

Techniques and Applications of Path Integration

L. S. SCHULMAN



TECHNIQUES and APPLICATIONS of PATH INTEGRATION

L. S. SCHULMAN

Clarkson University

**DOVER PUBLICATIONS, INC.
MINEOLA, NEW YORK**

Copyright

Copyright © 1981, 2005 by L. S. Schulman
All rights reserved.

Bibliographical Note

This Dover edition, first published in 2005, is an expanded republication of the 1996 printing of the work originally published by John Wiley & Sons, Inc., New York, in 1981. The author has provided a new fifty-six-page supplementary section (beginning on page 361) for the Dover edition.

Library of Congress Cataloging-in-Publication Data

Schulman, L. S. (Lawrence S.), 1941-

Techniques and applications of path integration / L.S. Schulman.
p. cm.

Originally published: New York : J. Wiley, 1981. With new supplementary section.

Includes bibliographical references and index.

ISBN 0-486-44528-3 (pbk.)

1. Path integrals. I. Title.

QC174.17.P27S38 2005

530.12—dc22

2005051788

Manufactured in the United States of America
Dover Publications, Inc., 31 East 2nd Street, Mineola, N.Y. 11501

Preface

This book originated in a course given at the Technion some 10 years ago during my first stay, as a visitor, in Israel. Things were different then. Path integrals were not in the mainstream of anything, and I think those who studied this topic did so more from an aesthetic turn of mind than for practical reasons. Either that, or they still carried forth the ideas of the 1950s when path integration had its great, early successes. My own interest in the subject is accidental—while reading an article in Schwinger's reprint collection on quantum electrodynamics the pages slipped and the book fell open to Feynman's *Reviews of Modern Physics* paper. This I read, and resolved, as a thesis topic, to try to produce a path integral for spin.

Path integration has come a long way in the 1970s. In statistical physics it was the basic framework for the first formulation of the renormalization group transformation. It is used extensively in studying systems with random impurities. In particle physics it is basic to the instanton industry and finds application in studies of gauge field theory (even though some of the methods used had been developed for other problems in the 1960s). In chemical, atomic, and nuclear physics path integrals have been applied to semiclassical approximation schemes for scattering theory. And in rigorous studies of quantum field theory and statistical mechanics the functional integral is used again and again.

This is a book of techniques and applications. My aim is to say what the path integral is and then by example to show how it can and has been used. The approach is that of a physicist with a weakness for but not an addiction to mathematics. The level is such that anyone with a reasonable first course in quantum mechanics should not find difficulty although some of the applications presuppose specialized knowledge; even then, on topics of special interest to me I have supplied background material unrelated to path integrals.

The implications of path integrals for a general understanding of quantum mechanics have been beautifully expounded in Feynman's origi-

nal *Reviews of Modern Physics* paper and in his book on path integrals with Hibbs. For this reason I have touched only lightly on these matters. The Feynman-Hibbs book also includes many applications of path integration, some of which have been given brief treatment here. The emphasis in that volume is on applications developed by Feynman himself, and while they form a considerable body of knowledge there is still enough left over for the present book.

The first part of the book develops the techniques of path integration. Our basic derivation of the path integral presents it as a mathematically justified consequence of the usual quantum mechanics formalism (via the Trotter product formula). Of course we also talk of summing the quantity $\exp(iS/\hbar)$ over all paths, despite the lack of rigorous justification for such terminology. In fact some of our work makes extensive use of this view. Nevertheless, while I have been willing to work without the full blessings of theorems at every step, I have tried to avoid some of the pitfalls that path integrals offer to the unwary. In particular there is a good deal of discussion of the relation $(\Delta \text{distance})^2 \sim (\Delta \text{time})$, a central property of paths entering the Feynman sum over histories. Some of the usual quantum formalism is recovered from the path integral but no great emphasis is placed on this goal. The explicitly solvable path integrals—the harmonic oscillator and variations thereof—are written out, and it is thus shown that the awesome task of summing over paths can in fact occasionally be done. At this early stage we also introduce the Wiener integral, formal first cousin of the path integral and legitimate integral over paths. Here we are able to indulge in an occasional rigorous proof and present a calculation of a first passage time, illustrating the profound connection provided by the Wiener integral between probability and potential theory.

The choice of applications that appear in this book requires a special apology. For a topic to be treated here, I had to first know about it, next understand it (or think I did), then find it amusing, exciting, fundamental, or possessing some similar quality, and finally have the time to present it. There are undoubtedly works that satisfy the third of my criteria but miss out on some other count. Section 32, being a brief treatment of some omissions, reflects the fact that the book had to be finished some time although many beautiful applications would not appear.

As to the applications that do appear.... A lot of space is devoted to the semiclassical approximation. Although the mathematical justification for the stationary phase approximation to the functional integral is not strong, this is an important application, at least in terms of consumer interest. Also, one of the features of Feynman's formulation of quantum mechanics that first impressed me was that the correspondence limit ($\hbar \rightarrow 0$) was a wave of the hand away (via the stationary phase approxima-

tion). Of course converting the hand waving arguments to mathematics is still an uncompleted job, but that does not detract from the beauty of the ideas. I must also confess that I am drawn to the semiclassical approximation not so much by consumer interest but rather by the way in which so many different strands of nineteenth and twentieth century mathematics are brought together. Between Sections 11 and 18 the following topics—all relevant to the matter at hand—are taken up: (1) variational principles of classical mechanics and *minimum* (rather than merely *extremum*) properties of paths—the Jacobi equation; (2) the Morse index theorem; (3) asymptotic analysis, order relations, and so on; (4) Sturm-Liouville theory; (5) Thom's catastrophe theory; (6) *uniform* asymptotic analysis.

