), \chi(T_2)) = 0$.

As it is, the state $\phi(-T_1)$ evolves to the scattering state $\Psi(t) = \Omega^{(+)}\phi(t)$ and may have a non-zero probability $P_{\xi\Psi}(t) = |(\xi(t), \Psi(t))|^2$ to be found in a free state $\xi(t)$ at time t. This probability is zero for large negative times, increases as interaction is "switched on" and gradually saturates to a constant value $P_{\xi\Psi} = |(\xi(T_2), \Psi(T_2))|^2 = |(\xi(T_2), \chi(T_2))|^2$ because for large positive times the state again evolves as a free state χ. Therefore the total number of transitions to state $\xi(T_2)$ at time T_2 are

$$N P_{\xi\Psi} = N|(\xi(T_2), \Psi(T_2))|^2.$$

If we wait for a time ΔT_2 more, the total number of transitions will be

$$N P_{\xi\Psi} = N|(\xi(T_2 + \Delta T_2), \Psi(T_2 + \Delta T_2))|^2.$$

Therefore the number of *transitions per unit time* or the *transition rate* is given by

$$n_\xi = N\frac{d}{dt}|(\xi(t), \Psi(t))|^2, \qquad (t \to \infty) \tag{3.21}$$

to be evaluated for large times.

The argument given above is based on the assumption that *once the system makes a transition to the free state $\xi(t)$ from $\Psi(t)$ it continues to evolve as a free state from then on, and does nor revert back again to a scattering state.*

Therefore, transitions to ξ occurring at different times keep accumulating and would be counted among the states which have already made transition to ξ.

At present we do not have a complete theory of measurement in quantum mechanics. These issues are not completely understood. But the formulas given below are based on these assumptions and they agree completely with experimental observations.

Let us put the transition rate in a more convenient form. For any complex function $A(t)$ of t such that $dA/dt = C/i$

$$\frac{d|A|^2}{dt} = A^*\frac{C}{i} - \frac{C^*}{i}A = 2\,\mathrm{Im}(A^*C).$$

Applying this simple identity to $A = \langle\xi(t)|\Psi(t)\rangle = \langle\xi(t)|\Omega^{(+)}|\phi(t)\rangle$ we obtain

$$\begin{aligned}\frac{dA}{dt} &= \frac{1}{i\hbar}\left[\left(i\hbar\frac{d}{dt}\langle\xi(t)|\right)|\Psi(t)\rangle + \langle\xi(t)|\left(i\hbar\frac{d}{dt}|\Psi(t)\rangle\right)\right] \\ &= \frac{1}{i\hbar}\left[\langle\xi(t)|(-H_0)|\Psi(t)\rangle + \langle\xi(t)|H|\Psi(t)\rangle\right] \\ &= \frac{1}{i\hbar}\langle\xi(t)|V|\Psi(t)\rangle, \qquad V = H - H_0\end{aligned}$$

$$= \frac{1}{i\hbar}\langle\xi(t)|V\Omega^{(+)}|\phi(t)\rangle$$

which identifies the number C of the identity as

$$C = \langle\xi(t)|V\Omega^{(+)}|\phi(t)\rangle/\hbar.$$

The transition rate at time t_0 is, therefore,

$$\begin{aligned} n_\xi &= N2\,\mathrm{Im}(A^*C) \\ &= \frac{2N}{\hbar}\mathrm{Im}\Big[\langle\xi(t_0)|\Omega^{(+)}|\phi(t_0)\rangle^*\langle\xi(t_0)|V\Omega^{(+)}|\phi(t_0)\rangle\Big], \\ &= \frac{2N}{\hbar}\mathrm{Im}\Big[\langle\phi(t_0)|\Omega^{(+)^\dagger}|\xi(t_0)\rangle\langle\xi(t_0)|V\Omega^{(+)}|\phi(t_0)\rangle\Big]. \end{aligned}$$

Thus,

$$n_\xi = \frac{2N}{\hbar}\mathrm{Im}\langle\phi(t_0)|B(t_0)|\phi(t_0)\rangle, \tag{3.22}$$

where

$$B(t_0) = \Omega^{(+)^\dagger}|\xi(t_0)\rangle\langle\xi(t_0)|V\Omega^{(+)}. \tag{3.23}$$

Note the appearance of the projection operator

$$P_{\xi(t_0)} = |\xi(t_0)\rangle\langle\xi(t_0)| \tag{3.24}$$

to the final states.

This is the formula for the number of transitions per unit time if all the N particles were in the same identical free state ϕ in remote past. When we apply this formula to an actual beam we have to replace the particles with the bunch average. That is

$$n_\xi = \frac{2N}{\hbar}\mathrm{Im}\left(\frac{1}{N}\sum_i\langle\phi_i(t_0)|B(t_0)|\phi_i(t_0)\rangle\right). \tag{3.25}$$

We can substitute the "bunch formula" inside the parentheses in the formula above depending on the situation of the specific case.

To evaluate the above average for large values of t we can choose the fixed 'origin' of time t_0 large enough to include the time of the duration of interaction.

In practice, both the free states $\phi(t)$ and $\xi(t)$ are stationary states with trivial time dependence $\exp(-iEt/\hbar)$. Therefore n_ξ is actually independent of time.

3.4 Evaluation of Transition Rate and Cross-Sections

3.4.1 Non-relativistic Scattering from a Potential

The initial free state ϕ (whose spatial translates constitute the bunch for a beam) is an eigenstate of the linear momentum $|\mathbf{k}\rangle$ with energy $E_{\mathbf{k}} = |\mathbf{k}|^2/2m$ where m is the mass of the particle. In potential scattering we also look for transitions to final states with definite energy as well.

With this in mind, let $\phi(t_0)$ represent a beam with sharp momentum around $\mathbf{k}$, density ρ, and energy $E_{\mathbf{k}}$. Let the final state have the sharp energy $E_{\mathbf{p}}$. Our bunch formula (3.4) will give the transitions rate

$$\begin{aligned} n_\xi &= \frac{2N}{\hbar}\mathrm{Im}\left(\frac{1}{N}\sum_i \langle\phi_i(t_0)|B(t_0)|\phi_i(t_0)\rangle\right) \\ &= \frac{2N}{\hbar}\frac{\rho(2\pi\hbar)^3}{N}\mathrm{Im}\,\langle\mathbf{k}\beta_0|B(t_0)|\mathbf{k}\beta_0\rangle \\ &= 16\pi^3\hbar^2\rho\,\mathrm{Im}\,\langle\mathbf{k}\beta_0|B(t_0)|\mathbf{k}\beta_0\rangle \\ &= 16\pi^3\hbar^2\rho\,\mathrm{Im}\,\langle\mathbf{k}\beta_0|{\Omega^{(+)}}^\dagger|\xi(t_0)\rangle\langle\xi(t_0)|V\Omega^{(+)}|\mathbf{k}\beta_0\rangle \\ &= 16\pi^3\hbar^2\rho\,\mathrm{Im}\left[\langle\xi(t_0)|\Omega^{(+)}|\mathbf{k}\beta_0\rangle^*\langle\xi(t_0)|V\Omega^{(+)}|\mathbf{k}\beta_0\rangle\right]. \end{aligned}$$

It should be noted that there is t_0 dependence of $\phi(t_0)$. Strictly speaking we should write $\langle\mathbf{k}\beta_0, t_0|$ and $|\mathbf{k}\beta_0, t_0\rangle$, but we omit it to simplify writing. It should be understood in the formulas above and below.

Now, we first calculate $\langle\xi(t_0)|\Omega^{(+)}|\mathbf{k}\beta_0\rangle$ using the integral equation (3.19). Since both $|\mathbf{k}\beta_0\rangle$ and ξ are free states at $t = t_0$ with energies $E_{\mathbf{k}}$ and E_f respectively, and they are chosen to be orthogonal, the first term in the integral equation, the identity term corresponding to 'no scattering', does not contribute. Therefore,

$$\begin{aligned} \langle\xi(t_0)|\Omega^{(+)}|\mathbf{k}\beta_0\rangle &= \frac{1}{i\hbar}\int_{-\infty}^{t_0}\langle\xi(t_0)|U_0(t-t_0)^{-1}V\Omega^{(+)}U_0(t-t_0)|\mathbf{k}\beta_0\rangle dt \\ &= \frac{1}{i\hbar}\left(\int_{-\infty}^{t_0} e^{i(E_f-E_{\mathbf{k}})(t-t_0)/\hbar}dt\right)\langle\xi(t_0)|V\Omega^{(+)}|\mathbf{k}\beta_0\rangle \\ &= \lim_{\epsilon\to 0}\frac{1}{E_{\mathbf{k}}-E_f+i\epsilon}\langle\xi(t_0)|V\Omega^{(+)}|\mathbf{k}\beta_0\rangle, \end{aligned}$$

where we have added an infinitesimal quantity $-i\epsilon$ to $E_f - E_{\mathbf{k}}$ in the exponential to make the singular integral convergent. Plugging the complex conjugate of this back this into our formula for n_ξ we obtain

$$n_\xi = 16\pi^3\hbar^2\rho\,\mathrm{Im}\left[\lim_{\epsilon\to 0}\frac{1}{E_{\mathbf{k}}-E_f-i\epsilon}\right]|\langle\xi(t_0)|V\Omega^{(+)}|\mathbf{k}\beta_0\rangle|^2.$$

Since we have chosen ξ to be an energy eigenstate, there is really no dependence of this transition rate on t_0, because the phase factors occur inside the modulus square. From the well known formula

$$\frac{1}{x - i\epsilon} = P\left(\frac{1}{x}\right) + i\pi\delta(x) \tag{3.26}$$

we calculate the imaginary part and get

$$n_\xi = 16\pi^4 \hbar^2 \rho\, \delta(E_{\mathbf{k}} - E_f)|T(\xi, \phi)|^2, \tag{3.27}$$

where we have introduced the transition amplitude

$$T(\xi, \phi) = \langle \xi(t_0)|V\Omega^{(+)}|\mathbf{k}\beta_0\rangle. \tag{3.28}$$

We choose the final states $|\xi(t_0)\rangle = |\mathbf{p}, \gamma\rangle$ into which transitions are taking place to lie in a narrow range Δ of momentum space $\mathbf{p}$ and sum over the number of transitions.

$$|\xi(t_0)\rangle = |\mathbf{p}, \gamma\rangle, \qquad E_{\mathbf{p}} = |\mathbf{p}|^2/2m, \qquad \mathrm{d}^3\mathbf{p} = m|\mathbf{p}|dE_{\mathbf{p}}d\Omega_{\mathbf{p}},$$

$$\int_\Delta \mathrm{d}^3\mathbf{p}\, n_\xi = 16\pi^4 \hbar^2 \rho m|\mathbf{k}| \int_\Omega |T(\mathbf{p}, \gamma; \mathbf{k}, \beta_0)|^2 d\Omega_{\mathbf{p}}. \tag{3.29}$$

The number of transitions per unit time is proportional to flux, that is density $\times$ velocity: $\rho|\mathbf{k}|/m$. The *cross section* σ_Δ is defined as the ratio of number of transitions in the desired final states (here those in momentum range Δ) to the initial flux. Therefore,

$$\sigma_\Delta = 16\pi^4 \hbar^2 m^2 \int_\Omega |T(\mathbf{p}, \gamma; \mathbf{k}, \beta_0)|^2 d\Omega_{\mathbf{p}} \tag{3.30}$$

It is well worth checking the physical dimensions of the cross section as of an area. Since the momentum eigenstates are normalized as $\langle \mathbf{p}|\mathbf{p}'\rangle = \delta^3(\mathbf{p} - \mathbf{p}')$, the scattering amplitude T being a matrix element of $V\Omega^{(+)}$ has dimensions of (energy)$\times$(momentum)$^{-3}$.

The importance of the integral equation (3.19) is that it can be used to calculate $\Omega^{(+)}$ by iteration if we can regard the interaction term V as small. To the lowest order, the *Born approximation*, $\Omega_0^{(+)}$ is the identity, and therefore the transition amplitude

$$T(\xi, \phi)\Big|_{\text{Born}} = \langle \mathbf{p}\gamma|V|\mathbf{k}\beta_0\rangle. \tag{3.31}$$

As an exercise one can check that for a central potential $V = V(r)$,

$$\langle \mathbf{k}|V|\mathbf{k}'\rangle = \frac{2}{(2\pi\hbar)^2\hbar} \int_0^\infty r^2 dr\, V(r) \frac{\sin(Kr)}{Kr}$$

where $K = |\mathbf{k} - \mathbf{k}'|/\hbar$. With this the differential cross section can be written in the standard form:

$$d\sigma = \left| \frac{2m}{\hbar^2} \int_0^\infty r^2 dr\, V(r) \frac{\sin(Kr)}{Kr} \right|^2 d\Omega_{\mathbf{k}'}. \tag{3.32}$$

We calculate the Rutherford scattering by a Coulomb potential that is

$$V(r) = \frac{C}{r}. \tag{3.33}$$

The integral in question is singular. We "screen" the Coulomb potential replacing C/r by $C\exp(-\epsilon r)/r$ with the understanding that $\epsilon \to 0$. Then

$$\int_0^\infty dr e^{-\epsilon r} \sin(Kr) = \lim_{\epsilon \to 0} \frac{K}{\epsilon^2 + K^2} = \frac{1}{K},$$

and the cross section formula is the familiar

$$\frac{d\sigma}{d\Omega} = \left(\frac{C}{4E}\right)^2 \frac{1}{\sin^4(\theta/2)}, \tag{3.34}$$

where $K^2\hbar^2 = |\mathbf{k} - \mathbf{k}'|^2 = 4k^2 \sin^2(\theta/2)$, θ is the angle between $\mathbf{k}$ and $\mathbf{k}'$, and E is the energy $E = k^2/2m$.

3.4.2 Colliding Non-relativistic Beams

A beam of particles collides with a bunch of more or less stationary particles called the target. This is when we say the scattering in taking place in a 'lab frame'. We can also have the situation when the incident beam collides with another beam.

The free momentum states basis in this case can be chosen to be normalized as

$$\langle \mathbf{p}_1\beta_1, \mathbf{p}_2\beta_2 | \mathbf{p'}_1\beta'_1, \mathbf{p'}_2\beta'_2 \rangle = \delta^3(\mathbf{p_1} - \mathbf{p'_1})\delta_{\beta_1\beta'_1}\delta^3(\mathbf{p_2} - \mathbf{p'_2})\delta_{\beta_2\beta'_2},$$

where the subscript 1 refers to beam particles and 2 to target particles (or particles of the other beam). For the sake of clarity, we call the particles of the second beam as target.

The quantity whose expectation value is to be measured on these bunches is B. We assume that B conserves, that is, commutes with, total linear momentum. Therefore, defining a reduced quantity b, write

$$\begin{aligned}\langle \mathbf{p}_1\beta_1, \mathbf{p}_2\beta_2 | B | \mathbf{p'}_1\beta'_1, \mathbf{p'}_2\beta'_2 \rangle = {} & \delta^3(\mathbf{p_1} + \mathbf{p_2} - \mathbf{p'_1} - \mathbf{p'_2}) \times \\ & b(\mathbf{p}_1\beta_1, \mathbf{p}_2\beta_2; \mathbf{p'}_1\beta'_1, \mathbf{p'}_2\beta'_2).\end{aligned}$$

The average value over the colliding bunch is,

$$\begin{aligned}\langle B\rangle &= \frac{1}{N_1N_2}\sum_{i,j}\langle\phi_{1i}\phi_{2j}|B|\phi_{1i}\phi_{2j}\rangle \\ &= \frac{\rho_1\rho_2}{N_1N_2}\int d^3\mathbf{r}_1\int d^3\mathbf{r}_2\int d^3\mathbf{p}_1\int d^3\mathbf{p}_2\int d^3\mathbf{p}'_1\int d^3\mathbf{p}'_2 \\ &\quad \exp[i(\mathbf{p}_1-\mathbf{p}'_1)\cdot\mathbf{r}_1/\hbar]\exp[i(\mathbf{p}_2-\mathbf{p}'_2)\cdot\mathbf{r}_2/\hbar]\times \\ &\quad f_1(\mathbf{p}_1)^*f_1(\mathbf{p}'_1)f_2(\mathbf{p}_2)^*f_2(\mathbf{p}'_2)\times \\ &\quad \delta^3(\mathbf{p_1}+\mathbf{p_2}-\mathbf{p'_1}-\mathbf{p'_2})b(\mathbf{p}_1\gamma_1,\mathbf{p}_2\gamma_2;\mathbf{p}'_1\gamma_1,\mathbf{p}'_2\gamma_2).\end{aligned}$$

Now integration over $\mathbf{r}_1$ gives $(2\pi\hbar)^3\delta^3(\mathbf{p_1}-\mathbf{p'_1})$ which removes integration on $\mathbf{p}'_1$. However, the momentum conserving delta function then becomes $\delta^3(\mathbf{p}_2-\mathbf{p}'_2)$. Thus $\mathbf{p}'_2$ integration can be done. Therefore the $\mathbf{r}_2$ integration is vacuous and $\rho_2\int d^3\mathbf{r}_2 = N_2$, which cancels with the $1/N_2$ outside the integral signs.