Starting from semiclassical results it is not difficult to derive both approximations for scattering theory (Section 19) and a path integral theory of optics (Section 20). The optics calls for some unnatural definitions but I think the reward is worth the temporary inelegance: semiclassical results for path integrals lead at once to geometrical (and even physical) optics with a possibility of getting Keller's “geometrical diffraction” theory too (that possibility is suggested but not carried out in this book).

Probably the most famous early application of path integration is to the polaron and we treat that here too. What makes the polaron special from the standpoint of selling path integrals is that it is one of the few places where the path integral not only helps you discover an answer, but also remains the best way to calculate the answer even after you know it. I like the polaron because it is a tractable field theory; the benefits obtained from using the path integral are entirely analogous to those gotten in quantum electrodynamics, but for the latter all steps are more difficult because of the infinities, the vector character of the field, and gauge problems. Results of the path integral treatment of Q.E.D. are mentioned briefly in Section 32.

Three sections are devoted to the problem of formulating a path integral for spin. Not surprisingly I place the most emphasis on the approach I myself have worked on. To be honest, if I had to solve the problem of a hydrogen atom in a magnetic field I would not use this formalism. Nevertheless, the method shows there is *some* way to treat spin by path integrals. It would also appear that some of the connections to homotopy theory developed in the course of working out a path integral for spin are turning out to be important in gauge theories. Unfortunately, path integral treatments of gauge theories get only the briefest mention in this book; this is one of the gaps I especially regret.

The section on relativistic propagators is both central to the book and an incidental side topic. It is central, because if you wish to think of path

integrals as telling you something fundamental about quantum mechanics then you had better have a relativistic formulation, since that's the way the world is. On the other hand, the most dramatic part of Section 25 deals with particles in the strong gravitational fields near black holes. This application certainly demonstrates the versatility of path integration. For this section some background in general relativity is needed.

Statistical mechanical applications of path integration could easily take up a book on their own. The partition function, the basic object of statistical physics, is most conveniently written as a functional integral for many physical systems. In Sections 26 to 30 both general developments and specific applications are treated. Systems with random impurities are studied although our greatest expositional efforts are not concentrated on the currently most popular developments. The references however cover some of the missing material. The instantons of the 1970s and all their aliases first appeared as critical droplets in path integral studies of metastability in first order phase transitions. Our own treatment takes a neutral view of the physics and presents the method as a way of doing an analytical continuation. The renormalization group and scale transformations have been an important recent application of functional integration and Section 30 deals with this.

The section on coherent states finds itself in the statistical mechanics department almost by accident: one of the applications of this form of the path integral is to the statistical mechanics of boson field theories. Rightfully this could have been put with the section on the phase space path integral with which it has strong ties. As to the section on the phase space path integral, I have included it somewhat reluctantly. Although the reader will soon find that I am not overly fussy about dotting all my mathematical i's, I think that phase space path integrals have more troubles than merely missing details. On this basis they should have been left out of the book; however, I often have conversations with people who use this form of the path integral and they want to know what all the fuss is about. Section 31 aims to answer that question.

The final section gives various applications that I just *couldn't* leave out, and of course, so as to finally finish the book (gestating lo these five years), some topics were left out.

Each section of the book has its own set of references and notes, since I felt that this gave the best opportunity to present background and ancillary material. For convenience however these references are included by author in the index.

There is some previously unpublished work in the book. Much of this occurs in the sections on the semiclassical approximation. The material on caustics was reported in a 1973 conference and was never fully written up, pending its appearance here.

It is a pleasure to thank the many students and colleagues who helped me in the completion of this book. I would like to mention especially Steve Coyne, Gianfausto Dell'Antonio, Mark Kac, Don Lichtenberg, Ady Mann, Rebecca McCraw, David McLaughlin, Charles Newman, Michael Revzen, Roman Shtokhamer, Barry Simon, David Wallace, Arthur Wightman, and Joshua Zak. Although many skilled hands have gone into the typing of the manuscript I am particularly grateful to Judy Huffaker Hammond and Gila Ezion for their help.

L. S. SCHULMAN

*Haifa, Israel
January 1981*

Contents

PART I INTRODUCTION	1
1 Introducing and Defining the Path Integral, 3	
Appendix: The Trotter Product Formula, 9	
2 Probabilities and Probability Amplitudes for Paths, 13	
3 Correspondence Limit for the Path Integral (Heuristic), 16	
Appendix: Useful Integrals, 20	
4 Vector Potentials and Another Proof of the Path Integral Formula, 22	
5 The Ito Integral and Gauge Transformations, 27	
6 Doing the Integral: Free Particle and Quadratic Lagrangians, 31	
Appendix: Exactness of the Sum over Classical Paths, 39	
7 Properties of Green's Functions; the Feynman-Kac Formula, 42	
8 Functional Derivatives and Commutation Relations, 49	
9 Brownian Motion and the Wiener Integral; Kac's Proof, 53	
10 Perturbation Theory and Feynman Diagrams, 65	

**PART II SELECTED APPLICATIONS OF
THE PATH INTEGRAL**

71

- 11 Asymptotic Analysis, 73**
- 12 The Calculus of Variations, 79**
- 13 The WKB Approximation and its Application to
the Anharmonic Oscillator, 92**
- 14 Detailed Presentation of the WKB Approximation, 108**
- 15 WKB Near Caustics, 118**
- 16 Caustics and Uniform Asymptotic Approximations, 131**
- 17 The Phase of the Semiclassical Amplitude, 143**
- 18 The Semiclassical Propagator as a Function of Energy, 147**
- 19 Scattering Theory, 158**
- 20 Geometrical Optics, 164**
- 21 The Polaron, 170**
- 22 Spin and Related Matters, 182**
 - 22.1 The Direct Method-Product Integrals or Time
Ordered Products, 182**
 - 22.2 Continuous Models for Spin, 185**
- 23 Path Integrals for Multiply Connected Spaces, 190**
 - 23.1 Particle Constrained to a Circle, 190**
 - 23.2 Rudiments of Homotopy Theory, 197**
 - 23.3 Homotopy Applied to the Path Integral, 205**
 - 23.4 Extensions of Symmetric Operators, 211**
- 24 Quantum Mechanics on Curved Spaces, 214**

25 Relativistic Propagators and Black Holes,	225
26 Applications to Statistical Mechanics,	237
27 Coherent State Representation,	242
28 Systems with Random Impurities,	259
29 Critical Droplets, Alias Instantons, and Metastability,	271
Appendix: Small Oscillations about the Instanton,	284
30 Renormalization and Scaling for Critical Phenomena,	290
31 Phase Space Path Integral,	303
32 Omissions, Miscellany, and Prejudices,	315
32.1 Field Theory,	315
32.2 Uncompleting the Square,	325
32.3 Rubber: Path Integral Formulation of a Polymer as a Random Walk,	332
32.4 Hard Sphere Gas Second Virial Coefficient,	333
32.5 Adding Paths by Computer,	339
32.6 A Perturbation Expansion Using the Path Integral,	344
32.7 Solvable Path Integral with the Potential $ax^2 + b/x^2$,	344
32.8 Superfluidity,	346
32.9 Fermions,	346
32.10 Books and Review Papers on Path Integrals,	347
Author Index	349
Subject Index	353
Supplements	361

**TECHNIQUES
and APPLICATIONS
of PATH INTEGRATION**

PART ONE

Introduction

ONE

Introducing and Defining the Path Integral

The best place to find out about path integrals is in Feynman's paper.* Our approach is not to use path integrals as a way of arriving at quantum mechanics, although Feynman has used this point of view in his book with Hibbs. Rather we assume knowledge of quantum mechanics and deduce the path integral formalism from it. This gets us into the subject quickly.

The wave function of a nonrelativistic spinless particle in one dimension evolves according to Schrödinger's equation

$$H\psi = i\hbar \frac{\partial \psi}{\partial t} \quad (1.1)$$

$$H = T + V = \frac{1}{2m} p^2 + V = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V \quad (1.2)$$

Our interest is in the propagator or Green's function G which satisfies the equation

$$\left(H - i\hbar \frac{\partial}{\partial t} \right) G(t, t_0) = -i\hbar \delta(t - t_0) \quad (1.3)$$

in operator notation. In coordinate space this is written

$$\left(H_x - i\hbar \frac{\partial}{\partial t} \right) G(x, t; y, t_0) = -i\hbar \delta(x - y) \delta(t - t_0) \quad (1.4)$$

The G 's are related by

$$G(x, t; y, t_0) = \langle x | G(t, t_0) | y \rangle \quad (1.5)$$

*For references see notes at the end of the section.

Knowing G means having a solution to the time dependent Schrödinger equation in the sense that if $\psi(t_0)$ is the state of the system at t_0 , $\psi(t)$, given by

$$\psi(t) = G(t, t_0)\psi(t_0) \quad (1.6)$$

is the state at t . For time independent H an operator solution of (1.3) can immediately be written down:

$$G(t, t_0) = \theta(t - t_0) \exp\left[-\frac{iH(t-t_0)}{\hbar}\right] \quad (1.7)$$

where θ is the step function. Since H is assumed to be time independent we can, without loss of generality, take $t_0=0$. Then for $t>0$ we have

$$G(x, t; y) = \langle x | e^{-iHt/\hbar} | y \rangle \quad (1.8)$$

where the argument 0 has been deleted.