Due to assumed sharp peaks in f_1 and f_2 at $\mathbf{k}_1$ and $\mathbf{k}_2$ respectively, and assuming the absence of sharp peak at these values in b, we get

$$\langle B\rangle = \frac{\rho_1(2\pi\hbar)^3}{N_1}b(\mathbf{k}_1\gamma_1,\mathbf{k}_2\gamma_2;\mathbf{k}_1\gamma_1,\mathbf{k}_2\gamma_2). \tag{3.35}$$

We are now ready to define cross section for colliding beams, or, as explained above, beam on target. Two beams with sharp values of momenta $\mathbf{k}_1$ and $\mathbf{k}_2$ respectively and other quantum numbers γ_1, γ_2 respectively collide and scatter. We are interested in the final states with projection operators

$$P_\xi = |\xi_1\xi_2\rangle\langle\xi_1\xi_2| = \int_\Delta d^3\mathbf{k}'_1\int d^3\mathbf{k}'_2|\mathbf{k}'_1\gamma'_1,\mathbf{k}'_2\gamma'_2\rangle\langle\mathbf{k}'_1\gamma'_1,\mathbf{k}'_2\gamma'_2|.$$

The number of particles making a transition into these final states is

$$n_{\xi_1\xi_2} = \frac{2N_1N_2}{\hbar}\mathrm{Im}\langle B\rangle,$$

with B as before given by

$$B = {\Omega^{(+)}}^\dagger P_\xi V\Omega^{(+)}.$$

The number becomes, using (3.35),

$$n_{\xi_1\xi_2} = \rho_1N_2(2\pi\hbar)^3\frac{2}{\hbar}\mathrm{Im}\,b(\mathbf{k}_1\gamma_1,\mathbf{k}_2\gamma_2;\mathbf{k}_1\gamma_1,\mathbf{k}_2\gamma_2)$$

where b is defined by ($\mathbf{K}=\mathbf{k}_1+\mathbf{k}_2$, etc.)

$$\langle \mathbf{k}_1\gamma_1, \mathbf{k}_2\gamma_2|B|\mathbf{k}_1''\gamma_1, \mathbf{k}_2''\gamma_2\rangle = \delta^3(\mathbf{K}-\mathbf{K}'')b(\mathbf{k}_1\gamma_1, \mathbf{k}_2\gamma_2; \mathbf{k}_1''\gamma_1, \mathbf{k}_2''\gamma_2).$$

To calculate this we begin with the general expression

$$\langle \mathbf{k}_1\gamma_1, \mathbf{k}_2\gamma_2|{\Omega^{(+)}}^\dagger P_\xi V\Omega^{(+)}|\mathbf{k}_1''\gamma_1, \mathbf{k}_2''\gamma_2\rangle$$

and put $\mathbf{k}_1'' = \mathbf{k}_1$, $\mathbf{k}_2'' = \mathbf{k}_2$ at the end. Use the (3.19) to write matrix element of ${\Omega^{(+)}}^\dagger$ in terms of complex conjugate of that of $V\Omega^{(+)}$. There are two momentum conserving delta functions, we can integrate over one by changing to variables

$$\int_\Delta d^3\mathbf{k}_1' \int d^3\mathbf{k}_2' = \int_\Delta d^3\mathbf{K}' \int d^3\mathbf{k}'$$

where $\mathbf{K}' = \mathbf{k}_1' + \mathbf{k}_2'$ is the total momentum and $\mathbf{k}' = (m_1\mathbf{k}_2' - m_2\mathbf{k}_1')/(m_1+m_2)$ the relative momentum. For free states the energies are

$$E_{\mathbf{k}_1\mathbf{k}_2} = \frac{|\mathbf{k}_1|^2}{2m_1} + \frac{|\mathbf{k}_2|^2}{2m_2} = \frac{|\mathbf{K}|^2}{2M} + \frac{|\mathbf{k}|^2}{2\mu} \equiv E_{\mathbf{K}} + e_{\mathbf{k}} = E_{\mathbf{Kk}}$$

with $M = m_1 + m_2$ as the total mass and $\mu = m_1m_2/(m_1+m_2)$ the reduced mass. We get, separating into total and relative momenta,

$$\langle \mathbf{p}_1\beta_1, \mathbf{p}_2\beta_2|V\Omega^{(+)}|\mathbf{p'}_1\beta_1', \mathbf{p'}_2\beta_2'\rangle \equiv \delta^3(\mathbf{P}-\mathbf{P}')T(\mathbf{p}\beta_1\beta_2, \mathbf{p}'\beta_1'\beta_2'; \mathbf{P}).$$

Therefore,

$$\begin{aligned}\langle \mathbf{k}_1\gamma_1, \mathbf{k}_2\gamma_2|{\Omega^{(+)}}^\dagger P_\xi V\Omega^{(+)}\mathbf{k}_1''\gamma_1, \mathbf{k}_2''\gamma_2\rangle = \delta^3(\mathbf{K}-\mathbf{K}'') \times \\ T(\mathbf{k}'\gamma_1'\gamma_2', \mathbf{k}\gamma_1\gamma_2; \mathbf{K})^* \times T(\mathbf{k}'\gamma_1'\gamma_2', \mathbf{k}''\gamma_1\gamma_2; \mathbf{K}') \times \\ (E_{\mathbf{Kk}} - E_{\mathbf{K}'\mathbf{k}'} - i\epsilon)^{-1}.\end{aligned}$$

This defines b in which we put $\mathbf{k}'' = \mathbf{k}$. We can then take the imaginary part, which gives,

$$\begin{aligned}n_{\xi_1\xi_2} = \rho_1 N_2 (2\pi\hbar)^3 \frac{2\pi}{\hbar} \times \\ \int_\Delta d^3\mathbf{k}' \delta(E_{\mathbf{Kk}} - E_{\mathbf{K}'\mathbf{k}'})|T(\mathbf{k}'\gamma_1'\gamma_2', \mathbf{k}\gamma_1\gamma_2; \mathbf{K})|^2.\end{aligned}$$

Because of $\mathbf{K} = \mathbf{K}'$ the energy corresponding to total momentum is already equal, so $E_{\mathbf{Kk}} - E_{\mathbf{K}'\mathbf{k}'} = e_{\mathbf{k}} - e_{\mathbf{k}'}$. As, in the case of potential scattering, we change

$$d^3\mathbf{k} = |\mathbf{k}|^2 d|\mathbf{k}| d\Omega_{\mathbf{k}} = \mu|\mathbf{k}| de_{\mathbf{k}} d\Omega_{\mathbf{k}}$$

this gives us the formula for cross section which is now defined as rate of transitions *per target particle*, that is a division by N_2. The flux is now given by $\rho_1 \times$ relative velocity $= \rho_1 |\mathbf{k}|/\mu$. We get

$$d\sigma = (2\pi)^4 \mu^2 \hbar^2 |T(\mathbf{k}'\gamma_1'\gamma_2', \mathbf{k}\gamma_1\gamma_2; \mathbf{K})|^2 d\Omega_{\mathbf{k}}' \tag{3.36}$$

This is the same formula as the one for the cross section in scattering by a fixed potential, with relative momentum $|\mathbf{k}| = |\mathbf{k}'|$ *for the particle momentum and the reduced mass in place of particle mass.* In the center of mass frame, we must put $\mathbf{K} = 0$.

3.4.3 Relativistic Scattering

Relativistic scattering differs from non-relativistic scattering in following respects. The process of scattering involves *creation* or *annihilation* of particles. Therefore the time evolution operator generators H_0 or H are defined on a much larger Hilbert space, the Fock space.

Relativistic one-particle states are given as momentum space wave functions in the basis $|\mathbf{p}\lambda\rangle$ with normalization

$$\langle \mathbf{p}\lambda | \mathbf{p}'\lambda' \rangle = 2\omega_{\mathbf{p}} \delta^3(\mathbf{p} - \mathbf{p}') \delta_{\lambda\lambda'}$$

where $\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2 c^2}$.

Multi-particle states are defined by tensor products but have to be properly symmetrized (or anti-symmetrized) over identical particles.

We consider two colliding beams with momenta sharply defined around $\mathbf{p}_1$ and $\mathbf{p}_2$ and spins σ_1 and σ_2 respectively.

The final state projection operator is taken as

$$P_\xi = \sum_{\{\lambda\}} \int \frac{d^3\mathbf{k}_1}{2\omega_{\mathbf{k}_1}} \cdots \frac{d^3\mathbf{k}_n}{2\omega_{\mathbf{k}_n}} |\mathbf{k}_1\lambda_1 \ldots \mathbf{k}_n\lambda_n\rangle \langle \mathbf{k}_1\lambda_1 \ldots \mathbf{k}_n\lambda_n|.$$

The argument for non-relativistic colliding beams can be repeated almost step by step except that the total momentum delta function is not integrated and the transition amplitude matrix is defined by

$$\begin{aligned} \langle \mathbf{k}_1\lambda_1 \ldots \mathbf{k}_n\lambda_n | S | \mathbf{p}_1\sigma_1\mathbf{p}_2\sigma_2 \rangle &= \langle \mathbf{k}_1\lambda_1 \ldots \mathbf{k}_n\lambda_n | \mathbf{p}_1\sigma_1\mathbf{p}_2\sigma_2 \rangle \\ &\quad - \frac{2\pi i}{c} \delta^4(P_f - P_i)\, M(\mathbf{k}_1\lambda_1 \ldots \mathbf{k}_n\lambda_n; \mathbf{p}_1\sigma_1\mathbf{p}_2\sigma_2). \end{aligned}$$

Here P_f and P_i are, respectively, the total 4-momenta of the final and the initial states, and a factor $1/c$ appears because, in our notation, the 0-component of 4-momentum

has the physical dimension of momentum, whereas the S-matrix contains a factor $\delta(E_f - E_i)$. This amplitude M (with two initial particles, n final particles, all with relativistic normalization) has the physical dimensions of (velocity)$\times$(momentum)$^{2-n}$.

The cross section can be calculated as (see Problem 3.3 below in Sect. 3.6):

$$\sigma = \frac{(2\pi)^4}{2\omega_{\mathbf{p}_1} 2\omega_{\mathbf{p}_2}} \frac{\hbar^2}{(v_{\text{rel}} c)} \sum_{\{\lambda\}} \int \frac{d^3\mathbf{k}_1}{2\omega_{\mathbf{k}_1}} \cdots \frac{d^3\mathbf{k}_n}{2\omega_{\mathbf{k}_n}} \delta^4(P_f - P_i) \times$$
$$|M(\mathbf{k}_1\lambda_1 \ldots \mathbf{k}_n\lambda_n; \mathbf{p}_1\sigma_1\mathbf{p}_2\sigma_2)|^2 \tag{3.37}$$

where the relative velocity of particles in one beam relative to those in the other is given by

$$v_{\text{rel}} = c \left| \frac{\mathbf{p}_1}{\omega_{\mathbf{p}_1}} - \frac{\mathbf{p}_2}{\omega_{\mathbf{p}_2}} \right|.$$

3.5 Comments on Formulas of Sect. 3.2.3

3.5.1 Moller Operators

1. Although $U(t - t_0)^{-1} U_0(t - t_0)$ are unitary for all finite t their limits $\Omega^{(\pm)}$ as $t \to \pm\infty$ may not be. For every scattering state there is a free state, but, there are, for example, bound states which do not go over to free states in remote past or future. Therefore the mappings $\Omega^{(\pm)}$ are not one-to-one and invertible as they should be if they were unitary.
2. $\Omega^{(\pm)}$ are *norm preserving* operators:

$$\|\Omega^{(+)}\phi\| = \|\phi\|, \qquad \text{and} \tag{3.38}$$
$$\|\Omega^{(-)}\phi\| = \|\phi\| \qquad \forall \phi \in \mathcal{H} \tag{3.39}$$

because they are the limits of norm preserving operators. Thus we can write

$$\Omega^{(\pm)^\dagger} \Omega^{(\pm)} = 1. \tag{3.40}$$

3. $\Omega^{(\pm)}$ are independent of time. This is because the limit $t \to \pm\infty$ remains the same even if the origin of t is changed by a finite constant.
4. $\Omega^{(\pm)^\dagger}$ annihilates bound states:

$$\Omega^{(\pm)^\dagger} \Psi_{\text{bd}} = 0.$$

This can be seen as follows. Let ϕ be any free state and $\Psi = \Omega^{(+)}\phi$ the corresponding scattering state. All scattering states are orthogonal to the bound states.

So, $(\Psi, \Psi_{\rm bd}) = 0$, and therefore $(\phi, {\Omega^{(+)}}^\dagger \Psi_{\rm bd}) = 0$ for any $\phi \in \mathcal{H}$. Similarly for ${\Omega^{(-)}}^\dagger$.

5. As $\Omega^{(\pm)} {\Omega^{(\pm)}}^\dagger \Psi_{\rm bd} = 0$ for all bound states, the operator is a projection on to the subspace orthogonal to the space of bound states. Thus

$$\Omega^{(\pm)} {\Omega^{(\pm)}}^\dagger = 1 - P_{\rm bd} = P_{\rm scatt}$$

where the operators $P_{\rm scatt}$ and $P_{\rm bd}$ are projection operators on the subspaces of scattering and bound states respectively.

6. For any fixed t,

$$\begin{aligned} U(t)\Omega^{(\pm)} &= \lim_{s \to \mp\infty} U(t)U(s)^{-1}U_0(s) \\ &= \lim_{s \to \mp\infty} U(t-s)U_0(s-t)U_0(t) \\ &= \lim_{u \to \mp\infty} U(u)^{-1}U_0(u)U_0(t), \quad u = t - s \\ &= \Omega^{(\pm)} U_0(t). \end{aligned} \tag{3.41}$$

Differentiating with respect to t and putting $t = 0$ gives another useful result,

$$H\Omega^{(\pm)} = \Omega^{(\pm)} H_0, \tag{3.42}$$

and its adjoint equation,

$${\Omega^{(\pm)}}^\dagger H = H_0 {\Omega^{(\pm)}}^\dagger.$$

Physically, this equation means that the energy spectrum of free particle states is contained in the spectrum of the total Hamiltonian: if ϕ_E is an eigenstate of H_0 with energy E, $\Omega^{(\pm)}\phi_E$ is an eigenstate of H with the same eigenvalue. One should appreciate that energy eigenvectors of H_0 and H may have the same labels, but those of H_0 span the whole space $\mathcal{H}$ whereas those of H span only the subspace of scattering states.

3.5.2 S-Matrix

S is unitary.

$$SS^\dagger = {\Omega^{(-)}}^\dagger \Omega^{(+)} {\Omega^{(+)}}^\dagger \Omega^{(-)} = {\Omega^{(-)}}^\dagger (1 - P_{\rm bd}) \Omega^{(-)} = 1$$

because ${\Omega^{(-)}}^\dagger P_{\rm bd} = 0$. Similarly, $S^\dagger S = 1$.

S commutes with the *free* hamiltonian H_0: using (3.15) and (3.16)

$$SH_0 = \Omega^{(-)^\dagger}\Omega^{(+)}H_0 = \Omega^{(-)^\dagger}H\Omega^{(+)} = H_0\Omega^{(-)^\dagger}\Omega^{(+)} = H_0S.$$

The S-matrix may not commute with the total Hamiltonian H.

3.5.3 Integral Equations

1. Define $V \equiv H - H_0$, then

$$\Omega^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^{0} U_0(t)^{-1}V\Omega^{(+)}U_0(t)dt. \tag{3.43}$$

2. Similarly,

$$S = 1 + \frac{1}{i\hbar}\int_{-\infty}^{\infty} U_0(t)^{-1}V\Omega^{(+)}U_0(t)dt. \tag{3.44}$$

As $U_0(t)$ is the free evolution operator, it is supposed to be known, and if V can be considered as small, the first equation (called the Lippmann–Schwinger equation) provides a way to calculate $\Omega^{(+)}$ by iteration. For example, the zero order approximation to $\Omega^{(+)}$ (for $V = 0$) is $\Omega_0^{(+)} = 1$, and the first order, called the 'first Born approximation' is, by substituting $\Omega_0^{(+)}$ on the right hand side,

$$\Omega_1^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^{0} U_0(t)^{-1}VU_0(t)dt. \tag{3.45}$$

Once $\Omega^{(+)}$ is known, S can be calculated from the second equation above.

Proof for the Integral Equations

Start from the trivial identity

$$\int_{-\infty}^{0}\frac{d}{dt}[U(t)^{-1}U_0(t)]dt = 1 - \Omega^{(+)}.$$

Inside the integral sign differentiation (using $i\hbar dU/dt = HU$ and $i\hbar dU_0/dt = H_0U_0$) gives

$$\Omega^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^{0} U(t)^{-1}VU_0(t)dt.$$

The adjoint of this equation is

$$\Omega^{(+)^\dagger} = 1 - \frac{1}{i\hbar}\int_{-\infty}^{0} U_0(t)^{-1}VU(t)dt. \tag{3.46}$$

Similarly,

$$\Omega^{(-)^\dagger} = 1 - \frac{1}{i\hbar}\int_0^\infty U_0(t)^{-1} V U(t) dt. \tag{3.47}$$

Multiply (3.46) by $\Omega^{(+)}$ on the right and use $\Omega^{(+)^\dagger}\Omega^{(+)} = 1$ to get

$$\Omega^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^0 U_0(t)^{-1} V U(t) \Omega^{(+)} dt. \tag{3.48}$$

This can be written in the desired form by using (3.14):

$$\Omega^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^0 U_0(t)^{-1} V \Omega^{(+)} U_0(t) dt. \tag{3.49}$$

Next, from (3.46) and (3.47),

$$\Omega^{(+)^\dagger} - \Omega^{(-)^\dagger} = -\frac{1}{i\hbar}\int_{-\infty}^\infty U_0(t)^{-1} V U(t) dt, \tag{3.50}$$

Multiply on the right by $\Omega^{(+)}$ to get (again using (3.14)),

$$S = \Omega^{(-)^\dagger}\Omega^{(+)} = 1 + \frac{1}{i\hbar}\int_{-\infty}^\infty U_0(t)^{-1} V \Omega^{(+)} U_0(t) dt. \tag{3.51}$$

3.5.4 $\Omega^{(+)}$ and S in Energy Basis

Let $|E\alpha\rangle$ be eigenvectors of H_0 with normalization

$$H_0|E\alpha\rangle = E|E\alpha\rangle \qquad \langle E\alpha|E'\alpha'\rangle = \delta(E - E')\delta_{\alpha\alpha'}$$

where α are observables other than energy needed to from a complete set of commuting observables. Sandwich the integral equation (3.19) in these states, and as

$$U_0(t)|E\alpha\rangle = \exp(-iEt/\hbar)|E\alpha\rangle,$$

we obtain,

$$\langle E\alpha|\Omega^{(+)}|E'\alpha'\rangle = \delta(E - E')\delta_{\alpha\alpha'} + \frac{\langle E\alpha|V\Omega^{(+)}|E'\alpha'\rangle}{E' - E + i\epsilon}, \tag{3.52}$$

where we interpret the singular integral

$$\int_{-\infty}^{0} \exp[i(E-E')t/\hbar]dt = \lim_{\epsilon\to 0}\int_{-\infty}^{0} \exp[i(E-E'-i\epsilon)t/\hbar]dt$$
$$= \lim_{\epsilon\to 0} \frac{i\hbar}{E'-E+i\epsilon}.$$

The sign of $i\epsilon$ is chosen to make the integral convergent. Similarly, the equation for the S-matrix (3.20) can be written

$$\langle E\alpha|S|E'\alpha'\rangle = \delta(E-E')\delta_{\alpha\alpha'} - 2\pi i\delta(E-E')\langle E\alpha|V\Omega^{(+)}|E'\alpha'\rangle$$
$$\equiv \delta(E-E')\delta_{\alpha\alpha'} - 2\pi i\delta(E-E')T(E\alpha, E'\alpha'). \tag{3.53}$$

The quantity $T(E\alpha, E'\alpha') \equiv \langle E\alpha|V\Omega^{(+)}|E'\alpha'\rangle$ occurs very frequently in scattering theory and is called the off-shell transition amplitude or the off-shell T-matrix. But the transition amplitude in the equation above occurs with the energy conserving delta function and is actually the *on-shell* transition amplitude or T-matrix $T_E(\alpha, \alpha') \equiv T(E\alpha, E\alpha')$.

3.6 Problems

Problem 3.1 Define the Green's function

$$G_E = \frac{1}{i\hbar}\int_{-\infty}^{0} dt\ \exp[i(H_0-E)t/\hbar].$$

Show that in the coordinate basis $|\mathbf{r}\rangle$ of a non-relativistic particle of mass m

$$\langle \mathbf{r}|G_E|\mathbf{r}'\rangle = -\frac{m}{2\pi\hbar^2}\frac{\exp(ipR/\hbar)}{R}, \qquad p = +\sqrt{2mE},\quad R = |\mathbf{r}-\mathbf{r}'|.$$

Problem 3.2 Most textbooks discuss non-relativistic potential scattering by assuming the 'Sommerfeld radiation condition' in the asymptotic region. That is, for an incoming beam represented by $A\exp(ikz)$, the wave function of the scattered 'wave' is assumed to have the form

$$A\left(e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right), \qquad kr\to\infty.$$

Justify this expression from the formal scattering theory.

Hint: The incoming state (at $t=0$) $\phi_E(0)$ is an eigenstate of energy with $E = k^2\hbar^2/2m$. Since $H\Omega^{(+)} = \Omega^{(+)}H_0$ (see (3.42)), the scattering state $\Psi_E(0) = \Omega^{(+)}\phi_E(0)$ is also an eigenstate of energy with the same value. Therefore, from the integral equation for $\Omega^{(+)}$ (3.48),

$$\Psi_E(0) = \phi_E(0) + \frac{1}{i\hbar}\int_{-\infty}^{0} \exp(iH_0t/\hbar)V\exp(-iEt/\hbar)\Psi_E(0)$$
$$= \phi_E(0) + G_E V\Psi_E(0),$$

where G_E is as defined in the previous problem. Multiply on the left by $\langle \mathbf{r}|$, use the result of the previous problem, and the fact that $V(\mathbf{r}')$ is non-zero only in a small region.

Problem 3.3 Complete the steps for derivation of the relativistic cross-section formula (3.37): (We have omitted the spin variables σ_1, σ_2 in the initial state $|\mathbf{p}_1\sigma_1\mathbf{p}_2\sigma_2\rangle$. They play no role in the derivation and can be restored in the end.)

$$\sigma = \frac{(2\pi)^4}{2\omega_{\mathbf{p}_1}2\omega_{\mathbf{p}_2}}\frac{\hbar^2}{(v_{\rm rel}c)}\sum_{\{\lambda\}}\int \frac{d^3\mathbf{k}_1}{2\omega_{\mathbf{k}_1}}\cdots\frac{d^3\mathbf{k}_n}{2\omega_{\mathbf{k}_n}}\delta^4(P_f - P_i)\,|M(\mathbf{k}_1\lambda_1\ldots\mathbf{k}_n\lambda_n;\,\mathbf{p}_1\mathbf{p}_2)|^2$$

where we define the relativistic transition matrix M as

$$\langle \mathbf{k}_1\lambda_1\ldots\mathbf{k}_n\lambda_n|V\Omega^{(+)}|\mathbf{p}_1\mathbf{p}_2\rangle = \delta^3(\mathbf{P}_f - \mathbf{p}_1 - \mathbf{p}_2)M(\mathbf{k}_1\lambda_1\ldots\mathbf{k}_n\lambda_n;\,\mathbf{p}_1\mathbf{p}_2),$$

or, equivalently, from

$$\langle f|S|i\rangle = \langle f|i\rangle - \frac{2\pi i}{c}\delta^4(P_f - P_i)\,\langle f|M|i\rangle.$$

Outline of solution: Let $|\xi\rangle$ be the final state and $|\phi_1\phi_2\rangle$ the two-particle initial state. The number of transitions per unit time for two bunches of size N_1 and N_2 is given by

$$n_\xi = N_1N_2\frac{2}{\hbar}\mathrm{Im}\langle\phi_1\phi_2|B|\phi_1\phi_2\rangle,\quad B = \Omega^{(+)^\dagger}|\xi\rangle\langle\xi|V\Omega^{(+)}.$$

The average over the bunches is

$$\langle\phi_1\phi_2|B|\phi_1\phi_2\rangle = \frac{1}{N_1N_2}\sum_{i,j}\langle\phi_{1i}\phi_{2j}|B|\phi_{1i}\phi_{2j}\rangle,$$

where $|\phi_{1i}\rangle$ are spatially displaced by $\mathbf{r}_i$ and $|\phi_{2j}\rangle$ by $\mathbf{r}_j$. Substituting the wave functions

$$\langle\phi_1\phi_2|B|\phi_1\phi_2\rangle = \frac{1}{N_1N_2}\sum_{i,j}\int\frac{\mathrm{d}^3\mathbf{p}_1'}{2\omega_{\mathbf{p}_1'}}\int\frac{\mathrm{d}^3\mathbf{p}_2'}{2\omega_{\mathbf{p}_2'}}\int\frac{\mathrm{d}^3\mathbf{p}_1}{2\omega_{\mathbf{p}_1}}\int\frac{\mathrm{d}^3\mathbf{p}_2}{2\omega_{\mathbf{p}_2}}$$
$$\exp(i(\mathbf{p}_1' - \mathbf{p}_1)\cdot\mathbf{r}_i)\exp(i(\mathbf{p}_2' - \mathbf{p}_2)\cdot\mathbf{r}_j)\times$$
$$f_1^*(\mathbf{p}_1')f_1(\mathbf{p}_1)f_2^*(\mathbf{p}_2')f_2(\mathbf{p}_2)\,\langle\mathbf{p}_1'\mathbf{p}_2'|B|\mathbf{p}_1\mathbf{p}_2\rangle \tag{3.54}$$

We first calculate

$$\begin{aligned}\langle \mathbf{p}_1' \mathbf{p}_2' | B | \mathbf{p}_1 \mathbf{p}_2 \rangle &= \langle \mathbf{p}_1' \mathbf{p}_2' | {\Omega^{(+)}}^\dagger | \xi \rangle \langle \xi | V \Omega^{(+)} | \mathbf{p}_1 \mathbf{p}_2 \rangle \\ &= (\langle \xi | \Omega^{(+)} | \mathbf{p}_1' \mathbf{p}_2' \rangle)^* \langle \xi | V \Omega^{(+)} | \mathbf{p}_1 \mathbf{p}_2 \rangle.\end{aligned}$$

We have seen from the integral equation for $\Omega^{(+)}$ that (as $\langle \xi | \mathbf{p}_1' \mathbf{p}_2' \rangle = 0$)

$$\langle \xi | \Omega^{(+)} | \mathbf{p}_1' \mathbf{p}_2' \rangle = \frac{\langle \xi | V \Omega^{(+)} | \mathbf{p}_1' \mathbf{p}_2' \rangle}{E' - E_\xi + i\epsilon}, \qquad E' = E_{\mathbf{p}_1'} + E_{\mathbf{p}_2'}.$$

Since total linear momentum is conserved,

$$\langle \xi | V \Omega^{(+)} | \mathbf{p}_1 \mathbf{p}_2 \rangle = \delta^3(\mathbf{P}_\xi - \mathbf{p}_1 - \mathbf{p}_2) M(\xi, \mathbf{p}_1 \mathbf{p}_2).$$

Therefore,

$$\begin{aligned}\langle \mathbf{p}_1' \mathbf{p}_2' | B | \mathbf{p}_1 \mathbf{p}_2 \rangle = \delta^3(\mathbf{P}_\xi - \mathbf{p}_1 - \mathbf{p}_2) \delta^3(\mathbf{P}_\xi - \mathbf{p}_1' - \mathbf{p}_2') \times \\ \frac{1}{E' - E_\xi - i\epsilon} M(\xi, \mathbf{p}_1' \mathbf{p}_2')^* M(\xi, \mathbf{p}_1 \mathbf{p}_2).\end{aligned}$$

The two delta factors above can also be written as

$$\delta^3(\mathbf{P}_\xi - \mathbf{p}_1 - \mathbf{p}_2) \delta^3(\mathbf{P}_\xi - \mathbf{p}_1' - \mathbf{p}_2') = \delta^3(\mathbf{P}_\xi - \mathbf{p}_1 - \mathbf{p}_2) \delta^3(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_1' - \mathbf{p}_2')$$

Substitute this in (3.54) and (1) convert sums over i and j as $\rho_1 \int \mathrm{d}^3\mathbf{r}_1$ and $\rho_2 \int \mathrm{d}^3\mathbf{r}_2$ respectively; (2) integration over $\mathbf{r}_1$ produces $(2\pi\hbar)^3 \delta(\mathbf{p}_1' - \mathbf{p}_1)$ which allows integral over $\mathbf{p}_1'$ to be performed; (3) since $\mathbf{p}_1' = \mathbf{p}_1$, the delta functions in $\langle B \rangle$ make a $\delta(\mathbf{p}_2' - \mathbf{p}_2)$ so that integral over $\mathbf{p}_2'$ can be performed; (4) the $\mathbf{r}_2$ integration is vacuous and $\rho_2 \int \mathrm{d}^3\mathbf{r}_2 = N_2$ which cancels the $1/N_2$ factor outside. Thus (3.54) becomes (recall the derivation for the non-relativistic case)

$$\langle \phi_1 \phi_2 | B | \phi_1 \phi_2 \rangle = \frac{\rho_1}{N_1} \frac{1}{2\omega_{\mathbf{p}_1} 2\omega_{\mathbf{p}_2}} \frac{(2\pi\hbar)^3}{E - E_\xi - i\epsilon} \delta^3(\mathbf{P}_\xi - \mathbf{p}_1 - \mathbf{p}_2) |M(\xi, \mathbf{p}_1 \mathbf{p}_2)|^2.$$

This can be substituted in the formula for n_ξ. The $-i\epsilon$ in the denominator gives the imaginary part of the factor as energy delta function times π by (3.26), and the N_2 factor can be omitted if we are calculating cross-section *per target particle*. The connection of M as a matrix element of $V\Omega^{(+)}$ related to the S-matrix is easily seen from the integral equation for S.

Problem 3.4 (*Relativistic decay rate*) The decay rate, of a relativistic particle of rest-mass m is defined as the fraction of particles making a transition per unit time from initial state $|\mathbf{k}_0\rangle$ to final states $|\xi\rangle = |\mathbf{k}_1 \lambda_1 \ldots \mathbf{k}_n \lambda_n\rangle$. Show that it is given by

$$\frac{1}{\tau} = \frac{2\pi}{\hbar}\frac{1}{2\omega_{\mathbf{k}}}\frac{1}{c}\int \frac{d^3\mathbf{k}_1}{2\omega_{\mathbf{k}_1}}\cdots\frac{d^3\mathbf{k}_n}{2\omega_{\mathbf{k}_n}}\delta^4(P_f - P_i)\,|M(\mathbf{k}_1\lambda_1\ldots\mathbf{k}_n\lambda_n;\mathbf{k}_0)|^2, \quad (3.55)$$

where P_i is the initial 4-momentum vector $(\omega_{\mathbf{k}_0}, \mathbf{k}_0)$ and P_f the total 4-momentum of the final state.

Hint: In this case there is no need to consider the bunch. The average of B in the initial state $|\phi\rangle$ is

$$\langle\phi|B|\phi\rangle = \int \frac{\mathrm{d}^3\mathbf{k}'}{2\omega_{\mathbf{k}'}}\int \frac{\mathrm{d}^3\mathbf{k}}{2\omega_{\mathbf{k}}}\, f^*(\mathbf{k}')f(\mathbf{k})\,\langle\mathbf{k}'|B|\mathbf{k}\rangle,$$

and, calling the final state as ξ,

$$\langle\mathbf{k}'|B|\mathbf{k}\rangle = \delta^3(\mathbf{p}_\xi - \mathbf{k})\delta^3(\mathbf{p}_\xi - \mathbf{k}')\frac{M(\xi;\mathbf{k}')^*M(\xi;\mathbf{k})}{E' - E_\xi - i\epsilon}.$$

The integral on $\mathbf{k}'$ can be done, which puts $\mathbf{k}' = \mathbf{k}$, and as $f(\mathbf{k})$ is peaked around $\mathbf{k}_0$, the matrix elements (including the delta function!) can be pulled out at value $\mathbf{k}_0$. So,

$$\begin{aligned} n_\xi &= N\frac{2}{\hbar}\mathrm{Im}\,\langle\phi|B|\phi\rangle \\ &= N\frac{2\pi}{\hbar}\frac{1}{2\omega_{\mathbf{k}_0}}\frac{1}{c}\delta^4(P_\xi - P_i)|M(\xi;\mathbf{k}_0)|^2. \end{aligned}$$

Divide by N to get the fraction and integrate over the final states to get the decay rate formula. Usually, the particle decays in its rest frame where $\mathbf{k}_0 = 0$ and $\omega_{\mathbf{k}_0} = mc$.

The physical dimensions of $M(\mathbf{k}_1\lambda_1\ldots\mathbf{k}_n\lambda_n;\mathbf{k}_0)$ is (energy)$\times$(momentum)$^{2-n}$.

3.7 Notes and References

The probability interpretation of quantum mechanics was first given by Born [1] in connection with scattering theory. The idea of 'coherent bunches' relating actual number densities of particles to matrix elements of sharp momentum eigenstates has been used by the present author in classroom teaching. The formal scattering theory given here closely follows the treatment in Newton's book [2].

References

1. M. Born, Z. Phys. **38**, 803 (1926), an English translation appears in G. Ludwig (ed.), *Wave Mechanics* (Pergamon Press, Oxford, 1968)
2. R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966)

Chapter 4
Star-Product Formulation of Quantum Mechanics

Star-product, $(f * g)(q, p)$ of two functions $f(q, p)$ and $g(q, p)$ on the phase space is a non-commutative product corresponding to the Hilbert space product $\hat{f}\hat{g}$ of operators $\hat{f}$ and $\hat{g}$. The quantum theory can be developed in analogy with the classical mechanics on the phase space, not with a Poisson bracket, but with a Moyal bracket $(f * g - g * f)/i\hbar$. This formulation is also called the deformation theory of quantization.

4.1 Weyl Ordering and the Star-Product

Quantization is the process of arriving at a quantum theory starting from a classical theory by a set of rules. The usual procedure of assigning to phase space canonical variables q and p the Hermitian operators $\hat{q}$ and $\hat{p}$ in a Hilbert space runs into 'ordering problems' when we seek to define observables other than the simplest ones because operators do not commute. For example, how is $q^2 p^2$ to be quantized? If we take $\hat{q}^2 \hat{p}^2$ as the corresponding operator, it is not Hermitian. There are many choices even for a Hermitian operator: $(\hat{q}^2 \hat{p}^2 + \hat{p}^2 \hat{q}^2)/2$ or $\hat{q}\,\hat{p^2}\,\hat{q}$ or $(\hat{q}\,\hat{p}\hat{q}\,\hat{p} + \hat{p}\hat{q}\,\hat{p}\hat{q})/2$ etc. As part of our quantization procedure we must also provide (at least for the physically relevant observables) a rule for defining ordering of operators when converting a classical phase space quantity into its quantum mechanical counterpart.

One of the oldest ordering rule is due to Weyl. It is simple to state: Let the classical quantity $f(q, p)$ to be quantized be written as a Fourier transform

$$f(q, p) = \int du\, dv \exp[i(uq + vp)/\hbar]\ \tilde{f}\ (u, v).$$

The operator $\hat{f}$ to be associated with $f(q, p)$ by this rule is obtained by this same formula replacing classical q and p by $\hat{q}$ and $\hat{p}$ respectively on the right hand side in the exponential:

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_4

$$\hat{f} = \int du\, dv \exp[i(u\hat{q} + v\hat{p})/\hbar]\ \tilde{f}\ (u, v). \tag{4.1}$$

We note down here the inverse formula which expresses the phase space function $f(q, p)$ given the operator $\hat{f}$:

$$f(q, p) = \int dx \exp(ipx/\hbar)\langle q - x/2|\hat{f}|q + x/2\rangle. \tag{4.2}$$

We will prove this formula in Sect. 4.4 below.

Note: Throughout this chapter we stick to one degree of freedom for writing formulas. The generalization to many degrees of freedom is straightforward.

Exercise 4.1 Show that if $f(q, p)$ is real then $\hat{f}$ as defined above will be Hermitian.

We can now ask the natural question: if $\hat{f}$ corresponds to $f(q, p)$ under Weyl ordering, and $\hat{g}$ to $g(q, p)$, is there a function, which under the Weyl rule, will correspond to the product $\hat{f}\hat{g}$?

$$\begin{array}{rl} f(q, p) & \longrightarrow \ \hat{f} \\ g(q, p) & \longrightarrow \ \hat{g} \\ ?? & \longrightarrow \hat{f}\hat{g} \end{array}$$

The answer is yes, and the function on phase space is written as $(f * g)(q, p)$ given by

$$(f * g)(q, p) = f(q, p) \exp\left[\frac{i\hbar}{2} \overleftrightarrow{P}\right] g(q, p) \tag{4.3}$$

where $\overleftrightarrow{P}$ is the Poisson bracket bi-differential operator

$$\overleftrightarrow{P} = \frac{\overleftarrow{\partial}}{\partial q}\frac{\overrightarrow{\partial}}{\partial p} - \frac{\overleftarrow{\partial}}{\partial p}\frac{\overrightarrow{\partial}}{\partial q} \tag{4.4}$$

the arrows indicating the direction in which the differential operators act. Thus,

$$f(q, p)(\overleftrightarrow{P})^n g(q, p) = \sum_{r=0}^{n} \frac{n!}{r!(n-r)!}(-1)^{n-r} \frac{\partial^n f}{\partial q^r \partial p^{n-r}} \frac{\partial^n g}{\partial p^r \partial q^{n-r}}$$

and so,

$$(f * g)(q, p) = \sum_{r,s=0}^{\infty} \left(\frac{i\hbar}{2}\right)^{r+s} \frac{(-1)^s}{r!s!} \frac{\partial^{r+s} f}{\partial q^r \partial p^s} \frac{\partial^{r+s} g}{\partial p^r \partial q^s}. \tag{4.5}$$

It is clear that this 'star-product' of f and g is not commutative $f * g \neq g * f$ unless $\hat{f}$ and $\hat{g}$ commute. And $f * g$ may not be real because $\hat{f}\hat{g}$ may not be Hermitian.

But the star-product is associative, $f * (g * h) = (f * g) * h$ because the product of operators $\hat{f}, \hat{g}, \hat{h}$ etc. is associative. A direct proof is still required, but we omit it here. See references at the end of this chapter.

4.2 Derivation for Star-Product Expression

The product $\hat{f}\hat{g}$

$$\hat{f}\hat{g} = \int du\, dv\, du'\, dv'\ \exp[i(u\hat{q} + v\hat{p})/\hbar] \exp[i(u'\hat{q} + v'\hat{p})/\hbar]\ \widetilde{f}\ (u, v)\ \widetilde{g}\ (u', v')$$

has two operator exponentials. But the commutator of the exponents is a c-number, commuting with all operators. Using the identity

$$\exp(\hat{A}) \exp(\hat{B}) = \exp(\hat{A} + \hat{B}) \exp([\hat{A}, \hat{B}]/2) \tag{4.6}$$

which holds when $[\hat{A}, \hat{B}]$ commutes with both $\hat{A}$ and $\hat{B}$, we obtain therefore

$$\begin{aligned}\hat{f}\hat{g} = \int du\, dv\, du'\, dv'\ &\exp[i((u + u')\hat{q} + (v + v')\hat{p})/\hbar] \\ &\times \exp[-i(uv' - vu')/2\hbar]\ \widetilde{f}\ (u, v)\ \widetilde{g}\ (u', v').\end{aligned} \tag{4.7}$$

This looks like the Weyl ordering of some phase space function whose Fourier transform goes with variables $(u + u')$ and $(v + v')$ and an extra factor of $\exp[-i(uv' - vu')/2\hbar]$. Instead of changing to those variables we use a trick that simplifies things enormously. We start with the ordinary product of $f(q, p)$ and $g(q', p')$ at two *different* points:

$$\begin{aligned}f(q, p)g(q', p') = \int du\, dv\, du'\, dv'\ &\exp[i(uq + vp + u'q' + v'p')/\hbar] \\ &\times\ \widetilde{f}\ (u, v)\ \widetilde{g}\ (u', v').\end{aligned}$$

A differential operator like

$$\left(\frac{\partial}{\partial q}\frac{\partial}{\partial p'} - \frac{\partial}{\partial p}\frac{\partial}{\partial q'}\right)$$

produces a factor $-(uv' - vu')/\hbar^2$ inside the integral sign when acting on $f(q, p)g(q', p')$. If we then put $q = q'$ and $p = p'$ then

$$\frac{i\hbar}{2}\left(\frac{\partial}{\partial q}\frac{\partial}{\partial p'}-\frac{\partial}{\partial p}\frac{\partial}{\partial q'}\right) f(q,p)g(q',p')\Big|_{q=q',p=p'}$$
$$=\int du\,dv\,du'\,dv'\ \exp[i((u+u')q+(v+v')p)/\hbar]$$
$$\times[-i(uv'-vu')/2\hbar]\ \widetilde{f}\ (u,v)\ \widetilde{g}\ (u',v').$$

We can apply repeated powers of this operator as in an exponential and get

$$\exp\left[\frac{i\hbar}{2}\left(\frac{\partial}{\partial q}\frac{\partial}{\partial p'}-\frac{\partial}{\partial p}\frac{\partial}{\partial q'}\right)\right] f(q,p)g(q',p')\Big|_{q=q',p=p'}$$
$$=\int du\,dv\,du'\,dv'\ \exp[i((u+u')q+(v+v')p)/\hbar]$$
$$\times\exp[-i(uv'-vu')/2\hbar]\ \widetilde{f}\ (u,v)\ \widetilde{g}\ (u',v').$$

Comparing it with (4.7) above, we see that the left hand side is what we have defined as $(f*g)(q,p)$ because converting q and p on the right hand side into $\hat{q}$ and $\hat{p}$ we get precisely $\hat{f}\hat{g}$. This completes the proof for the expression for the star product.