The path integral arises from the fact that

$$e^A = (e^{A/N})^N \quad (1.9)$$

Letting $\lambda = it/\hbar$ yields

$$G(x, t; y) = \langle x | e^{-\lambda(T+V)/N} e^{-\lambda(T+V)/N} \dots e^{-\lambda(T+V)/N} | y \rangle \quad (1.10)$$

with the product in the brackets taken N times. Now we make use of a fundamental fact about the exponential of two operators, namely

$$e^{-\lambda(T+V)/N} = e^{-\lambda T/N} e^{-\lambda V/N} + O\left(\frac{\lambda^2}{N^2}\right) \quad (1.11)$$

This is proved easily enough,* and in a power series expansion the coefficient of the λ^2/N^2 term is

$$A = \frac{1}{2} [V, T]$$

In subsequent manipulations we assume that the $O(1/N^2)$ term is well behaved, that it stays bounded when applied to states, and so on. For reasonable potentials this assumption is justified; more is said on this topic in the appendix.

*An expansion is conveniently generated by looking at derivatives of $\exp(\lambda T/N) \exp(-\lambda(T+V)/N) \exp(\lambda V/N)$.

What we are now aiming for is to replace the term

$$\left[e^{-\lambda(T+V)/N} \right]^N = \left[e^{-\lambda T/N} e^{-\lambda V/N} + O(1/N^2) \right]^N \quad (1.12)$$

by the term

$$\left[e^{-\lambda T/N} e^{-\lambda V/N} \right]^N \quad (1.13)$$

For real numbers (rather than operators) this replacement is a reflection of a fundamental fact about the exponential. The expression

$$\left(1 + \frac{x+y_n}{n} \right)^n$$

converges to e^x despite the presence of y_n so long as $y_n \rightarrow 0$ as $n \rightarrow \infty$. (A proof of this assertion can be had by taking large enough n that $|y_n| < \delta$ and by using the bound

$$\left(1 + \frac{x-\delta}{n} \right)^n < \left(1 + \frac{x+y_n}{n} \right)^n < \left(1 + \frac{x+\delta}{n} \right)^n$$

and assuming n large enough that $|(x-\delta)/n| < 1$.)

For operators a bit of care is required, and the trick is to express the difference of (1.12) and (1.13) in a peculiar way:

$$\begin{aligned} & (e^{-\lambda T/N} e^{-\lambda V/N})^N - (e^{-\lambda(T+V)/N})^N \\ &= [e^{-\lambda T/N} e^{-\lambda V/N} - e^{-\lambda(T+V)/N}] (e^{-\lambda(T+V)/N})^{N-1} \\ &\quad + e^{-\lambda T/N} e^{-\lambda V/N} [e^{-\lambda T/N} e^{-\lambda V/N} - e^{-\lambda(T+V)/N}] e^{-\lambda(T+V)(N-2)/N} \\ &\quad + \cdots + (e^{-\lambda T/N} e^{-\lambda V/N})^{N-1} [e^{-\lambda T/N} e^{-\lambda V/N} - e^{-\lambda(T+V)/N}] \end{aligned} \quad (1.14)$$

Equation 1.14 is an identity. It contains N terms, each of which has the factor $\exp(-\lambda T/N) \exp(-\lambda V/N) - \exp(-\lambda(T+V)/N)$, which by (1.11) is of order $1/N^2$. Hence in the limit the difference is zero. (In an appendix mention is made of various finer points in the estimate.)

We have therefore justified the replacement of (1.10) by

$$G(x, t; y) = \lim_{N \rightarrow \infty} \langle x | (e^{-\lambda T/N} e^{-\lambda V/N})^N | y \rangle \quad (1.15)$$

In effect we have given a heuristic proof of the Trotter product formula. From here, getting the path integral is just a few easy steps. The identity

operator, in the form

$$\int dx_j |x_j\rangle \langle x_j|, \quad j=1, \dots, N-1 \quad (1.16)$$

is inserted between each term in the product in (1.15), yielding

$$G(x, t; y) = \lim_{N \rightarrow \infty} \int dx_1 \cdots dx_{N-1} \prod_{j=0}^{N-1} \langle x_{j+1} | e^{-\lambda T/N} e^{-\lambda V/N} | x_j \rangle \quad (1.17)$$

(for convenience we have taken $y = x_0$, $x = x_N$). The multiplication operator V is diagonal in coordinate space so that

$$\exp\left(-\frac{\lambda V}{N}\right) |x_j\rangle = |x_j\rangle \exp\left(-\frac{\lambda V(x_j)}{N}\right) \quad (1.18)$$

Next we require coordinate space matrix elements of $\exp(-\lambda T/N)$ (between states $\langle \eta |$ and $|\xi \rangle$, say), and to obtain these we insert a complete set of momentum states

$$1 = \int dp |p\rangle \langle p| \quad \text{with} \quad \langle p | \xi \rangle = (2\pi\hbar)^{-1/2} \exp\left(-\frac{ip\xi}{\hbar}\right) \quad (1.19)$$

This gives

$$\begin{aligned} \langle \eta | e^{-\lambda T/N} | \xi \rangle &= \int dp \langle \eta | e^{-\lambda T/N} | p \rangle \langle p | \xi \rangle \\ &= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp e^{-\lambda p^2/2mN} e^{ip(\eta - \xi)/\hbar} \end{aligned} \quad (1.20)$$

This is our first Gaussian integral of the book, but far from the last. The general formula is

$$\int_{-\infty}^{\infty} e^{-ay^2 + by} dy = \sqrt{\frac{\pi}{a}} e^{b^2/4a} \quad (1.21)$$