4.3 Wigner Distribution Function

Dynamics, whether classical or quantum mechanical, requires three things: specification of state, specification of observables, and equations of motion. We have seen the correspondence between the observables on the phase space and the Hermitian (or self-adjoint) operators on Hilbert space through the Weyl ordering. What can we say about specification of states?

In classical mechanics the variables (q,p) play a double role: they are both coordinates on the phase space as well as, like other observables, *functions* on it. Strictly speaking, it is not functions q, p that specify the state, rather it is their specific *values* that identify the state. This is clarified if we define the classical state by a distribution function. For example

$$\rho_{\rm cl}(q,p)=\delta(q-q_0)\delta(p-p_0)$$

is the state corresponding to the phase space point with coordinates (q_0,p_0). In a general case the distribution function may not be a sharp Dirac delta function, but a probability distribution $\rho(q,p)$, positive definite and giving unity when integrated over the whole phase space. Dynamics will determine trajectories $t\to(q(t),p(t))$ for all the points of the phase space, leading to evolution of the probability distribution $\rho(q,p)$.

For quantum theory we represent the state by a unit vector ψ in the Hilbert space of the system. But as we discussed earlier, it is the unit ray which determines the state,

all vectors in the ray being equally qualified to represent the same physical state. A better way to represent the state is to use the projection operator $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$.

We therefore look for a phase-space function $\rho_\psi(q, p)$ which under the Weyl ordering will produce $\hat{\rho}_\psi$:

$$\begin{aligned}\hat{\rho}_\psi = |\psi\rangle\langle\psi| &= \int du\, dv \exp[i(u\hat{q} + v\hat{p})/\hbar]\, \widetilde{\rho}\,(u, v) \\ &= \int du\, dv \exp[iu\hat{q}/\hbar] \exp[iv\hat{p}/\hbar] \exp[iuv/2\hbar]\, \widetilde{\rho}\,(u, v),\end{aligned}$$

where in the second step we have separated the exponents using the identity (4.6). Once we identify $\widetilde{\rho}\,(u, v)$ we can construct the phase space function $\rho_\psi(q, p)$. Let $|q'\rangle$ etc. be the eigenstates of $\hat{q}$. By taking matrix element with these eigenstates

$$\psi(q')\psi^*(q'') = \int du\, dv \exp[iuv/2\hbar] \exp[iuq'/\hbar]\, \langle q'| \exp[iv\hat{p}/\hbar]|q''\rangle\, \widetilde{\rho}\,(u, v).$$

As

$$\langle q'| \exp[iv\hat{p}/\hbar]|q''\rangle = \delta(v + q' - q'')$$

we get

$$\psi(q')\psi^*(q'') = \int du \exp[iu(q' + q'')/2\hbar]\, \widetilde{\rho}\,(u, q'' - q').$$

We can choose variables $Q = (q' + q'')/2$ and $x = q'' - q'$ and invert the Fourier transform to obtain $\widetilde{\rho}$ and from there

$$\rho_\psi(q, p) = \int dx \exp(ixp/\hbar)\psi(q - x/2)\psi^*(q + x/2). \tag{4.8}$$

The function $\rho_\psi(q, p)$ on phase space is called the Wigner distribution function.

The Wigner distribution function when integrated over the whole phase space (with a suitable factor) gives unity:

$$\begin{aligned}\int \frac{dq\, dp}{(2\pi\hbar)} \rho_\psi(q, p) &= \int \frac{dq\, dp}{2\pi\hbar} \int dx \exp(ixp/\hbar)\psi(q - x/2)\psi^*(q + x/2) \\ &= \int dq |\psi(q)|^2 \\ &= 1. \end{aligned} \tag{4.9}$$

Exercise 4.2 The function $\rho_\psi(q, p)$ can be interpreted as a 'quasi probability' density on the phase space with marginal probabilities for q and p equal to $|\langle q|\psi\rangle|^2$ and $|\langle p|\psi\rangle|^2$ respectively:

$$\int \frac{dp}{(2\pi\hbar)}\rho_\psi(q, p) = |\langle q|\psi\rangle|^2,$$
$$\int \frac{dq}{(2\pi\hbar)}\rho_\psi(q, p) = |\langle p|\psi\rangle|^2.$$

As we shall see in Sect. 4.5 below, Wigner distributions for two normalized states ϕ and ψ allow us to calculate the transition probability

$$|\langle\phi|\psi\rangle|^2 = \mathrm{Tr}(|\phi\rangle\langle\phi|\,|\psi\rangle\langle\psi|) = \int \frac{dq\,dp}{(2\pi\hbar)}\rho_\phi(q, p)\rho_\psi(q, p). \tag{4.10}$$

This, incidentally, also shows that $\rho_\psi(q, p)$ (for any ψ) cannot be a true positive-definite probability distribution, because if ϕ and ψ are orthogonal, the left hand side will be zero. It can be proved that the Wigner function is positive definite only for Gaussian wave functions.

Exercise 4.3 Calculate the Wigner function for the ground state ψ_0 and for the first level ψ_1 of a one-dimensional harmonic oscillator. Estimate the area of phase space where $\rho_{\psi_1}(q, p)$ is negative.

4.4 Trace of $\hat{f}$

Let $\hat{f}$ be an observable with the corresponding phase space function $f(q, p)$. We show that the trace, (which is actually independent of the basis in which it is taken, but here we choose $|q'\rangle$ for simplicity), is equal to the integral of $f(q, p)$ over the phase space:

$$\mathrm{Tr}\,\hat{f} = \int dq'\langle q'|\hat{f}|q'\rangle = \int \frac{dq\,dp}{(2\pi\hbar)}f(q, p). \tag{4.11}$$

The matrix element of $\hat{f}$ as given by (4.1) is

$$\begin{aligned}\langle q'|\hat{f}|q''\rangle &= \int du\,dv\langle q'|\exp[i(u\hat{q} + v\hat{p})/\hbar]|q''\rangle\ \widetilde{f}\ (u, v),\\ &= \int du\exp[iu(q' + q'')/2\hbar]\ \widetilde{f}\ (u, q'' - q'),\end{aligned}$$

where we have taken exactly the same steps as in calculating the matrix elements of $|\psi\rangle\langle\psi|$ for the Wigner distribution. Substituting the expression for the inverse Fourier transform

$$\widetilde{f}\ (u, q'' - q') = \int \frac{dq\,dp}{2\pi\hbar} exp[-iuq/\hbar - i(q'' - q')p/\hbar] f(q, p)$$

in the above equation and integrating over u,

$$\langle q'|\hat{f}|q''\rangle = \int \frac{dq\,dp}{(2\pi\hbar)}\,\delta\left(q - \frac{q'' + q'}{2}\right) \exp[-ip(q'' - q')/\hbar] f(q, p). \quad (4.12)$$

Before we complete the proof for the trace formula, let us prove the Weyl inverse correspondence (4.2) quoted in the very beginning of this chapter. For this purpose define $Q = (q' + q'')/2$ and $\widetilde{q} = q'' - q'$, then

$$\langle Q - \widetilde{q}\ /2|\hat{f}|Q + \widetilde{q}\ /2\rangle = \int \frac{dq\,dp}{(2\pi\hbar)}\,\delta(q - Q) \exp(-ip\,\widetilde{q}\ /\hbar) f(q, p).$$

Integrating over q, multiplying both sides with $\exp(i P\,\widetilde{q}\ /\hbar)$ and integrating with respect to $\widetilde{q}$ gives the formula (4.2), written for Q, P in place of q, p.

We can now come back to Eq. (4.12) above. Put $q'' = q'$ in this equation and integrate over q' to give us the result.

4.5 Trace of a Product: Expectation Values

From the formula for the trace (4.11) we know that

$$\mathrm{Tr}(\hat{f}\hat{g}) = \int \frac{dq\,dp}{(2\pi\hbar)}(f * g)(q, p) = \int \frac{dq\,dp}{(2\pi\hbar)}(g * f)(q, p). \quad (4.13)$$

We now show that provided one of the functions, say, g, vanishes along with its derivatives at infinitely large values of q and p, then

$$\int \frac{dq\,dp}{(2\pi\hbar)}(f * g)(q, p) = \int \frac{dq\,dp}{(2\pi\hbar)} f(q, p) g(q, p). \quad (4.14)$$

The proof is based on the star-product formula and using integration by parts repeatedly.

$$\begin{aligned}\int \frac{dq\,dp}{(2\pi\hbar)}(f * g) &= \int \frac{dq\,dp}{(2\pi\hbar)} \sum_{r,s} \frac{(-1)^s}{r!s!} \frac{\partial^{r+s} f}{\partial q^r \partial p^s} \frac{\partial^{r+s} g}{\partial p^r \partial q^s} \left(\frac{i\hbar}{2}\right)^{r+s} \\ &= \int \frac{dq\,dp}{(2\pi\hbar)} \sum_{r,s} \frac{(-1)^s}{r!s!} (-1)^{r+s} \frac{\partial^{2(r+s)} f}{\partial q^{r+s} \partial p^{r+s}}\, g \left(\frac{i\hbar}{2}\right)^{r+s}\end{aligned}$$

$$= \int \frac{dq\,dp}{(2\pi\hbar)} f g + \int \frac{dq\,dp}{(2\pi\hbar)} g \sum_{n=1}^{\infty} \left(-\frac{i\hbar}{2}\right)^n \frac{1}{n!} \frac{\partial^n f}{\partial q^n \partial p^n} \sum_{r=0}^{n} \frac{n!}{r!s!} (-1)^{n-r} = \int \frac{dq\,dp}{(2\pi\hbar)} f g.$$

All terms $n = 1$ onwards drop out because of the factor $(1-1)^n$.

The main use of the trace of product formula is in calculating expectation values. If ρ_ψ is the function corresponding to pure normalized state ψ

$$\langle\psi|\hat{f}|\psi\rangle = \mathrm{Tr}(|\psi\rangle\langle\psi|\hat{f}) = \int \frac{dq\,dp}{(2\pi\hbar)} \rho_\psi(q,p) f(q,p). \tag{4.15}$$

4.6 Eigenvalues

The eigenvalue equation for an observable $\hat{A}\psi_a = a\psi_a$ can be written as

$$\hat{A}|\psi_a\rangle\langle\psi_a| = a|\psi_a\rangle\langle\psi_a| = |\psi_a\rangle\langle\psi_a|\hat{A}.$$

Its counterpart in the phase space is

$$A(q,p) * \rho_a = \rho_a * A(q,p) = a\rho_a.$$

However, the eigenvalue problem in the phase space is neither convenient nor very useful.

4.7 Dynamics and the Moyal Bracket

The Schrödinger equation, written for a normalized state $|\psi\rangle$ in terms of its projection operator $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$ is

$$i\hbar\frac{d\hat{\rho}_\psi}{dt} = i\hbar\frac{d}{dt}(|\psi\rangle\langle\psi|) = \hat{H}|\psi\rangle\langle\psi| - |\psi\rangle\langle\psi|\hat{H} = [\hat{H}, \hat{\rho}_\psi]. \tag{4.16}$$

The projection operator is also called the 'density matrix' corresponding to the 'pure state' $|\psi\rangle$. If $|r\rangle$ is an orthonormal basis then

$$\hat{\rho}_\psi = \sum_{r,s} (\psi_r \psi_s^*)\ |r\rangle\langle s| = \sum_r (|\psi_r|^2)|r\rangle\langle r| + \sum_{r\neq s} (\psi_r\psi_s^*)\ |r\rangle\langle s|.$$

The first term on the extreme right shows the probabilities of ψ to be in the states $|r\rangle$, and the second contains the quantum interference terms. An operator of the form $\rho = \sum p_r |r\rangle\langle r|$ without the interference terms is said to represent a 'mixed' state if p_r are probabilities, and there are more than one term in the sum. For a pure state there is no orthonormal basis in which it can be put in the form of a mixed state.

The density matrix equation (4.16) above is analogous to the Liouville equation of classical probability distribution on the phase space:

$$\frac{d\rho_{\rm cl}}{dt} = \{H, \rho_{\rm cl}\}.$$

For this reason it is called *quantum Liouville equation*. This analogy becomes even more close when $\hat{\rho}_\psi = |\psi\rangle\langle\psi|$ is replaced by its counterpart in the phase space as we see below.

Let $\rho_\psi(q, p)$ be the Wigner function corresponding to the pure state $|\psi\rangle$. Then, the quantum Liouville equation (4.16) can be written in the phase space as

$$\frac{d\rho_\psi}{dt} = \frac{1}{i\hbar}(H * \rho_\psi - \rho_\psi * H) \equiv [H, \rho_\psi]_M.$$

Here we have defined the *Moyal bracket* between two phase space functions as

$$[A, B]_M = \frac{1}{i\hbar}(A * B - B * A) = \frac{2}{\hbar} A(q, p) \sin\left(\frac{\hbar}{2} \overleftrightarrow{P}\right) B(q, p). \tag{4.17}$$

If we were to expand the sine function, the first term is the classical Poisson bracket followed by terms with higher powers of $\hbar$. The quantum Liouville equation in this version looks like a series in powers of $\hbar$ whose limit as $\hbar \to 0$ is the classical Liouville equation.

The above discussion is in the Schrödinger picture as it appears in phase space. If we were to keep ρ_ψ independent of time we can write the equations for any observable f as

$$\frac{df}{dt} = [f, H]_M. \tag{4.18}$$

The equations of motion for q and p are like the classical equations

$$\frac{dq}{dt} = [q, H]_M = \frac{\partial H}{\partial p}, \qquad \frac{dp}{dt} = [p, H]_M = -\frac{\partial H}{\partial q}.$$

Because of this, the phase-space volume (the 'Liouville measure') is preserved by time evolution. But in quantum mechanics, trajectories like these do not make sense. What corresponds to a quantum state is its Wigner function. And that is just a quasi-probability distribution, which cannot be squeezed in both coordinates and momenta

at the same time. Moreover, as the Moyal bracket involves infinitely many differentiations for arbitrary functions, it is not clear that the time evolution for general functions is local.

4.8 Star-Exponential and the Path Integral

The one-to-one correspondence between the Hilbert space operators and phase space functions corresponding to them allows us to construct the exponential evolution operator

$$\exp^*(-itH/\hbar)(q,\,p) = \sum_n \left(\frac{t}{i\hbar}\right)^n \frac{1}{n!}(H * H * \cdots * H)(q,\,p).$$

From the inverse formula (4.2) we can relate it to the matrix element needed for the propagator $\langle q''|\exp(-it\hat{H}/\hbar)|q'\rangle$. Start with

$$\int \frac{dp}{2\pi\hbar}\exp(-ipy/\hbar)\exp^*(-itH/\hbar)(q,\,p)$$

and substitute for the expression for $\exp^*(-itH/\hbar)(q,\,p)$. Then,

$$\begin{aligned}&\int \frac{dp}{2\pi\hbar}\exp(-ipy/\hbar)\exp^*(-itH/\hbar)(q,\,p)\\ &= \int dx \int \frac{dp}{2\pi\hbar}\exp[ip(x-y)/\hbar]\langle q - x/2|\exp(-it\hat{H}/\hbar)|q + x/2\rangle\end{aligned}$$

The integration on p on the right hand side gives the delta function putting $x = y$. Redefining $q'' = q - y/2$ and $q' = q + y/2$,

$$\begin{aligned}&\langle q''|\exp(-it\hat{H}/\hbar)|q'\rangle\\ &\quad= \int \frac{dp}{2\pi\hbar}\exp[ip(q''-q')/\hbar]\exp^*(-itH/\hbar)\left(\frac{q''+q'}{2},\,p\right).\end{aligned} \tag{4.19}$$

The left hand side is related to the path integral of action $S([q(\tau)])$ for all paths $q(\tau)$ with $q(0) = q'$ and $q(t) = q''$:

$$\begin{aligned}&\int d[q]\exp(iS([q])/\hbar)\\ &\quad= \int \frac{dp}{2\pi\hbar}\exp[ip(q''-q')/\hbar]\exp^*(-itH/\hbar)\left(\frac{q''+q'}{2},\,p\right).\end{aligned} \tag{4.20}$$

The usefulness of the formula is dependent, of course, on the feasibility of calculating the star exponential in the phase space.

4.9 Problems

Problem 4.1 Show that the Weyl ordering for q^2p^2 is

$$(q^2p^2)^\wedge = \frac{1}{6}[\hat{q}^2\hat{p}^2 + \hat{q}\hat{p}^2\hat{q} + \hat{q}\hat{p}\hat{q}\hat{p} + \hat{p}\hat{q}\hat{p}\hat{q} + \hat{p}\hat{q}^2\hat{p} + \hat{p}^2\hat{q}^2].$$

By using $[\hat{q}, \hat{p}] = i\hbar$ appropriately, show that we can also write,

$$\begin{aligned}(q^2p^2)^\wedge &= \frac{1}{4}[\hat{q}^2\hat{p}^2 + 2\hat{q}\hat{p}^2\hat{q} + \hat{p}^2\hat{q}^2] \\ &= \frac{1}{4}[\hat{p}^2\hat{q}^2 + 2\hat{p}\hat{q}^2\hat{p} + \hat{q}^2\hat{p}^2].\end{aligned}$$

Hint: Write

$$\delta(u)\delta(v) = \frac{1}{(2\pi)^2}\int dqdp\,\exp[-i(qu + pv)]$$

and differentiate both sides with respect to u and v an appropriate number of times to get the Fourier transform $\widetilde{f}$ of monomials like $f = q^n p^m$. The Weyl ordering formula (4.1) will then involve differentiating $\exp[i(\hat{q}u + \hat{p}v)]$ with respect to u and v a suitable number of times before putting both u and v to zero. Effectively, for $f = q^n p^m$ it will involve looking for the coefficient of $u^n v^m$ in the term $(\hat{q}u + \hat{p}v)^{n+m}/(n + m)!$.

Remark: The Weyl ordering for formula $q^n p^m$ is more convenient in the McCoy [1] form,

$$\begin{aligned}q^n p^m \to (q^n p^m)^\wedge &= \frac{1}{2^n}\sum_{r=0}^{n}\frac{n!}{r!(n-r)!}\hat{q}^{n-r}\hat{p}^m\hat{q}^r \\ &= \frac{1}{2^n}\sum_{r=0}^{m}\frac{m!}{r!(m-r)!}\hat{p}^{m-r}\hat{q}^n\hat{p}^r.\end{aligned}$$

For a proof of the formula see McCoy, Ref. [1]. McCoy had corrected an error in a similar formula suggested by Born and Jordan.