Using (1.21), (1.20) becomes

$$\langle \eta | e^{-\lambda T/N} | \xi \rangle = \sqrt{\frac{mN}{2\pi\lambda\hbar^2}} e^{-mN(\eta - \xi)^2/2\lambda\hbar^2} \quad (1.22)$$

Equations 1.18 and 1.22 are inserted in (1.17) to yield

$$G(x, t; y) = \lim_{N \rightarrow \infty} \int dx_1 \cdots dx_{N-1} \left(\frac{mN}{2\pi\lambda\hbar^2} \right)^{N/2} \times \prod_{j=0}^{N-1} \exp \left[-\frac{m(x_{j+1}-x_j)^2 N}{2\lambda\hbar^2} - \frac{\lambda V(x_j)}{N} \right] \quad (1.23)$$

Now let $\varepsilon = t/N = \hbar\lambda/iN$ and combine the exponentials in (1.23):

$$G(x, t; y) = \lim_{N \rightarrow \infty} \int dx_1 \cdots dx_{N-1} \left(\frac{m}{2\pi i\hbar\varepsilon} \right)^{N/2} \times \exp \left[\frac{i\varepsilon}{\hbar} \sum_{j=0}^{N-1} \left[\frac{m}{2} \left(\frac{x_{j+1}-x_j}{\varepsilon} \right)^2 - V(x_j) \right] \right] \quad (1.24)$$

Equation 1.24 is the path integral expression for the propagator. A few words are in order, however, on why this is called a “path integral” or “sum over histories.”

Imagine that the points $y, x_1, \dots, x_{N-1}, x$ are connected by lines. Then we have a broken line path from y to x . The sum in the exponential of (1.24) can be interpreted as a Riemann sum of a certain integral along that path:

$$\sum_{j=0}^{N-1} \varepsilon \left[\frac{m}{2} \left(\frac{x_{j+1}-x_j}{\varepsilon} \right)^2 - V(x_j) \right] \sim \int_0^t d\tau \left[\frac{1}{2} m \left(\frac{dx}{d\tau} \right)^2 - V(x) \right] \quad (1.25)$$

The integrand in (1.25) is well known in classical mechanics. It is just the Lagrangian

$$L = \frac{1}{2} m \left(\frac{dx}{dt} \right)^2 - V(x) \quad (1.26)$$

of the classical system, which when quantized has the Hamiltonian (1.2). Furthermore, the action

$$S = \int L dt \quad (1.27)$$

is no less prominent an object in classical mechanics. The argument of the exponential in (1.24) is thus iS/\hbar , with S evaluated along the broken line path connecting $y, x_1, x_2, \dots, x_{N-1}, x$.

The integrals over the quantities $x_1, \dots x_{N-1}$ can be interpreted as summing over all possible broken line paths connecting y and x . Since any continuous path can be approximated by a broken line path, and considering the fact that the limit $N \rightarrow \infty$ is taken, one might be optimistic enough to interpret the integrals as a sum over all paths. A final cosmetic transformation on (1.24) is to let

$$C = \left(\frac{m}{2\pi i\hbar\epsilon} \right)^{N/2} \quad (1.28)$$

and call this a “normalization constant” which, though infinite in the limit $N \rightarrow \infty$, serves merely to make sure G is a unitary operator (more about this later). Equation 1.24 has become

$$G(x, t; y) = C \sum_{\substack{x(\cdot) \\ x(0)=y \\ x(t)=x}} e^{iS[x(\cdot)]/\hbar} \quad (1.29)$$

This is a sum over paths, or histories, of $e^{iS[x(\cdot)]/\hbar}$ with all paths satisfying $x(0)=y$, $x(t)=x$, entering the sum. The capital Σ is used to avoid giving the impression that we have a bona fide measure, and the present state of knowledge on path integration is such that when you encounter expressions like (1.29) they mean neither more nor less than what appears in (1.24)—despite the more suggestive notation some authors prefer to use.

Exercise: Derive a path integral expression for the Green’s function in the case of a time dependent Hamiltonian of the form

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

HINTS: (1) Although (1.7) is no longer true, we still have

$$\begin{aligned} G(t_1, t_0) &= \lim_{\Delta t \rightarrow 0} \exp\left(\frac{\Delta t H(t_1)}{i\hbar}\right) \exp\left(\frac{\Delta t H(t_1 - \Delta t)}{i\hbar}\right) \times \dots \\ &\quad \times \exp\left(\frac{\Delta t H(t_0 + 2\Delta t)}{i\hbar}\right) \exp\left(\frac{\Delta t H(t_0 + \Delta t)}{i\hbar}\right) \end{aligned} \quad (1.30)$$

with $(t_1 - t_0)/\Delta t$ factors. Equation 1.30 is just a time ordered product, about which more later.

(2) It may be useful to change time variables.

APPENDIX: THE TROTTER PRODUCT FORMULA

THEOREM: (Trotter product formula) Let A and B be linear operators on a Banach space X such that A , B and $A+B$ are the infinitesimal generators of the contraction semigroups P^t , Q^t , and R^t respectively. Then for all $\psi \in X$

$$R^t\psi = \lim_{n \rightarrow \infty} (P^{t/n}Q^{t/n})^n \psi$$

In this appendix we define some of the terms used above, indicate how a proof of the theorem goes, and examine some of its consequences.