Problem 4.2 For $H(q, p) = (p^2 + q^2)/2$ and $F(q, p) = f(H)$, show that

$$(H * F)(q, p) = HF - \frac{\hbar^2}{4}f'(H) - \frac{\hbar^2}{4}Hf''(H).$$

Problem 4.3 Choose H as in the previous problem and $f(H,t) = \exp^*(-itH/\hbar)$. Show that f satisfies

$$\frac{\partial f}{\partial t} = \frac{1}{i\hbar}\left[Hf - \frac{\hbar^2}{4}f' - \frac{\hbar^2}{4}Hf''\right]$$

and verify that it has the solution

$$f = \frac{1}{\cos(t/2)}\exp\left(\frac{2H}{i\hbar}\tan(t/2)\right).$$

Substitute f in the path integral formula (4.20) and obtain the path integral.

4.10 Notes and References

4.10.1 Weyl Correspondence and Wigner Distribution

The Weyl ordering appears in the classic book "The theory of groups and quantum mechanics" [2]. The quantum mechanics of phase space can be said to have been started by Wigner [3] when he introduced the phase space function described here. The original motivation was to apply it to problems of statistical mechanics in situations where quantum effects can be considered small. As an example, an expansion of the partition function in powers of $\hbar$ can be made to approximate physical quantities. This so-called Wigner–Kirkwood expansion (see Kirkwood [4]) lends itself for symbolic computation through *-product for higher order terms, as in Sharan [5].

4.10.2 Star-Product and Moyal Bracket

Moyal bracket was introduced by Moyal in Ref. [6].

Orderings other than the Weyl ordering can be considered to define phase space quantities. A good reference is the paper by Agarwal and Wolf [7]. The associative property of the star product can be followed, apart from other details, in Mehta [8] and Jordan and Sudarshan [9].

The Moyal bracket when expanded in a series has the first term as the Poisson bracket, and then other terms have powers of $\hbar$. This is called by mathematicians as a deformation of the symplectic structure. An extensive introduction to this 'deformation theory of quantization' can be found in the paper by Bayen, Flato, Fronsdal, Lichnerowicz and Sternheimer [10].

The relation of the path-integral and the star exponential was given in Sharan [11].

References

1. N. McCoy, Proc. Natl. Acad. Sci. (USA) **18**, 674 (1932)
2. H. Weyl, *Theory of Groups and Quantum Mechanics* (Dover Publications, New York, 1950)
3. E.P. Wigner, Phys. Rev. **40**, 749 (1932); A review on various distribution functions is M. Hillery, R.F. O'Connel, M.O. Scully, E.P. Wigner, Phys. Rep. **106**, 121 (1984)
4. J.G. Kirkwood, Phys. Rev. **44**, 31 (1933)
5. P. Sharan, Comput. Phys. Commun. **69**, 235 (1992)
6. J.E. Moyal, Proc. Camb. Philos. Soc. **45**, 99 (1949)
7. G.S. Agarwal, E. Wolf, Phys. Rev. D **2**, 2161 (1970)
8. C.L. Mehta, J. Math. Phys. **5**, 677 (1969)
9. T.F. Jordan, E.C.G. Sudarshan, Rev. Mod. Phys. **33**, 515 (1961)
10. F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, D. Sternheimer, Ann. Phys. **111**, 61–150 (1978)
11. P. Sharan, Phys. Rev. D **20**, 414 (1979)

Chapter 5
Can There Be a Non-linear Quantum Mechanics?

Quantum mechanics can be looked upon as a Hamiltonian theory with linear equations of motion by choosing the real and imaginary parts of the Schrödinger wave function as phase-space variables with expectation value of the quantum Hamiltonian as the classical Hamiltonian. One can ask the question: does there exist a non-linear generalization of this formalism?

5.1 Hamiltonian Equations in Quantum Mechanics

Let us choose an orthonormal basis $|r\rangle$, $r, s = 1, 2, \ldots$ and write $\psi_r = \langle r|\psi\rangle$ as representatives of the state vector $|\psi\rangle$. The Schrödinger equation can be expressed as

$$i\hbar\dot{\psi}_r = \sum_s H_{rs}\psi_s, \qquad r = 1, 2, \ldots \tag{5.1}$$

where $H_{rs} = \langle r|\hat{H}|s\rangle$ is the matrix of the Hamiltonian $\hat{H}$ in this basis.

For a vector ψ, define real quantities q_r, p_r, and ρ_r, θ_r by

$$\psi_r = \frac{(q_r + ip_r)}{\sqrt{2\hbar}} = \rho_r \exp(i\theta_r),$$

and write the expectation value $\langle\psi|H|\psi\rangle$ in terms of q_r's and p_r's as

$$\begin{aligned} H(q, p) &= \sum_{r,s} \psi_r^* H_{rs}\psi_s, \\ &= \frac{1}{2\hbar}\sum_{r,s}\Big[H_{rs}^{+}(q_r q_s + p_r p_s) + H_{rs}^{-}(p_r q_s - p_s q_r)\Big], \end{aligned} \tag{5.2}$$

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_5

$$= \sum_{r,s} \rho_r \rho_s \Big[H^+_{rs} \cos(\theta_r - \theta_s) + H^-_{rs} \sin(\theta_r - \theta_s) \Big], \tag{5.3}$$

where

$$H_{rs} = H^+_{rs} + i H^-_{rs}, \qquad H^+_{rs} = H^+_{sr}, \qquad H^-_{rs} = -H^-_{sr}.$$

The real and imaginary parts of H_{rs} are, as shown, a real symmetric matrix H^+ and a real anti-symmetric matrix H^-.

The Schrödinger equation (5.1) and its complex conjugate can be written as

$$\dot{q}_r = \frac{\partial H}{\partial p_r}, \qquad \dot{p}_r = -\frac{\partial H}{\partial q_r}. \tag{5.4}$$

The norm,

$$\|\psi\|^2 = \sum_r |\psi_r|^2 = \frac{1}{2\hbar} \sum_r [q_r^2 + p_r^2] = \sum_r \rho_r^2$$

is preserved by these equations, as expected. Therefore *all motion is restricted to the surface*

$$\sum_r [q_r^2 + p_r^2] = \text{constant}, \quad \text{or} \quad \sum_r \rho_e^2 = \text{constant}.$$

5.2 Observables and Poisson Bracket

Any observable represented by a Hermitian matrix can be separated into real and imaginary parts:

$$A_{rs} = A^+_{rs} + i A^-_{rs},$$

that is, into real symmetric and real anti-symmetric matrices A^+ and A^- respectively. The expectation value of A has the standard form

$$\begin{aligned} A(q, p) &= (\psi, A\psi) \\ &= \frac{1}{2\hbar} \sum_{r,s} \big[A^+_{rs}(q_r q_s + p_r p_s) + A^-_{rs}(p_r q_s - p_s q_r) \big] \qquad (5.5) \\ &= \sum_{r,s} \rho_r \rho_s \Big[A^+_{rs} \cos(\theta_r - \theta_s) + A^-_{rs} \sin(\theta_r - \theta_s) \Big]. \qquad (5.6) \end{aligned}$$

Note that just as ψ is not directly observable, q_r, p_r are also not observables.

Exercise 5.1 Show that if a normalized ψ corresponds to (q, p) and normalized Ψ to (Q, P) then the transition probability can be written as

$$|\langle\Psi|\psi\rangle|^2 = \frac{1}{(2\hbar)^2}\left[\left(\sum_r (q_r Q_r + p_r P_r)\right)^2 + \left(\sum_r (p_r Q_r - q_r P_r)\right)^2\right] \tag{5.7}$$

$$= \frac{1}{(2\hbar)^2}\sum_{r,s}\big[(q_r q_s + p_r p_s)(Q_r Q_s + P_r P_s) + (p_r q_s - p_s p_r)(P_r Q_s - P_s Q_r)\big]. \tag{5.8}$$

Let A, B be two Hermitian matrices, and $A(q, p)$ and $B(q, p)$ the observables constructed as above from the expectation values, and let

$$[A, B] = i\hbar C.$$

Then the Poisson bracket

$$\{A, B\} = \sum_r \left[\frac{\partial A}{\partial q_r}\frac{\partial B}{\partial p_r} - \frac{\partial A}{\partial p_r}\frac{\partial B}{\partial q_r}\right] \tag{5.9}$$

corresponds to C.

Proof Writing in matrix notation, let the real part (symmetric) and imaginary part (antisymmetric) of matrices A and B be respectively a, b, c, d

$$A = a + ib, \qquad B = c + id,$$

then

$$i\hbar C = [A, B] = ([a, c] - [b, d]) + i([a, d] + [b, c])$$
$$= i[([a, d] + [b, c]) + i([b, d] - [a, c])].$$

This shows that the real and imaginary parts of the matrix C are

$$C = \frac{1}{\hbar}[([a, d] + [b, c]) + i([b, d] - [a, c])].$$

On the other hand, using

$$A(q, p) = \frac{1}{2\hbar}\sum_{r,s}[a_{rs}(q_r q_s + p_r p_s) + b_{rs}(p_r q_s - p_s q_r)],$$

$$B(q, p) = \frac{1}{2\hbar}\sum_{r,s}[c_{rs}(q_r q_s + p_r p_s) + d_{rs}(p_r q_s - p_s q_r)]$$

the Poisson bracket can be calculated

$$\{A(q,\,p),\,B(q,\,p)\} = \frac{1}{\hbar^2}\sum_{r,s}\Big[q_r[a,\,c]_{rs}\,p_s + p_r[b,\,d]_{rs}q_s + q_r(ad-cb)_{rs}q_s + p_r(bc-da)_{rs}\,p_s\Big].$$

The first two terms on the right hand side can be combined as follows: $[a, c]$ and $[b, d]$ are antisymmetric matrices because both a and c are symmetric and both b and d are anti-symmetric. As there is a summation over r and s, only the anti-symmetric part $(p_r q_s - p_s q_r)/2$ survives in the $[b, d]$ term and $(p_s q_r - p_r q_s)/2$ in the $[a, c]$ term. Interchanging dummy indices r, s we get the first two terms as

$$-([a,\,c]-[b,\,d])_{rs}(p_r q_s - p_s q_r)/2.$$

Similarly, in the last two terms,

$$q_r(ad)_{rs}q_s = (ad)_{rs}q_r q_s = (1/2)[(ad)_{rs} + (ad)_{sr}]q_r q_s = (1/2)[a,\,d]_{rs}q_r q_s$$

where in the last step we use the antisymmetry of d and symmetry of a

$$(ad)_{sr}q_r q_s = a_{st}d_{tr}q_r q_s = -d_{rt}a_{ts}q_r q_s.$$

Similarly for the $p_r p_s$ term. Thus,

$$\begin{aligned}\{A(q,\,p),\,B(q,\,p)\} &= \frac{1}{2\hbar^2}[([a,\,d]+[b,\,c])_{rs}(q_r q_s + p_r p_s) + ([b,\,d]-[a,\,c])_{rs}\,p_r q_s]\\ &= C(q,\,p).\end{aligned}$$

Exercise 5.2 Use the definition of the Poisson brackets in the $(q_r,\,p_r)$ coordinates to calculate the bracket in variables ρ_r^2, θ_r as

$$\{A,\,B\} = \frac{1}{\hbar}\sum_r\left(\frac{\partial A}{\partial\rho_r^2}\frac{\partial B}{\partial\theta_r} - \frac{\partial A}{\partial\theta_r}\frac{\partial B}{\partial\rho_r^2}\right) \tag{5.10}$$

and

$$\{\rho_r^2,\,\theta_s\} = \frac{1}{\hbar}\delta_{rs}. \tag{5.11}$$

5.3 Symmetry Transformations

An infinitesimal canonical transformation is generated by an observable $A(q, p)$ through the Poisson bracket:

$$\delta q_t = \epsilon\{q_t, A(q, p)\}, \qquad \delta p_t = \epsilon\{p_t, A(q, p)\}$$

which, for an observable constructed of the standard form (5.5) in terms of $A^{\pm}_{rs}$ is

$$\delta q_t = \frac{\epsilon}{\hbar}\sum_s \left(A^+_{ts} p_s + A^-_{ts} q_s\right)$$
$$\delta p_t = -\frac{\epsilon}{\hbar}\sum_s \left(A^+_{ts} q_s - A^-_{ts} p_s\right)$$

leading to an infinitesimal unitary transformation

$$\psi_t + \delta\psi_t = \psi_t - \frac{i\epsilon}{\hbar}\sum_s (A^+_{ts} + iA^-_{ts})\psi_s = \sum_s \left(\mathbf{1} - \frac{i\epsilon}{\hbar}A\right)_{ts}\psi_s.$$

as expected.

Exercise 5.3 A unitary transformation $\psi' = U\psi$ with U_{rs} as its matrix is given. U_{rs} can be separated into its real and imaginary parts as

$$U_{rs} = u_{rs} + iv_{rs}.$$

Show these matrices satisfy (T denotes transpose)

$$uu^T + vv^T = \mathbf{1} = u^T u + v^T v, \qquad u^T v - v^T u = 0 = uv^T - vu^T,$$

and give rise to a linear canonical transformation:

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \begin{pmatrix} u & -v \\ v & u \end{pmatrix}\begin{pmatrix} q \\ p \end{pmatrix}.$$

All observables, since they correspond to generators of infinitesimal unitary transformations, preserve the norm $\|\psi\|^2$. This shows up as the vanishing of the Poisson bracket of the norm square

$$n = \|\psi\|^2 = \frac{1}{2\hbar}\sum_r (q_r^2 + p_r^2) = \sum_t \rho_t^2$$

with any $A(q, p)$ in the standard form. It can be checked immediately. The bracket

$$\begin{aligned}\{n, A\} &= \sum_{t,rs}\{\rho_t^2, \rho_r\rho_s(A_{rs}^+\cos(\theta_r-\theta_s)+A_{rs}^-\sin(\theta_r-\theta_s))\}\\ &= \frac{1}{\hbar}\sum_{t,rs}\rho_r\rho_s\frac{\partial}{\partial\theta_t}(A_{rs}^+\cos(\theta_r-\theta_s)+A_{rs}^-\sin(\theta_r-\theta_s))\\ &= 0,\end{aligned}$$

because the derivative with respect to θ turns symmetric cosine into antisymmetric sine (and vice versa) causing the sum of a product of symmetric and antisymmetric quantities to zero.

5.4 Eigenvalues

Eigenvector of an observable A in quantum mechanics can be determined as that ψ which is an extremum of its average value $(\psi, A\psi)$ under variation $\psi \to \psi + \delta\psi$ subject to constraint that $\|\psi\|^2$ is kept constant. If we introduce a Lagrange's multiplier λ then the condition of eigenvector amounts to demanding the extremum of

$$(\psi, A\psi) - \lambda(\psi, \psi),$$

under free variation of ψ. Let ψ_0 be one such extremal point then the corresponding eigenvalue is just the value of

$$\left.\frac{(\psi, A\psi)}{(\psi, \psi)}\right|_{\psi_0}.$$

In the classical language we are using here, this can be taken over as the extremal points of

$$A(q, p) - \lambda n(q, p),$$

under the free variations of $q_r, p_r, r = 1, 2, \ldots$. The eigenvalue will correspond to the value of $A(q, p)/n(q, p)$ at the extremal point.

5.5 Non-linear Quantum Mechanics?

What we did in the last few sections is simply to put the ordinary quantum mechanics into a language of classical mechanics. This is a linear Hamiltonian system of infinite number of dimensions in the general case, and it has the following features:

1. The original coordinates q_r, p_r or ρ_r, θ_r do not represent the state completely. The coordinates q_r, p_r themselves are not observables.
2. All observables are quadratic in q, p and of the standard form (5.5) or (5.6). Their eigenvalues can be determined by variational principle applied to A restricted to the surface of constant n, or by diagonalizing the Hermitian matrix $A = A^+ + iA^-$ where

$$A^+_{rs} = \frac{\hbar}{2}\left[\frac{\partial^2 A}{\partial q_r \partial q_s} + \frac{\partial^2 A}{\partial p_r \partial p_s}\right], \qquad A^-_{rs} = \frac{\hbar}{2}\left[\frac{\partial^2 A}{\partial p_r \partial q_s} - \frac{\partial^2 A}{\partial p_s \partial q_r}\right].$$

3. Symmetry transformations are determined by linear canonical transformations. In particular, the equations of motion for observables are linear.

One can naturally ask the question, if quantum theory is equivalent to a classical mechanical system with a phase space, Poisson bracket, and *linear* Hamiltonian equations of motion, could it be that it is a special case of a more general, non-linear theory?

One can ask a counter question: is there a need to look for a general theory? Quantum mechanics is well established with no experiment suggesting any conflict so far despite the fact that its interpretation by measurement theory, projection postulate or collapse of the wave-packet, non-separability etc. are still not well understood.

There are two reasons worth considering for such general theories.

When special relativity was well established, the motivation for searching for a general theory of relativity came from the fact that the gravitational field (the only other classical field theory apart form electrodynamics) could not be accommodated with special theory. The quantum theory is very well adapted to special relativity theory, but has failed so far to accommodate the general relativistic theory of gravitation. Could it be due to the conflict of an essentially linear quantum theory with an essentially non-linear general relativity?

The second reason, due to Weinberg [3], is this: quantum mechanics has been verified excessively, but not been *tested* enough in the sense that there are no alternative theories which can give predictions different from quantum mechanics. A non-linear quantum theory offers a chance: hence the interest.

So, which features of standard quantum theory should we carry over to the general theory? Following Weinberg we make the following assumptions

1. *The non-linear effects are small, therefore the observables have the standard homogeneous quadratic form plus a non-linear term.*
2. *The non-linear term to be added are also homogeneous of second degree in* q_r's *and* p_r's, *such that*

$$\sum_r \left(q_r \frac{\partial A}{\partial q_r} + p_r \frac{\partial A}{\partial p_r}\right) = 2A, \qquad \sum_r \left(p_r \frac{\partial A}{\partial q_r} - q_r \frac{\partial A}{\partial p_r}\right) = 0. \quad (5.12)$$

3. *All symmetry transformations preserve the norm n.*

The condition 2 above, which we call the 'Weinberg condition' ensures that the state ψ and $z\psi$ represent the same physical state with the same average values for all observables for any complex number z. If we treat ψ_r and ψ_r^* as independent variables, the Weinberg condition for $A(q, p) = A(\psi, \psi^*)$ would be written as

$$\sum_r \psi_r \frac{\partial A}{\partial \psi_r} = A = \sum_r \psi_r^* \frac{\partial A}{\partial \psi_r^*}. \tag{5.13}$$

Exercise 5.4 Show that the Weinberg condition can be satisfied in the ρ, θ expression for an observable A if

$$\sum_r \rho_r^2 \frac{\partial A}{\partial \rho_r^2} = A, \qquad \sum_r \frac{\partial A}{\partial \theta_r} = \hbar\{n, A\} = 0. \tag{5.14}$$

In the next section we see how one can implement these requirements and see the non-linear effects in some very simple models.

5.6 Non-linear Terms: Simple Examples

In order to introduce non-linear terms satisfying the Weinberg condition we can simply take the observables in the standard form (5.5) or (5.6) and substitute the constant matrices $A^{\pm}$ into similar (that is, symmetric and anti-symmetric) matrices, but dependent on (q, p) or (ρ, θ). We take some simple examples to illustrate the idea.

5.6.1 Example 1: Extra Energy Level

We take a two-level system in the basis of eigenstates E_1, E_2 of a linear part H_0 of the full Hamiltonian. In the (ρ, θ) coordinates

$$H_0 = E_1\rho_1^2 + E_2\rho_2^2$$

which corresponds to the 2×2 symmetric matrix H_0^+ being diagonal with elements E_1 and E_2, and the antisymmetric matrix being zero. We now add a non-linear part to the symmetric matrix:

$$\begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} + g \begin{pmatrix} 0 & \rho_1\rho_2/(\rho_1^2 + \rho_2^2) \\ \rho_1\rho_2/(\rho_1^2 + \rho_2^2) & 0 \end{pmatrix}$$

so that the Hamiltonian becomes

$$H = E_1\rho_1^2 + E_2\rho_2^2 + 2g\frac{\rho_1^2\rho_2^2}{\rho_1^2+\rho_2^2}.$$

As there is no dependence on θ_1 or θ_2, ρ_1, ρ_2 are constants in time and the θ's grow linearly with t:

$$\dot{\theta}_1 = -\frac{1}{\hbar}\frac{\partial H}{\partial \rho_1^2}, \qquad \dot{\theta}_2 = -\frac{1}{\hbar}\frac{\partial H}{\partial \rho_2^2}$$

or,

$$-\hbar\dot{\theta}_1 = E_1 + 2g\left(\frac{\rho_2^2}{\rho_1^2+\rho_2^2}\right)^2, \qquad -\hbar\dot{\theta}_2 = E_2 + 2g\left(\frac{\rho_1^2}{\rho_1^2+\rho_2^2}\right)^2$$

It is clear that the original eigenstates of H_0 are also the eigenstates of the non-linear Hamiltonian with the same eigenvalues E_1 and E_2 because the additional term is zero if either ρ_1^2 or ρ_2^2 is zero. This can also be checked directly. A normalized state starting with (ρ_1, ρ_2), $\rho_1^2 + \rho_2^2 = 1$ at time $t = 0$ will become

$$\begin{pmatrix}\psi_1(t)\\ \psi_2(t)\end{pmatrix} = \begin{pmatrix}\rho_1 \exp(-it(E_1 + 2g\rho_2^4)/\hbar)\\ \rho_2 \exp(-it(E_2 + 2g\rho_1^4)/\hbar)\end{pmatrix}. \tag{5.15}$$

Therefore, except for the eigenstates corresponding to E_1, E_2 ($\rho_1 = 1, \rho_2 = 0$ or $\rho_2 = 1, \rho_1 = 0$) which have the usual time dependence $\exp(-itE_1/\hbar)$ or $\exp(-it E_2/\hbar)$ respectively, for all other states *the relative phase of the two components is dependent on the initial value of* ρ_1 (or ρ_2).

Let us assume $E_2 > E_1$ without any loss of generality, and write $E_2 - E_1 = \Delta E$. Then, there is an additional eigenvalue provided $2g \geq \Delta E$:

$$E_3 = \frac{E_1 + E_2}{2} + \frac{g}{2} + \frac{(\Delta E)^2}{8g}$$

corresponding to the eigenvector (in the Hilbert space language of ψ_r),

$$\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}\rho_1\\ \rho_2\end{pmatrix}\exp(-iE_3t/\hbar),$$

where

$$\rho_1 = \frac{1}{\sqrt{2}}\left(1 - \frac{\Delta E}{2g}\right)^{1/2}, \qquad \rho_2 = \frac{1}{\sqrt{2}}\left(1 + \frac{\Delta E}{2g}\right)^{1/2}.$$

It should be noted that the extra eigenvalue appears only for $2g > \Delta E$.

5.6.2 Example 2: Asymptotic States

As a second simple example choose

$$H = E_1\rho_1^2 + E_2\rho_2^2 + g(\rho_1^2 - \rho_2^2)(\theta_1 - \theta_2).$$

This corresponds to non-linearity in the θ variables due to choosing the 2×2 anti-symmetric H_{rs}^- with

$$H_{12}^- = g\frac{\rho_1^2 - \rho_2^2}{\rho_1\rho_2}(\theta_1 - \theta_2).$$

We are also allowing the phases to take all values, not being limited to the range $(0, 2\pi)$. In regular quantum mechanics we only require periodic functions of the phases. The equations of motion are

$$\dot{\rho}_1^2 = \frac{1}{\hbar}\frac{\partial H}{\partial\theta_1} = \frac{g}{\hbar}(\rho_1^2 - \rho_2^2)$$
$$\dot{\rho}_2^2 = \frac{1}{\hbar}\frac{\partial H}{\partial\theta_2} = -\frac{g}{\hbar}(\rho_1^2 - \rho_2^2)$$

which shows that $(\rho_1^2 + \rho_2^2)$ remains constant and can be chosen to be equal to 1, and so

$$\rho_1^2(t) = \rho_1^2(0)\exp(2gt/\hbar) + (1 - \exp(2gt/\hbar))/2$$
$$\rho_2^2(t) = \rho_2^2(0)\exp(2gt/\hbar) + (1 - \exp(2gt/\hbar))/2.$$

For θ's

$$\dot{\theta}_1 = -\frac{1}{\hbar}\frac{\partial H}{\partial\rho_1^2} = -\frac{1}{\hbar}[E_1 + g(\theta_1 - \theta_2)]$$
$$\dot{\theta}_2 = -\frac{1}{\hbar}\frac{\partial H}{\partial\rho_2^2} = -\frac{1}{\hbar}[E_2 - g(\theta_1 - \theta_2)].$$

Therefore, (if we write $\Delta\theta = \theta_2 - \theta_1$ and $\Delta E = E_2 - E_1$)

$$(\theta_1 + \theta_2)(t) = (\theta_1 + \theta_2)(0) - t(E_1 + E_2)/\hbar.$$

and

$$\Delta\theta(t) = \Delta\theta(0)\exp(-2gt/\hbar) - \frac{\Delta E}{2g}(1 - \exp(-2gt/\hbar)).$$

As $(\rho_1^2 + \rho_2^2) = 1$, the solution for ρ_1^2 and ρ_2^2 shows that the coupling constant g has to be negative. Therefore all states grow or decay asymptotically to a state with $\rho_1^2 = 1/2, \rho_2^2 = 1/2$.

This is an unphysical case as θ_1 and θ_2 can take arbitrarily large values making energy negative,

5.7 Problems

Problem 5.1 Every Hamiltonian system preserves the volume elements of the phase space under Hamiltonian evolution. In the formalism of Sect. 5.1 verify this statement. How can one handle a volume element for the infinite dimensional Hilbert space? Does it make sense to consider probability distributions on this classical space?

Problem 5.2 The Hermitian observables of quantum mechanics have the standard form (5.5). What will be the expression for the (in general non-Hermitian) product AB where A and B are Hermitian? Will the expression satisfy the Weinberg condition? Show that the expressions satisfy the associative law $(AB)C = A(BC)$, as expected.

Problem 5.3 Verify the expressions in Sect. 5.6.1 for the eigenvalue E_3 and the corresponding eigenvector.

5.8 Notes and References

5.8.1 Non-linear Quantum Mechanics

The fact that quantum mechanics is a linear Hamiltonian theory has been known from the very beginning. Dirac [1] in 1927 wrote the Hamiltonian equations derived from the expectation value of a quantum mechanical operator, identified the canonical variables and quantized once again, thereby achieving a 'second quantization' for the first time, although the term came to be used much later!

The material in the first four sections of this chapter is based on a 1983 unpublished Jamia Millia Islamia preprint of the author and Choudhry [2]. In 1989, Weinberg's "Testing Quantum Mechanics" paper [3] practically exhausted the possibilities of non-linear quantum theory, with experimental bounds. The reasoning for such models goes like this: one of the features of non-linear quantum mechanics is the dependence of the relative phase between two level system (as in the example in Sect. 5.6.1 above) on the initial state, Under fairly general assumptions this would lead to a broadening of the incident absorption or emission frequency of radiation causing transitions between the two levels. A measurement by Bollinger et al. [4] gives an upper limit on the relative size of the non-linear term in the Hamiltonian as less than 10^{-20} eV.

References

1. P.A.M. Dirac, The quantum theory of the emission and absorption of radiation. Proc. R. Soc. A **114**, 243 (1927)
2. P. Sharan, V. Choudhry, Quantum mechanics in the 'classical' language: a step towards a non-linear quantum theory. Jamia Millia Islamia, New Delhi. Preprint (1983)
3. S. Weinberg, Phys. Rev. Lett. **62**, 485 (1989), Testing quantum mechanics. Ann. Phys. **194**, 330 (1989)
4. J.J. Bollinger, J.D. Prestage, S.M. Itano, D.J. Weinland, Phys. Rev. Lett. **63**, 1031 (1989)

Chapter 6
Interaction = Exchange of Quanta

Interaction between a particle with a potential or between two particles illustrates the fundamental fact of quantum theory that exchange of a quantum leads to interaction.

6.1 Non-relativistic 'Potential'

The concept of force in classical theory or quantum theory is the same. A particle changes its linear momentum under the influence of a force.

Let $|\mathbf{p}\rangle$ be the state of a particle with momentum $\mathbf{p}$. If the Hamiltonian is just $\hat{\mathbf{P}}^2/2m$ the momentum is constant in time. Since the Hamiltonian determines the time development, we must include in the Hamiltonian terms which change momentum in order to introduce force. What kind of operators change momentum? Let us define an operator $h(\mathbf{k})$ for a fixed $\mathbf{k}$, acting on the basis $\{|\mathbf{p}\rangle\}$ as

$$\hat{h}(\mathbf{k})|\mathbf{p}\rangle = |\mathbf{p}+\mathbf{k}\rangle.$$

This is not Hermitian, (it is unitary actually), and its adjoint is

$$\hat{h}^\dagger(\mathbf{k})|\mathbf{p}\rangle = |\mathbf{p}-\mathbf{k}\rangle,$$

which shows that $\hat{h}^\dagger(\mathbf{k}) = \hat{h}(-\mathbf{k})$.

Of course $\hat{h}(\mathbf{k})$ gives just a jump, or a kick, to the momentum by a fixed amount $\mathbf{k}$. A general Hamiltonian will include terms for all $\mathbf{k}$. Let $v(\mathbf{k})$ be a complex number representing the strength or amplitude of the force for a kick by amount $\mathbf{k}$. We add to the Hamiltonian the following term:

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_6

$$\begin{aligned}\hat{H}_1 &= \int d^3\mathbf{k}[v(\mathbf{k})\hat{h}(\mathbf{k}) + v^*(\mathbf{k})\hat{h}^\dagger(\mathbf{k})] \\ &= \int d^3\mathbf{k}[v(\mathbf{k})\hat{h}(\mathbf{k}) + v^*(\mathbf{k})\hat{h}(-\mathbf{k})] \\ &= \int d^3\mathbf{k}[v(\mathbf{k}) + v^*(-\mathbf{k})]\,\hat{h}(\mathbf{k}) \\ &\equiv \int d^3\mathbf{k}V(\mathbf{k})\hat{h}(\mathbf{k}).\end{aligned}$$

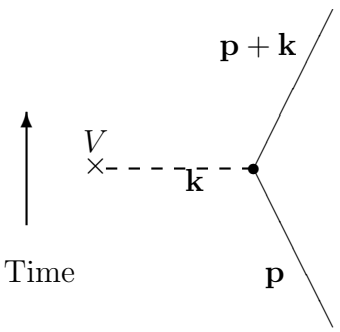
From Chap. 1 we know what the operator $\hat{h}(\mathbf{k})$ is:

$$\hat{h}(\mathbf{k}) = \exp(i\mathbf{k}\cdot\hat{\mathbf{X}}/\hbar),$$

where $\hat{\mathbf{X}}$ is the position operator. Therefore the term to be added to the Hamiltonian in the Schrödinger representation acts on the wave functions as

$$\langle\mathbf{x}|\hat{H}_1|\psi\rangle = \int d^3\mathbf{k}\; V(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{x}/\hbar)\psi(\mathbf{x}) \equiv \mathcal{V}(\mathbf{x})\psi(\mathbf{x}).$$

The function $\mathcal{V}(\mathbf{x})$, is called the *potential* of the force. It is sensible to look at the potential as a classical field which absorbs or emits quanta which can change the momentum of the particle interacting with it. The quanta of the field are not quantized part of the system, but the picture much looks like the following figure:



6.2 The Simplest Model of Quanta Exchange

We take two oscillators with frequencies ω and Ω. One can think of these systems as 'fields' of 'particles' or quanta with energy $\hbar\omega$ and $\hbar\Omega$. Thus, the oscillator with n quanta has energy $n\hbar\omega$ above the ground state and similarly, the second with N quanta has energy $N\hbar\Omega$.

A simple, exactly solvable, toy model of these two interacting oscillators illustrates many features.

6.2.1 The Hamiltonian

Let the two types of 'particles' be called B and A. They are described by operators $b, b^\dagger$ and $a, a^\dagger$, the usual 'ladder' operators of harmonic oscillators with commutation rules

$$[b, b^\dagger] = 1, \qquad [a, a^\dagger] = 1, \qquad [a, b] = 0, \qquad [a, b^\dagger] = 0.$$

The 'free', non-interacting, part of the Hamiltonian is

$$H_0 = \hbar\Omega b^\dagger b + \hbar\omega a^\dagger a, \tag{6.1}$$

omitting the constant ground state energies $\hbar\omega/2$ and $\hbar\Omega/2$. Energy levels of this free system are labeled by two integers N and n

$$E_{N,n} = N\hbar\Omega + n\hbar\omega$$

The interpretation of these levels is that there are n quanta of the type A and N of B.

The total Hamiltonian including the interaction is now chosen as

$$H = H_0 + \hbar\lambda b^\dagger b(a + a^\dagger), \tag{6.2}$$

where λ is a coupling constant of the dimensions of a frequency. The 'vacuum' or zero quanta state of free Hamiltonian is defined by

$$b\Phi_0 = 0, \qquad a\Phi_0 = 0. \tag{6.3}$$

This state also happens to be the eigenstate of the total Hamiltonian:

$$H\Phi_0 = 0,$$

which follows because $a^\dagger$ and b commute.

6.2.2 Bare and Dressed States

The number operators for the two particles are $N_B = b^\dagger b$ and $N_A = a^\dagger a$. We see that

$$[H, N_B] = 0.$$

This means that under time evolution a state with some fixed number of the B quanta retains this number. That is not the case with A quanta because $[H, N_A] \neq 0$. Under an infinitesimal time evolution $\psi \to (1 - i dt\, H/\hbar)\psi$ the interaction term $\lambda b^\dagger b(a + a^\dagger)$ causes annihilation or creation of A-quanta but keeps the number of B-quanta the same. An eigenstate of H with one B-quantum has an indefinite number of A-quanta around it. The B-quantum acquires a 'cloud' of A-quanta around it and we say that this eigenstate of H is a *dressed state*. In contrast, the eigenstates of H_0 are said to contain *bare* B quanta.

6.2.3 Single Dressed B-Particle

We now calculate a single dressed B-particle state. It is a state that contains one bare B quantum and will have any number of A-quanta in the cloud surrounding it. Let Ψ be a one B-particle eigenstate of H,

$$H\Psi = E\Psi.$$

Choose Ψ as a superposition of states with one B-particle, and $0, 1, 2, \ldots$ etc. of A-particles:

$$\Psi = (d_0 + d_1 a^\dagger + d_2 (a^\dagger)^2 + \cdots) b^\dagger \Phi_0.$$

We substitute Ψ and compare the coefficients of various powers of $a^\dagger$:

$$\begin{aligned}(H - E)\Psi = {} & (\hbar\Omega - E)(d_0 + d_1 a^\dagger + d_2 (a^\dagger)^2 + \cdots) b^\dagger \Phi_0 \\ & + \hbar\omega (d_1 a^\dagger + 2 d_2 (a^\dagger)^2 + \cdots) b^\dagger \Phi_0 \\ & + \hbar\lambda (d_1 + 2 d_2 a^\dagger + 3 d_3 (a^\dagger)^2 + \cdots) b^\dagger \Phi_0 \\ & + \hbar\lambda (d_0 a^\dagger + d_1 (a^\dagger)^2 + \cdots) b^\dagger \Phi_0.\end{aligned}$$

This gives us a sequence of equations, the first three of which are

$$(\hbar\Omega - E) d_0 + \hbar\lambda d_1 = 0, \tag{6.4}$$

$$(\hbar\Omega - E) d_1 + \hbar\omega d_1 + \hbar\lambda (2 d_2 + d_0) = 0, \tag{6.5}$$

$$(\hbar\Omega - E) d_2 + \hbar\omega 2 d_2 + \hbar\lambda (3 d_3 + d_1) = 0. \tag{6.6}$$

If we argue from a perturbation theory point of view, then as $\lambda \to 0$, the state Ψ must become the bare one B-particle state proportional to $b^\dagger \Phi_0$. Therefore all d_i except d_0 must go to zero as $\lambda \to 0$. In fact we expect $d_1 = O(\lambda)$, $d_2 = O(\lambda^2)$ and so on.

The first equation of the sequence of equations then tells us that $(\hbar\Omega - E) = O(\lambda^2)$. A look at the second equation tells us that it consists of terms of $O(\lambda^3)$ and $O(\lambda)$, which should separately be equated to zero. Therefore

$$\begin{aligned} \omega d_1 + \lambda d_0 &= 0, \\ (\hbar\Omega - E)d_1 + \hbar\lambda 2 d_2 &= 0. \end{aligned}$$

This fixes

$$d_1 = -\frac{\lambda}{\omega} d_0,$$

which determines the value of $(\hbar\Omega - E)$ from the first Eq. (6.4)

$$E = \hbar(\Omega - \lambda^2/\omega),$$

as well as, from (6.5),

$$d_2 = \frac{1}{2!}\left(-\frac{\lambda}{\omega}\right)^2 d_0.$$

In the third Eq. (6.6) there are second and fourth order terms in λ. The second order terms $\omega 2 d_2 + \lambda d_1$ is identically zero, and the fourth order terms give

$$d_3 = \frac{1}{3!}\left(-\frac{\lambda}{\omega}\right)^3 d_0.$$

The general solution is not difficult to guess:

$$\Psi = d_0 \exp\left[-\frac{\lambda}{\omega} a^\dagger\right] b^\dagger \Phi_0.$$

The normalization constant d_0 can be determined from the condition

$$\begin{aligned} 1 = (\Psi, \Psi) &= |d_0|^2 (\Phi_0, \exp[-\lambda a/\omega] \exp[-\lambda a^\dagger/\omega] b b^\dagger \Phi_0) \\ &= |d_0|^2 \exp[\lambda^2/\omega^2], \end{aligned}$$

where we use the identity

$$\exp(A)\exp(B) = \exp([A, B])\exp(B)\exp(A).$$

Thus the one B-particle dressed state is given by (dropping the over-all phase in Ψ),

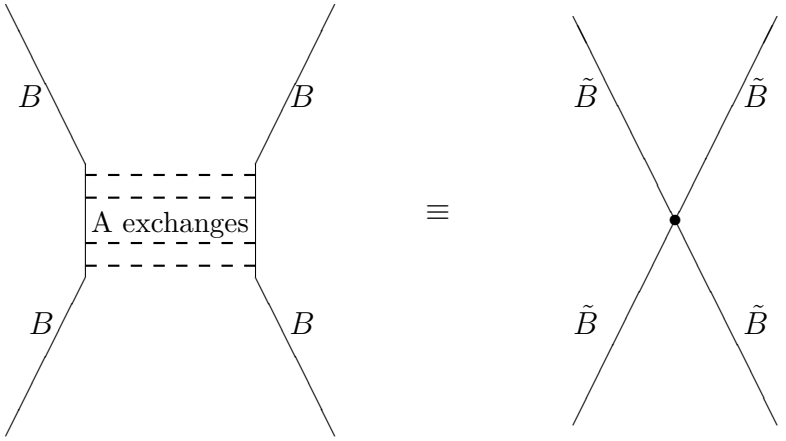
$$\Psi = \exp[-\lambda^2/2\omega^2] \exp[-\lambda a^\dagger/\omega] b^\dagger \Phi_0, \tag{6.7}$$

with eigenvalue

$$E = \hbar(\Omega - \lambda^2/\omega). \tag{6.8}$$

6.2.4 B-B Effective Interaction

Now we take two B-particles. In perturbation theory the second order terms in interaction will involve creation of an A-quantum by one B-particle and its absorption by the other. This back-and-forth exchange of A by B particles leads to an interaction between the B particles. *The A-exchange can be summed or 'integrated out' and replaced by an effective B-B interaction* as shown symbolically in the diagram below.