A semigroup is a set closed under a binary, associative operation. Were inverses required to be in the set it would form a group. A semigroup may or may not possess an identity.

Definition: A contraction semigroup on Banach space X is a family of bounded everywhere defined linear operators P^t , $0 \leq t < \infty$ mapping $X \rightarrow X$ such that

$$\begin{aligned} P^0 &= 1; & P^t P^s &= P^{t+s} & 0 \leq t, s \leq \infty \\ \|P^t\| &\leq 1 & 0 \leq t < \infty \\ \lim_{t \rightarrow 0} P^t \psi &= \psi & \psi \in X \end{aligned}$$

The norm used above is defined as follows:

$$\|Q\| = \inf_{\beta \in B} \beta \quad \text{where } B = \{\beta \mid \|Qx\| \leq \beta \|x\| \text{ for all } x \in X\}$$

and $\|x\|$ is the norm in X . The term “contraction” comes from the fact that $\|P^t\| \leq 1$, since vectors do not grow as they evolve under P^t . The infinitesimal generator A of P^t is defined by

$$A\psi = \lim_{t \rightarrow 0} \frac{1}{t} (P^t \psi - \psi)$$

on the domain $D(A)$ of all $\psi \in X$ for which the limit exists.

Remarks on the Proof of the Theorem (leaving out most statements about domains): let h be a positive real number and let P, Q, R be as defined in the statement of the theorem. By the definition of generators we have

$$(P^h Q^h - 1)\psi = (P^h - 1)\psi + P^h(Q^h - 1)\psi = h(A + B)\psi + o(h)$$

where $\text{o}(h)$ denotes vectors x such that $\lim_{h \rightarrow 0} \|x\|/h = 0$. Then since

$$(R^h - 1)\psi = h(A + B)\psi + \text{o}(h)$$

it follows that

$$(P^h Q^h - R^h)\psi = \text{o}(h).$$

Now we must establish the uniformity of the bound $\text{o}(h)$. By using properties necessarily possessed by infinitesimal generators we show that for ψ in some compact subset of $D(A + B)$, $h^{-1}\|(P^h Q^h - R^h)\psi\|$ is uniformly bounded. For some $\psi \in D(A + B)$, $\{R^s\psi\}$, $0 \leq s \leq t$, is compact and in $D(A + B)$, hence $\|(P^h Q^h - R^h)R^s\psi\| = \text{o}(h)$ uniformly in s .

Let $h = t/n$. Then we wish to show that

$$\|((P^h Q^h)^n - R^{hn})\psi\| \rightarrow 0 \quad \text{for } n \rightarrow \infty$$

To this end we examine

$$\begin{aligned} (P^h Q^h)^n - R^{hn} &= (P^h Q^h - R^h)R^{h(n-1)} + P^h Q^h(P^h Q^h - R^h)R^{h(n-2)} + \dots \\ &\quad + (P^h Q^h)^{n-1}(P^h Q^h - R^h) \end{aligned}$$

Next apply this to ψ , and use the fact that $\|P^h Q^h\| \leq 1$. This implies

$$\|((P^h Q^h)^n - R^{hn})\psi\| \leq \|(P^h Q^h - R^h)R^{h(n-1)}\psi\| + \dots = n\text{o}\left(\frac{t}{n}\right) \rightarrow 0,$$

where the limit is uniform.

For the physical application of this formula we let

$$\begin{aligned} A &= i\Delta, \quad B = -iV \\ -(A + B) &= i(-\Delta + V) = iH \end{aligned}$$

Thus it is necessary to know whether A , B , and $A + B$ generate contractive semigroups. Basically what is involved is examining $\|e^{Ct}\|$ where C is the proposed infinitesimal generator. Recall that

$$\|Q\| = \sup_{\psi \in X} \frac{\|Q\psi\|}{\|\psi\|}$$

For V there is no problem; V is a multiplication operator and

$$\|e^{-iVt}\|^2 = \int dx e^{iV(x)} \psi^*(x) e^{-iV(x)} \psi(x) = \int dx |\psi(x)|^2 = \|\psi\|^2$$

implies $\|e^{-iVt}\| = 1$ for all t .

To show that the Laplacian Δ generates a contractive semigroup we review some elementary facts and intuitions about the norm. Generally speaking the norm looks for the largest eigenvalue. If X is finite (M) dimensional we have $Q\psi_i = q_i \psi_i$, $i = 1, \dots, M$. Then the worst case is the biggest $|q_i|$, call it $|\tilde{q}|$ and its eigenvector $\tilde{\psi}$, for then $\|Q\| = \max_{\psi} \|Q\psi\|/\|\psi\| = \|Q\tilde{\psi}\|/\|\tilde{\psi}\| = |\tilde{q}|$. Say $Q = e^C$. Its eigenvalues are e^{C_i} where C_i are the eigenvalues of C . Then $\|e^C\| = \max_i |e^{C_i}| = \max_i e^{\operatorname{Re} C_i}$. Thus

$$\|e^{tC}\| = \max_i e^{t \operatorname{Re} C_i}$$

so that the condition for C to be the generator of a contractive semigroup is that $\operatorname{Re} C_i \leq 0$ for all i .