We notice that the one B-particle dressed state is obtained from the bare one particle state $b^\dagger\Phi_0$ by operating by $\exp[-\lambda^2/2\omega^2]\exp[-\lambda a^\dagger/\omega]$. We do a little manipulation as follows.

$$\begin{aligned}\Psi &= \exp[-\lambda^2/2\omega^2]\exp[-\lambda a^\dagger/\omega]b^\dagger\Phi_0 \\ &= \exp[-\lambda^2/2\omega^2]\exp[-\lambda a^\dagger/\omega]\exp[\lambda a/\omega]b^\dagger\Phi_0\end{aligned}$$

because $\exp[\lambda a/\omega]\Phi_0 = \Phi_0$. Now use

$$\exp(A)\exp(B) = \exp([A, B]/2)\exp(A + B)$$

which holds whenever $[A, B]$ commutes with both A and B to get

$$\Psi = \exp[\lambda(a - a^\dagger)/\omega]b^\dagger\Phi_0$$

The advantage of this manipulation is that the operator $\exp[\lambda(a - a^\dagger)/\omega]$ is unitary.

The two particle dressed-state will similarly involve $\exp[2\lambda(a-a^\dagger)/\omega]$ acting on the two particle bare state because there are two factors of $\exp[\lambda(a-a^\dagger)/\omega]$. Similarly the three particle dressed state will be a similar operator with 2 replaced by 3 in the exponent. Let us define a unitary operator

$$U = \exp\left[\frac{\lambda}{\omega} b^\dagger b(a-a^\dagger)\right]$$

so that it gives the right factor in the exponent because $b^\dagger b$ is the number operator. This unitary operator changes the bare to dressed states. If we apply it to all our observables, then the transformed operators are

$$\begin{aligned}\tilde{a} &= UaU^\dagger \\ &= \exp\left[\frac{\lambda}{\omega} b^\dagger b(a-a^\dagger)\right] a \exp\left[-\frac{\lambda}{\omega} b^\dagger b(a-a^\dagger)\right] \\ &= a + \frac{\lambda}{\omega} b^\dagger b\end{aligned}$$

where we use the identity

$$\exp[S]A\exp[-S] = A + [S,A] + \frac{1}{2!}[S,[S,A]] + \cdots .$$

Similarly

$$\tilde{a}^\dagger = a^\dagger + \frac{\lambda}{\omega} b^\dagger b,$$

and

$$\begin{aligned}\tilde{b} &= \exp\left[\frac{\lambda}{\omega} b^\dagger b(a-a^\dagger)\right] b \exp\left[-\frac{\lambda}{\omega} b^\dagger b(a-a^\dagger)\right] \\ &= b - \frac{\lambda}{\omega}(a-a^\dagger)b + \frac{1}{2!}\left(\frac{\lambda}{\omega}\right)^2 (a-a^\dagger)^2 b - \cdots \\ &= \exp\left[-\frac{\lambda}{\omega}(a-a^\dagger)\right] b\,,\end{aligned}$$

and

$$\tilde{b}^\dagger = b^\dagger \exp\left[\frac{\lambda}{\omega}(a-a^\dagger)\right].$$

We can transform the Hamiltonian in terms of these new operators, by writing the inverted formulas

$$
\begin{aligned}
a &= \tilde{a} - \frac{\lambda}{\omega}\tilde{b}^\dagger\tilde{b},\\
a^\dagger &= \tilde{a}^\dagger - \frac{\lambda}{\omega}\tilde{b}^\dagger\tilde{b},\\
b &= \exp\left[\frac{\lambda}{\omega}(a - a^\dagger)\right]\tilde{b},\\
b^\dagger &= \tilde{b}^\dagger \exp\left[-\frac{\lambda}{\omega}(a - a^\dagger)\right],
\end{aligned}
$$

so that

$$
\begin{aligned}
\frac{1}{\hbar}H &= \Omega b^\dagger b + \omega a^\dagger a + \lambda b^\dagger b(a + a^\dagger)\\
&= \left(\Omega - \frac{\lambda^2}{\omega}\right)\tilde{b}^\dagger\tilde{b} + \omega\tilde{a}^\dagger\tilde{a} - \frac{\lambda^2}{\omega}\tilde{b}^\dagger\tilde{b}^\dagger\tilde{b}\tilde{b}.
\end{aligned}
$$

The Hamiltonian has separated into dressed particles, call them by $\tilde{B}$ created by $\tilde{b}^\dagger$ which interact with themselves by the $\tilde{b}^\dagger\tilde{b}^\dagger\tilde{b}\tilde{b}$ term and a species of free quanta, call them $\tilde{A}$, which are created by $\tilde{a}^\dagger$!

6.3 Problems

Problem 6.1 (*Annihilation and creation operators of a particle in a potential*) The formalism of Sect. 6.1 indirectly uses the annihilation operators $a(\mathbf{k})$ and their adjoint, the creation operators, $a^\dagger(\mathbf{k})$ acting on a background state (ground state or 'vacuum'), $|0\rangle$, and one particle states as follows:

$$
\begin{aligned}
a(\mathbf{k})|0\rangle &= 0, \qquad \text{for all } \mathbf{k},\\
a^\dagger(\mathbf{k})|0\rangle &= |\mathbf{k}\rangle, \qquad \text{for all } \mathbf{k},\\
\left[a(\mathbf{k}), a(\mathbf{k}')\right] = 0, \quad & [a(\mathbf{k}), a^\dagger(\mathbf{k}')] = \delta^3(\mathbf{k} - \mathbf{k}'), \qquad \text{for all } \mathbf{k}, \mathbf{k}'.
\end{aligned}
$$

Write the expression for the term to be added to the Hamiltonian to generate the effect of the classical potential.

Problem 6.2 Derive an expression for amplitudes $V(\mathbf{k})$ in (6.1) for transitions $|\mathbf{p}\rangle \to |\mathbf{p} + \mathbf{k}\rangle$ for the potential (spherical square-well): $\mathcal{V}(r) = V_0$ for $r \le R_0$ and zero for $r > R_0$.

Problem 6.3 (*Project*) Follow the arguments of Sects. 6.2–6.4 for a more realistic model. B is a massive particle, a 'spinless fermion' of rest mass M, and A a much lighter particle of rest mass μ, a 'meson'. The free Hamiltonian is

$$H_0 = Mc^2 \int \mathrm{d}^3\mathbf{p}\, b^\dagger(\mathbf{p}) b(\mathbf{p}) + \int \mathrm{d}^3\mathbf{p}\, E_\mathbf{p}\, a^\dagger(\mathbf{p}) a(\mathbf{p})$$

where $E_\mathbf{p} = c\sqrt{\mathbf{p}^2 + \mu^2 c^2}$. The interaction part of the Hamiltonian is

$$H_I = \lambda \int \mathrm{d}^3\mathbf{p}\mathrm{d}^3\mathbf{k} \left(f(\mathbf{k}) b^\dagger(\mathbf{p}+\mathbf{k}) b(\mathbf{p}) a(\mathbf{k}) + \text{h.c.} \right)$$

where 'h.c.' denotes the Hermitian conjugate of the previous term, λ an appropriate coupling constant, and $f(\mathbf{k})$ the coupling strength at momentum $\mathbf{k}$. The annihilation and creation operators of particles B and A are the standard commutation/anticommutation relations:

$$\begin{aligned}
\left[a(\mathbf{k}), a(\mathbf{k}')\right] &= 0, \quad [a(\mathbf{k}), a^\dagger(\mathbf{k}')] = \delta^3(\mathbf{k}-\mathbf{k}'), \qquad \text{for all } \mathbf{k}, \mathbf{k}', \\
\left[b(\mathbf{k}), b(\mathbf{k}')\right]_+ &= 0, \quad [b(\mathbf{k}), b^\dagger(\mathbf{k}')]_+ = \delta^3(\mathbf{k}-\mathbf{k}'), \qquad \text{for all } \mathbf{k}, \mathbf{k}', \\
\left[a(\mathbf{k}), b(\mathbf{k}')\right] &= 0, \quad [a(\mathbf{k}), b^\dagger(\mathbf{k}')] = 0, \qquad \text{for all } \mathbf{k}, \mathbf{k}'.
\end{aligned}$$

The energy of the particle B is assumed to be independent of its momentum.

Refer to Sect. (12a) of Schweber [1] for details.

6.4 Notes and References

It is a pity that quantum theory which had its beginnings in the emission and absorption of quanta is taught as if non-relativistic quantum mechanics does not require any mention of creation, annihilation (or emission/absorption) of quanta. This is due to the extraordinary emphasis on the wave function in the configuration space with the potential function taken as unchanged classical potential. There is a total disregard for the physically meaningful momentum space. For a proper understanding, both the configuration space and the momentum space are needed. And the so-called 'wave-particle' duality, such a favorite of textbooks, should properly be called and interpreted as 'field-particle' duality.

The model of A and B particles discussed here is a brutally simplified version of a solvable model discussed in Schweber's classic book [1].

Reference

1. S.S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Harper and Row, New York, 1961)

Chapter 7
Proof of Wigner's Theorem

Wigner's theorem which relates a unitary or anti-unitary operator to a symmetry transformation is proved following V. Bargmann's version of the proof.

7.1 Rays and Symmetry Transformation

A *ray* in a Hilbert space $\mathcal{H}$ is the set $\{f\}$ obtained by multiplying a non-zero vector $f \in \mathcal{H}$ with all complex numbers of modulus unity. All vectors in a ray have the same norm. A vector $f \in \{f\}$ is called a *representative* of the ray. A *unit ray* is a ray such that all the vectors belonging to it have unit norm. A unit ray can be obtained by taking a unit vector and obtaining the set of all its multiples by phase factors. The set of all unit rays will be denoted by $\mathcal{R}_0$.

A ray $\{f\}$ multiplied by two different positive numbers $\alpha > 0, \beta > 0, \alpha \neq \beta$ gives two distinct rays.

We cannot talk about *change* in quantities like rays because they cannot be added or subtracted. We say the state is described by a unit ray, but how do we define the rate of change of state unless we can subtract the original quantity from the changed quantity? That is why, although there is a one-to-one correspondence between physical states and unit rays, we must choose a representative vector to work with in practice.

A *symmetry transformation* is a mapping $s : \mathcal{R}_0 \to \mathcal{R}_0$ from unit rays to unit rays such that if $\{f\} \to \{f'\} = s\{f\}$ and $\{g\} \to \{g'\} = s\{g\}$, then for any vectors f, g, f', g' belonging to their respective rays,

$$|(f, g)|^2 = |(f', g')|^2.$$

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9_7

7.2 Wigner's Theorem

For the sake of completeness, let us recall that a unitary operator U is a one-to-one, invertible mapping of the Hilbert space such that for any $f, g \in \mathcal{H}$, $(Uf, Ug) = (f, g)$. An anti-unitary operator A is a one-to-one, invertible mapping of the Hilbert space such that for any $f, g \in \mathcal{H}$, $(Af, Ag) = (f, g)^*$.

Wigner's theorem allows us to work with vectors instead of rays.

Let $s : \mathcal{R}_0 \to \mathcal{R}_0$ be a symmetry and T a vector mapping $T : \mathcal{H} \to \mathcal{H}$. We say that T is *compatible* with symmetry s if T maps vectors in the unit ray $\{f\}$ into vectors in the ray $s\{f\}$ for every unit ray $\{f\} \in \mathcal{R}_0$.

Exercise 7.1 Which of the following vector mappings can be compatible with some symmetry transformation?

(i) The constant linear operator $f \to 2f$ for every $f \in \mathcal{H}$. (No)
(ii) If $\{e_i\}, i = 1, \ldots, \infty$ is an orthonormal basis in $\mathcal{H}$ and $f = \sum_i c_i e_i$ then T is defined by $Tf = \sum_i c_i^* e_i$. (Yes)
(iii) T is a unitary operator. (Yes)
(iv) T is an anti-unitary operator. (Yes)
(v) A projection operator $P_{\mathcal{M}}$ onto a subspace $\mathcal{M} \subset \mathcal{H}$. (No)
(vi) A bounded self-adjoint operator A satisfying $A^2 = 1$. (Yes)
(vii) A bounded linear operator J satisfying $J^2 = -1$. (Yes)

Wigner's theorem is stated as follows:

Every symmetry transformation $s : \mathcal{R}_0 \to \mathcal{R}_0$ determines either a unitary or an anti-unitary operator compatible with it.

The unitary or anti-unitary nature of the operator is determined by the mapping s itself and the operator is determined uniquely except for a factor of modulus unity.

The symmetry transformation which has been defined only on the set of unit rays $\mathcal{R}_0$, can be extended to the set of all rays $\mathcal{R}$ as follows. Let $\{f\} \in \mathcal{R}$. Then $\{f/\|f\|\}$ is a unit ray. Define

$$s\{f\} \equiv \|f\| s\{f/\|f\|\}.$$

7.3 Bargmann Invariant

Let $f_1, f_2, f_3 \in \mathcal{H}$ be three unit vectors. Define a number

$$\Delta(f_1, f_2, f_3) \equiv (f_1, f_2)(f_2, f_3)(f_3, f_1).$$

This is a complex number which actually depends only on the three unit rays $\{f_1\}, \{f_2\}, \{f_3\}$ as can be checked by choosing any other representatives $e^{i\alpha} f_1, e^{i\beta} f_2$, $e^{i\gamma} f_3$ in place of f_1, f_2, f_3. We write therefore

$$\Delta(f_1, f_2, f_3) = \Delta(\{f_1\}, \{f_2\}, \{f_3\}).$$

If s is a symmetry transformation we can calculate

$$\Delta' = \Delta(s\{f_1\}, s\{f_2\}, s\{f_3\}).$$

Given a symmetry mapping s, Δ can be used to test whether the operator determined by the Wigner theorem is a unitary or an anti-unitary operator. If T is the operator compatible with s then for unitary T,

$$(Tf_1, Tf_2) = (f_1, f_2),$$

whereas for anti-unitary T,

$$(Tf_1, Tf_2) = (f_1, f_2)^*.$$

Thus in one case $\Delta' = \Delta$ and in the other $\Delta' = \Delta^*$.

Except for the trivial case of a one-dimensional Hilbert space, there will always be at least three unit rays to check the unitary or anti-unitary nature of a symmetry transformation by the Bargmann invariant.

7.4 A Lemma

This simple result is at the heart of the proof.

Let $e_1, \ldots, e_n \in \mathcal{H}$ be a finite orthonormal set. We call the set of unit rays $\{e_1\}, \ldots, \{e_n\} \in \mathcal{R}_0$ an *orthonormal set of rays*. Although there is no inner product defined for rays, we can still define two rays as being orthogonal.

Let $\{e_1\}, \ldots, \{e_n\} \in \mathcal{R}_0$ be an orthonormal set of rays, and s a symmetry transformation. Let $f \in \mathcal{H}$ be a vector constructed as a linear combination from representatives of the rays of the set:

$$f = \sum_i c_i f_i, \qquad f_i \in \{e_i\}.$$

Note that $f_1, \ldots, f_n$ is also an orthonormal set of vectors and $c_i = (f_i, f)$. The lemma is about the structure of a typical vector f' in the ray $s\{f\}$ mapped by the symmetry transformation.

Lemma
Any representative vector $f' \in s\{f\}$ can be written as

$$f' = \sum_i c_i' f_i'$$

where $f_i' \in s\{e_i\}$ and $|c_i'| = |c_i|,\ i = 1, \ldots, n$.

Proof of the Lemma
The set of rays $s\{e_1\}, \ldots, s\{e_n\} \in \mathcal{R}_0$ is also orthonormal because s being a symmetry transformation preserves modulus of the inner product. Choose any vectors $f_i' \in s\{e_i\}$ and define

$$c_i' = (f_i', f')$$

Then

$$|c_i'| = |(f_i', f')| = |(f_i, f)| = |c_i|.$$

And, as $|(f, f)|^2 = |(f', f')|^2$ because f and f' belong to rays mapped by s,

$$\begin{aligned} \|f' - \sum c_i' f_i'\|^2 &= \|f'\|^2 - \sum |c_i'|^2 \\ &= \|f\|^2 - \sum |c_i|^2 \\ &= 0. \end{aligned}$$

This proves that $f' = \sum_i c_i' f_i'$.

7.5 Proof of the Theorem

Construction of a vector mapping T compatible with the symmetry transformation s is carried out in nine steps.

Step 1
Choose a unit vector $e_0 \in \mathcal{H}$ and let $\{e_0\}$ be the corresponding unit ray. Choose any vector e_0' belonging to $s\{e_0\}$ and define

$$Te_0 = e_0'. \tag{7.1}$$

This defines the operator T on a single vector e_0. The choice of T is obviously arbitrary up to a phase factor.

This is the sole arbitrariness in the definition and construction of T.

Step 2
Let $\mathcal{M}$ be the subspace orthogonal to e_0

$$\mathcal{M} = \{f \in \mathcal{H} | (e_0, f) = 0\}.$$

Choose a unit vector e_1 in $\mathcal{M}$, that is, e_1 is a unit vector orthogonal to e_0. We will now define T on the vector $e_0 + e_1$.

As $\{e_0\}$ and $\{e_1\}$ are orthogonal unit rays, according to the Lemma above any representative of $s(\{e_0 + e_1\})$ is of the form $a_0 e_0' + a_1 e_1''$ where $|a_0| = 1$, $|a_1| = 1$, $e_1'' \in s\{e_1\}$ and where the vector e_0' has already been fixed in step 1 as the representative from $s\{e_0\}$.

Of these vectors $a_0 e_0' + a_1 e_1''$ in $s(\{e_0 + e_1\})$, there is exactly one vector of the form $e_0' + e_1'$, because we can just choose any $a_0 e_0' + a_1 e_1''$ and divide it by a_0. Define

$$T(e_0 + e_1) = e_0' + e_1'. \tag{7.2}$$

This step fixes the vector e_1'.

Step 3

We proceed to define T on a vector of type $e_0 + be_1$ where b is a complex number. T will map $e_0 + be_1$ to a vector in $s\{e_0 + be_1\}$.

A vector $f \in s\{e_0 + be_1\}$ has the form, according to the Lemma, $c_0 e_0' + c_1 e_1'$ with $|c_0| = 1$, $|c_1| = |b|$. Again, among these vectors is a *unique* vector of the form $e_0' + b' e_1'$ with $|b'| = |b|$. Define

$$T(e_0 + be_1) = e_0' + b' e_1' \tag{7.3}$$

We now show that not only $|b'| = |b|$, but their real parts are equal and imaginary parts equal up to a sign: $\mathrm{Re}\, b' = \mathrm{Re}\, b$ and $\mathrm{Im}\, b' = \pm \mathrm{Im}\, b$. This follows from the property of symmetry transformation

$$|1 + b|^2 = |(e_0 + e_1, e_0 + be_1)|^2 = |(e_0' + e_1', e_0' + b' e_1')|^2 = |1 + b'|^2,$$

which implies $\mathrm{Re}\, b = \mathrm{Re}\, b'$. Moreover, as

$$|b|^2 = (\mathrm{Re}\, b)^2 + (\mathrm{Im}\, b)^2 = (\mathrm{Re}\, b')^2 + (\mathrm{Im}\, b')^2 = |b'|^2,$$

we also have $\mathrm{Im}\, b' = \pm \mathrm{Im}\, b$.

For the fixed number b, the choice of b' was uniquely determined by our construction. *Therefore the sign in* $\mathrm{Im}(b') = \pm \mathrm{Im}(b)$ *is also fixed.* Let us write

$$b' = \mathrm{Re}\, b + \epsilon_b \mathrm{Im}\, b,$$

where the sign factor ϵ_b is $+1$ or -1.

Actually ϵ_b does not depend on b. To see this take two different vectors $e_0 + b_1 e_1$ and $e_0 + b_2 e_1$ and determine $e_0' + b_1' e_1'$ and $e_0' + b_2' e_1'$ by the above procedure. We have

$$b_1' = \mathrm{Re}(b_1) + \epsilon_{b_1} \mathrm{Im}(b_1), \qquad b_2' = \mathrm{Re}(b_2) + \epsilon_{b_2} \mathrm{Im}(b_2)$$

Using

$$|(e_0 + b_1 e_1, e_0 + b_2 e_1)|^2 = |(e'_0 + b'_1 e'_1, e'_0 + b'_2 e'_1)|^2$$

gives $|1 + b_1^* b_2|^2 = |1 + {b'_1}^* b'_2|^2$ which is equivalent to

$$\mathrm{Re}(b_1)\mathrm{Re}(b_2) + \mathrm{Im}(b_1)\mathrm{Im}(b_2) = \mathrm{Re}(b'_1)\mathrm{Re}(b'_2) + \mathrm{Im}(b'_1)\mathrm{Im}(b'_2).$$

But $\mathrm{Re}(b_1) = \mathrm{Re}(b'_1)$ and $\mathrm{Re}(b_2) = \mathrm{Re}(b'_2)$ therefore,

$$\mathrm{Im}(b_1)\mathrm{Im}(b_2) = \mathrm{Im}(b'_1)\mathrm{Im}(b'_2) = \epsilon_{b_1}\epsilon_{b_2}\mathrm{Im}(b_1)\mathrm{Im}(b_2).$$

This means

$$\epsilon_{b_1}\epsilon_{b_2} = 1.$$

In other words, either both ϵ_{b_1} and ϵ_{b_2} are $+1$ or both -1. We write the common value of ϵ_b as ϵ_1 where the index 1 is to remind that the choice of b' may still depend on the choice of e_1. From now on we write

$$T(e_0 + be_1) = e'_0 + b'e'_1, \qquad b' = \mathrm{Re}(b) + \epsilon_1 \mathrm{Im}(b). \tag{7.4}$$

Step 4
Let e_2 be another unit vector in $\mathcal{M}$ orthogonal to both e_0 and e_1. By an exactly similar procedure that we used in the last step we can define

$$T(e_0 + ce_2) = e'_0 + c'e'_2, \qquad c' = \mathrm{Re}(c) + \epsilon_2 \mathrm{Im}(c) \tag{7.5}$$

where e'_2 is the unique vector in the ray $s\{e_0 + e_2\}$ of the form $e'_0 + e'_2$.

Step 5
Now consider the three orthonormal rays $\{e_0\}, \{e_1\}, \{e_2\}$ and apply the Lemma to vector $e_0 + e_1 + e_2$. Choose the vector $e'_0 + c_1 e'_1 + c_2 e'_2$ in the ray $s\{e_0 + e_1 + e_2\}$. We must have $|c_1| = 1 = |c_2|$. But as

$$|(e_0 + e_1, e_0 + e_1 + e_2)|^2 = |(e'_0 + e'_1, e'_0 + c_1 e'_1 + c_2 e'_2)|^2$$

therefore $4 = |1 + c_1|^2$. This means $\mathrm{Re}(c_1) = 1$ and so $c_1 = 1$ because $|c_1|$ is already equal to one. Similarly, taking $|(e_0 + e_2, e_0 + e_1 + e_2)|^2$ we can prove $c_2 = 1$.

Therefore,

$$T(e_0 + e_1 + e_2) = e'_0 + e'_1 + e'_2. \tag{7.6}$$

Step 6

We try next to define T on a vector of type $e_0 + be_1 + ce_2$. Choose the unique vector $e'_0 + b''e'_1 + c''e'_2$ in the ray $s\{e_0 + be_1 + ce_2\}$. By taking

$$|(e_0 + e_1, e_0 + be_1 + ce_2)|^2 = |(e'_0 + e'_1, e'_0 + b''e'_1 + c''e'_2)|^2$$

we get using the by-now-familiar argument $b'' = \mathrm{Re}(b) + \epsilon_3 \mathrm{Im}(b)$ where this ϵ_3 is peculiar to the ray $s\{e_0 + be_1 + ce_2\}$.

We will now discuss the equality

$$|(e_0 + be_1, e_0 + be_1 + ce_2)|^2 = |(e'_0 + b'e'_1, e'_0 + b''e'_1 + c''e'_2)|^2 \tag{7.7}$$

with $b' = \mathrm{Re}(b) + \epsilon_1 \mathrm{Im}(b)$ and $b'' = \mathrm{Re}(b) + \epsilon_3 \mathrm{Im}(b)$.

Note that although we know that b'' is determined by the condition that $e'_0 + b''e'_1 + c''e'_2$ is in the ray $s\{e_0 + be_1 + ce_2\}$, b' is determined by the condition that $e'_0 + b'e'_1$ is in the ray $s\{e_0 + be_1\}$. We show below that $b'' = b'$.

We get from the above equality (7.7) of transition probabilities

$$|1 + |b|^2|^2 = |1 + b'^* b''|^2$$

or $\mathrm{Re}(b'^* b'') = |b|^2$. This means $\epsilon_1 \epsilon_3 = 1$ or $\epsilon_3 = \epsilon_1$. Thus $b'' = \mathrm{Re}(b) + \epsilon_3 \mathrm{Im}(b) = b'$.

Similarly considering

$$|(e_0 + e_2, e_0 + be_1 + ce_2)|^2 = |(e'_0 + e'_2, e'_0 + b'e'_1 + c''e'_2)|^2$$

we get $c'' = \mathrm{Re}(c) + \epsilon_4 \mathrm{Im}(c)$ and by

$$|(e_0 + ce_2, e_0 + be_1 + ce_2)|^2 = |(e'_0 + c'e'_2, e'_0 + b''e'_1 + c''e'_2)|^2$$

where $c' = \mathrm{Re}(c) + i\epsilon_2 \mathrm{Im}(c)$ we get again $\epsilon_4 = \epsilon_2$ and $c'' = c'$.

Therefore from $T(e_0 + be_1) = e'_0 + b'e'_1$ and $T(e_0 + ce_2) = e'_0 + c'e'_2$ we infer that the vector $e'_0 + b'e'_1 + c'e'_2$ is in the ray $s\{e_0 + be_1 + ce_2\}$. We therefore define

$$T(e_0 + be_1 + ce_2) = e'_0 + b'e'_1 + c'e'_2 \tag{7.8}$$

$$b' = \mathrm{Re}(b) + i\epsilon_1 \mathrm{Im}(b), \qquad c' = \mathrm{Re}(c) + i\epsilon_2 \mathrm{Im}(c). \tag{7.9}$$

Step 7

We now show that actually $\epsilon_1 = \epsilon_2$.

Consider the unit vector $e_3 = (e_1 + e_2)/\sqrt{2}$. Define $e'_3 = (e'_1 + e'_2)/\sqrt{2}$. We show that e'_3 is in the ray $s\{e_3\}$.

According to the lemma a typical vector in $s\{e_3\}$ is of the form $c_1 e'_1 + c_2 e'_2$ where $|c_1| = |c_2| = 1/\sqrt{2}$. Choose the unique vector $e'_1/\sqrt{2} + c'_2 e'_2$. From $|(e_0 + e_1 +$

$e_2, e_3)|^2 = |(e'_0 + e'_1 + e'_2, e'_1/\sqrt{2} + c'_2 e'_2)|^2$ we get $2 = |1/\sqrt{2} + c'_2|^2$ which implies $c'_2 = 1/\sqrt{2}$.

We already have

$$T(e_0 + e_3) = e'_0 + e'_3. \tag{7.10}$$

Let us choose from $s\{e_0 + de_3\}$ the unique vector $e'_0 + d'e'_3$ with $d' = \mathrm{Re}(d) + i\epsilon_5 \mathrm{Im}(d)$ and define

$$T(e_0 + de_3) = e'_0 + d'e'_3. \tag{7.11}$$

On the other hand

$$T(e_0 + de_3) = T(e_0 + de_1/\sqrt{2} + de_2/\sqrt{2}) = e'_0 + d''e'_1/\sqrt{2} + d'''e'_2/\sqrt{2}$$

with

$$d'' = \mathrm{Re}(d) + i\epsilon_1 \mathrm{Im}(d), \qquad d''' = \mathrm{Re}(d) + i\epsilon_2 \mathrm{Im}(d).$$

The comparison of these two definitions gives

$$\epsilon_1 = \epsilon_5 = \epsilon_2.$$

Since e_1 and e_2 could be any two arbitrary orthogonal unit vectors in $\mathcal{M}$, it follows that *there is a common value ϵ for the whole of $\mathcal{M}$.*

Let us define a function χ of complex numbers

$$\chi(c) = \mathrm{Re}(c) + i\epsilon \mathrm{Im}(c). \tag{7.12}$$

This function which is either $\chi(c) = c$ for $\epsilon = 1$ and $\chi(c) = c^*$ for $\epsilon = -1$, has the obvious properties

$$\chi(c + c') = \chi(c) + \chi(c') \tag{7.13}$$

$$\chi(cc') = \chi(c)\chi(c') \tag{7.14}$$

$$\chi(c^*) = \chi(c)^*, \qquad |\chi(c)| = |c|. \tag{7.15}$$

We can summarize the progress so far by saying that T has been defined on all vectors of the type $e_0 + f$ where $f \in \mathcal{M}$ (recall that $\mathcal{M}$ is the subspace orthogonal to e_0) by the equation

$$T(e_0 + f) = e'_0 + f' \tag{7.16}$$

where the vector f' is uniquely determined. Let us therefore give the definition of T on all vectors of $\mathcal{M}$ by

$$T(f) = f', \tag{7.17}$$

This mapping on vectors of $\mathcal{M}$ satisfies the following rules

$$T(f + f') = T(f) + T(f'), \tag{7.18}$$

$$T(cf) = \chi(c)T(f). \tag{7.19}$$

Step 8
Finally, define T on any vector of $\mathcal{H}$ of the type $ae_0 + f$ where $f \in \mathcal{M}$ by

$$T(ae_0 + f) = \chi(a)e_0' + T(f). \tag{7.20}$$

This completes the construction of T. As defined it satisfies

$$T(f + g) = T(f) + T(g), \tag{7.21}$$

$$T(cf) = \chi(c)T(f). \tag{7.22}$$

If e is a unit vector then $T(e)$ is also a unit vector and for any numbers a and b

$$(T(ae), T(be)) = \chi(a)^*\chi(b) = \chi(a^*b).$$

Since every vector can be expanded in an orthonormal basis from $f = \sum a_i e_i$ and $g = \sum b_j e_j$ we get

$$(T(f), T(g)) = \sum \chi(a_i^* b_i) = \chi((f, g)). \tag{7.23}$$

The vector mapping T is the required compatible unitary or anti-unitary operator, depending on whether $\chi(c) = c$ or $\chi(c) = c^*$.

But we must remember the entire construction was made starting from the initial vector e_0. We must prove that it does not matter which vector e_0 is chosen in the beginning.

7.6 The Final Step

We first prove that if T_1 and T_2 are two vector mappings compatible with the same symmetry transformation then they can differ only by a phase factor. Strictly speaking this holds in Hilbert spaces of dimension greater than one. For one dimensional vector space there can be a unitary as well as an anti-unitary mapping.

Let us first note that if f and g are two linearly independent vectors then the Schwarz inequality is a strict inequality. The equality holds if and only if the vectors are proportional. So

$$|(f, g)|^2 < (f, f)(g, g).$$

This implies for any vector mapping compatible with a symmetry transformation

$$|(T(f), T(g))|^2 < (T(f), T(f))(T(g), T(g)).$$

This shows that if f, g are linearly independent then so are $T(f)$ and $T(g)$.

Now let T_1 and T_2 be two operators both compatible with the same symmetry transformation. As $T_1(f)$ and $T_2(f)$ belong to the same ray they differ by a phase factor. Let the factor for vector f be written as $\omega(f)$.

$$T_1(f) = \omega(f)T_2(f)$$

where $|\omega(f)| = 1$.

We show that $\omega(f)$ actually does not depend on f. Let f and g be two linearly independent vectors. Then

$$T_1(f + g) = \omega(f + g)T_2(f + g) = \omega(f + g)T_2(f) + \omega(f + g)T_2(g)$$

On the other hand

$$T_1(f + g) = T_1(f) + T_1(g) = \omega(f)T_2(f) + \omega(g)T_2(g)$$

Equating the two and using the linear independence of the vectors $T_2(f)$ and $T_2(g)$ we get

$$\omega(f) = \omega(g) = \omega(f + g)$$

This proves the independence of our construction on any particular choice of e_0.

And this completes the proof.

7.7 Problems

In addition to the continuous spacetime symmetries of the Poincare group, discussed in the first chapter, the discrete symmetries like the space-inversion or parity, and time reversal also keep the metric, or the line element of special relativity invariant. Denote by $I_s = -\eta$ and $I_t = \eta$ the 4×4 parity and time reversal matrices. Their product, the space-time inversion $I_s I_t \equiv I = -\mathbf{1}$ is also a symmetry.

$$I_t = -I_s = \eta = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}. \tag{7.24}$$

Their group multiplication laws with Poincare transformations $(a, \Lambda),\ \Lambda \in \mathcal{L}_+^\uparrow$ are

$$\begin{aligned} I_s(a, \Lambda) I_s^{-1} &= (I_s a, I_s \Lambda I_s^{-1}), \\ I_t(a, \Lambda) I_t^{-1} &= (I_t a, I_t \Lambda I_t^{-1}), \\ I(a, \Lambda) I^{-1} &= (-a, \Lambda). \end{aligned} \tag{7.25}$$

Wigner's theorem assigns operators $\mathcal{P},\ \mathcal{T}$ and $\mathcal{I}$ to the symmetries defined by $I_s,\ I_t$ and I respectively. These operators could be unitary or anti-unitary, and undetermined up to a phase factor. Since the squares of all three $I_s,\ I_t$ and I are equal to the identity of the group, (which is denoted by the identity operator on the Hilbert space of states), the operators $\mathcal{P}^2,\ \mathcal{T}^2$ and $\mathcal{I}^2$ must be equal to phase factors ω_s, ω_t and ω_I times the identity operator, respectively. We are free to choose the phase factor of $\mathcal{I}$ so that (as $I = I_s I_t$),

$$\begin{aligned} \mathcal{I} &= \mathcal{P}\mathcal{T} \\ \mathcal{P}^2 &= \omega_s = 1, \qquad \text{chosen by convention} \\ \mathcal{T}^2 &= \omega_t \\ \mathcal{I}^2 &= \omega_I. \end{aligned}$$

The choice $\mathcal{P}^2 = \mathbf{1}$ still leaves the parity operator undetermined up to ± 1 which can only be fixed in specific cases.

Problem 7.1 Show that as $I_s(a, \mathbf{1}) I_s^{-1} = (I_s a, \mathbf{1})$ and $I_t(a, \mathbf{1}) I_t^{-1} = (I_t a, \mathbf{1})$, we must choose $\mathcal{P}$ to be unitary and $\mathcal{T}$ to be anti-unitary to avoid these operators mapping states of positive energy into negative energy states.

Problem 7.2 As $\mathcal{T}$ is anti-unitary, (and so is $\mathcal{I}$) show that phases ω_t and ω_I are real, and therefore equal to ± 1.

Problem 7.3 Derive the multiplication table of the discrete symmetries:

$$\begin{array}{c|ccc} & \mathcal{P} & \mathcal{T} & \mathcal{I} \\ \hline \mathcal{P} & \mathbf{1} & \mathcal{I} & \mathcal{T} \\ \mathcal{T} & \omega_t\omega_I\mathcal{I} & \omega_t\mathbf{1} & \omega_I\mathcal{P} \\ \mathcal{I} & \omega_t\omega_I\mathcal{T} & \omega_t\mathcal{P} & \omega_I\mathbf{1} \end{array} . \tag{7.26}$$

Hint: $\mathcal{P}^{-1} = \mathcal{P}$, $\mathcal{T}^{-1} = \omega_t\mathcal{T}$ and $\mathcal{I}^{-1} = \omega_I\mathcal{I}$.

Problem 7.4 Work out the relations (7.25) for infinitesimal generators P^μ, $\mathbf{J}$ and $\mathbf{K}$ of Poincare group and $\mathcal{P}$ and $\mathcal{T}$.

$$\begin{aligned} \mathcal{P}P^\mu\mathcal{P}^{-1} &= (I_s P)^\mu, \\ \mathcal{P}\mathbf{J}\mathcal{P}^{-1} &= \mathbf{J}, \\ \mathcal{P}\mathbf{K}\mathcal{P}^{-1} &= -\mathbf{K}, \\ \mathcal{T}P^\mu\mathcal{T}^{-1} &= (I_s P)^\mu, \qquad I_s \text{ and not } I_t! \\ \mathcal{T}\mathbf{J}\mathcal{T}^{-1} &= -\mathbf{J}, \\ \mathcal{T}\mathbf{K}\mathcal{T}^{-1} &= \mathbf{K}. \end{aligned}$$

Problem 7.5 (*Project*) Follow the argument for proof of the Wigner theorem as given in Refs. [3, 4] below.

7.8 Notes and References

Wigner's theorem was first proved in his original German book of 1931. The English translation is [1]. We follow the proof of the theorem by Bargmann in Ref. [2]. More recent and elegant proofs where references to other, earlier, proofs can be found is a paper by Simon, Mukunda, Chaturvedi and Srinivasan [3], and its follow-up by Simon, Mukunda, Chaturvedi, Srinivasan and Hamhalter [4]. Among textbooks, Wigner's theorem is discussed in Weinberg [5].

References

1. E.P. Wigner, *Group Theory* (Academic, New York, 1959)
2. V. Bargmann, J. Math. Phys. **5**, 862 (1964)
3. R. Simon, N. Mukunda, S. Chaturvedi, V. Srinivasan, Phys. Lett. A **372**, 6847 (2008)
4. R. Simon, N. Mukunda, S. Chaturvedi, V. Srinivasan, J. Hamhalter, Phys. Lett. A **378**, 2332 (2014)
5. S. Weinberg, *The Quantum Theory of Fields I* (Cambridge, 1995)

Index

P. Sharan, *Some Unusual Topics in Quantum Mechanics*,
SpringerBriefs in Physics,
https://doi.org/10.1007/978-3-030-60418-9