On Hilbert space, the thing that would be the eigenvector is not always in the space and the definition of the norm as a limit ("sup") must be invoked. For example, let M be the multiplication operator by the function $\exp(-x^2)$. Thus

$$M\psi(x) = e^{-x^2}\psi(x)$$

Clearly $\|M\psi(x)\| \leq \|\psi\|$ so that $\|M\| \leq 1$. Let

$$\psi = \frac{N}{2} \theta\left(\frac{1}{N} - |x|\right)$$

then

$$\|\psi\|^2 = \frac{N^2}{4} \frac{2}{N} = \frac{N}{2} < \infty$$

For all finite N

$$\frac{\|e^{-x^2}\psi\|^2}{\|\psi\|^2} = \frac{N}{2} \int_{-1/N}^{1/N} e^{-x^2} dx > \frac{N}{2} \int_{-1/N}^{1/N} (1-x^2) dx = 1 - \frac{1}{3N^2}$$

(See 11.20 for proof of the inequality.) But for large enough N this gets arbitrarily close to 1, hence $\|M\| = 1$. However, there is no ψ in the Hilbert space such that $M\psi = \psi$. Thus 1 is not an eigenvalue of M , but it does have some special properties with respect to M : it is in the spectrum. The spectrum is defined as the complement of the resolvent set where the resolvent set of an operator A is the set of λ for which $(\lambda - A)^{-1}$ exists.

Above we had a condition on the eigenvalues of a finite dimensional matrix C so that it generated a contractive semigroup. In a Hilbert space it is most convenient to state the condition in terms of the spectrum. The

condition on an operator C is that $\operatorname{Re} \lambda \leq 0$ for λ in the spectrum. If $C = iK$, in terms of the eigenvalues (or spectrum) k_i of K this means $\operatorname{Im} k_i \geq 0$ for all i . If e^{-tC} is also to be a contractive semigroup—evolution in both directions—we must have

$$\operatorname{Im} k = 0$$

That is, K has only real spectrum, a condition guaranteed by the usual requirement that the Hamiltonian be self-adjoint. Thus $\|e^{-it\Delta}\| = 1$ which is the statement that free particle propagation is norm preserving. To determine whether $A + B$ generates a contractive semigroup we examine whether $-\Delta + V$ is self-adjoint (so that its spectrum would be real). This question is what mathematicians call perturbation theory (and physicists never bother to ask). If it is self-adjoint, then the conditions for the Trotter formula are satisfied.

As shown above the Feynman integral is justified by the foregoing theorem, where by sum over histories is meant the specific way of doing this sum which amounts to the taking of the n th power of (or the n fold iteration of a certain approximation to the Green's function for) a finite operator. This justification of the path integral loses some of the intuitive appeal of the Feynman's formulation.

NOTES

The original paper on path integrals is

R. P. Feynman, *Rev. Mod. Phys.* **20**, 367 (1948)

A general exposition with emphasis on those applications developed by Feynman himself can be found in

R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965

The relation of the Trotter product formula to path integration is lucidly presented in

E. Nelson, *J. Math. Phys.* **5**, 332 (1964)

and our appendix follows that paper.

TWO

Probabilities and Probability Amplitudes for Paths

When you learn quantum mechanics, you're told to forget about your naive, classical idea of particles traveling on trajectories. A particle might be here at one time, and there a bit later, but to speak in terms of the path it took from here to there is to invite contradiction and confusion. The chief parable on which to unlearn your classical intuition is the tale of the two slit experiment. An electron passes through either (or both) of two slits on its way to a screen. The screen is a detector for position and an interference pattern can be seen. Any attempt to verify that the electron went through one slit or the other by a localization of the electron will destroy the interference pattern. How can this lesson, well learned, be consistent with the sum over paths that we have just derived?

The secret lies in the distinction between probability and probability amplitudes. In probability theory there is a rule for conditional probabilities. Let $P(a|b)$ denote the probability of an event a , given that event b occurred, with similar definitions for $P(a|c)$ and $P(b|c)$. Then it is true that

$$P(a|c) = \sum_b P(a|b)P(b|c) \quad (2.1)$$

where the sum is over all events (or states) b that can occur between c and a . In quantum mechanics there is a different rule. We work with a probability amplitude, say φ , which, like $P(a|b)$ depends on two states and

satisfies the relation

$$\varphi_{ac} = \sum_b \varphi_{ab} \varphi_{bc} \quad (2.2)$$

The sum is again over all possible states b . The difference in the two situations is that φ is not a probability, and it is only the square of its absolute value that is interpreted as a probability. This profound difference is discussed at length by Feynman.

Thus although in (2.2) a sum over b is written it would not be correct to say that the system was either in b_1 or b_2 or \dots , since then we ought to be using (2.1) for $|\varphi_{ab}|^2$.

The same caution applies to the path integral. True, we have found a sum over paths, but there is no assertion that the system followed definite paths with certain probabilities. Rather, we compute probability amplitudes for the paths and sum the amplitudes.

Exercise: Analyze the two slit experiment in these terms.

At this point one can turn the entire inquiry around and, starting from (2.2) as a postulate, derive quantum mechanics. Usually in working with (2.2) one is summing over some finite number of spin states, or even some continuum of states. But to get the dynamics of the Schrödinger equation from (2.2) one must try a far more daring sort of sum. Thus we want the amplitude for a (nonrelativistic, spinless) particle to go from y at time 0 to x at time t . We must sum over all intermediate possibilities, that is, our sum must include a contribution for each trajectory from $(y, 0)$ to (x, t) . For the trajectory $x(\tau)$ ($x(0)=y$, $x(t)=x$), call the amplitude $\varphi(x(\tau))$. The total amplitude is

$$\varphi_{(x, t; y, 0)} = \sum_{x(\tau)} \varphi(x(\tau)) \quad (2.3)$$

Exercise: Use this interpretation plus the meaning of the wave function $\psi(x)$ as a probability amplitude for position to show that $\varphi_{(x, t; y, 0)}$ is the Green's function.

The only remaining question is what to take for $\varphi(x(\tau))$. It appears that it was Dirac who put Feynman on the right track for finding this functional. In an early paper and also in later editions of his quantum mechanics book, Dirac found that the Green's function (he calls it "transformation function") looks something like

$$G = e^{iS/\hbar} \quad (2.4)$$

with S the solution of the Hamilton-Jacobi equation. Now everyone knows that $\exp(iS/\hbar)$ is a good approximation for small \hbar ; that's WKB. What Dirac further observed is that $\exp(iS/\hbar)$ is also a good approximation when the time interval over which G is supposed to propagate goes to zero. Consequently, for short times ϵ , the propagator from y to x is approximated by

$$\exp(iS/\hbar) = \exp\left[\frac{i}{\hbar}\epsilon\left(\frac{m}{2}\left(\frac{x-y}{\epsilon}\right)^2 - V\right)\right]$$

Then, when one sums over intermediate positions for finite time propagators (in the spirit of (2.2)) one gets (1.24).

This, I believe, was the main precursor to Feynman's work.

NOTES

The early paper by Dirac is

P. A. M. Dirac, *Physikalische Zeitschrift der Sowjetunion*, 3, No. 1 (1933)

This paper, as well as Feynman's original *Rev. Mod. Phys.* paper can be found reprinted in

J. Schwinger, *Quantum Electrodynamics*, Dover, New York, 1958

Additional material along these lines can be found in later editions of the famous text

P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford, London

In particular see Sections 31 and 32, pp. 121–130 in the fourth (1958) edition.

In the usual formulation of quantum mechanics $\psi(x)$ is the amplitude to find the particle at x where “ x ,” the position, has an operator corresponding to it and one can in principle localize at x . Given $\varphi(x(\tau))$ of (2.3), for example, one might ask if one could in principle determine (i.e., measure) whether a particle has taken a particular trajectory. A way of doing this is presented in

Y. Aharonov and M. Vardi, Meaning of an Individual “Feynman Path,” *Phys. Rev. D* 21, 2235 (1980)

Correspondence Limit for the Path Integral (Heuristic)

We have obtained the propagator in the form of a sum

$$G(x, t; y) = C \sum_{x(\cdot)} \exp\left(\frac{iS(x(\cdot))}{\hbar}\right) \quad (3.1)$$

where $x(\cdot)$ is a path starting at y and ending at x . Starting from (3.1) the correspondence limit of quantum mechanics is a wave of the hand away. The way to get the classical limit of (3.1) is the method of stationary phase and the hand waving consists of assuming that what works for a one dimensional integral works for the “sum” in (3.1).

In the method of stationary phase one considers integrals of the form

$$F(\lambda) = \int_{-\infty}^{\infty} dt \exp[i\lambda f(t)] \quad (3.2)$$

What is sought is the dominant contribution to F as $\lambda \rightarrow \infty$. Later we have occasion to define this goal more precisely, in terms of order symbols and asymptotic expansions. The incantation used for simplifying (3.2) is as follows. For large λ , the phase of $e^{i\lambda f(t)}$ will vary rapidly unless $f' = 0$. This implies that the dominant contribution to the integral comes from regions of t where f' vanishes. Suppose f' vanishes at only one point t_0 . Neglecting contributions to the integral from regions far from t_0 , f is expanded about t_0 ,

$$F(\lambda) = \int dt \exp\left[i\lambda f(t_0) + \frac{1}{2}i\lambda(t-t_0)^2 f''(t_0) + \dots\right] \quad (3.3)$$

If cubic and higher order terms in $(t-t_0)$ are neglected and the integral in

t taken from $-\infty$ to $+\infty$, the result is

$$F(\lambda) = \sqrt{\frac{2\pi i}{\lambda f''(t_0)}} e^{i\lambda f(t_0)} \quad (3.4)$$

These operations can be justified* for sufficiently well behaved functions, f .

What happens when $f''(t_0)$ also vanishes is itself an interesting question and in the context of path integration leads to an examination of caustics in electron optics. This is dealt with later in the book.

Concerning the cubic, quartic, and higher powers of $(t - t_0)$, it is possible to give some heuristic arguments justifying their neglect. Consider the integral

$$K(\lambda) \equiv \int_{-\infty}^{\infty} e^{i\lambda t^2} e^{ia\lambda t^3} e^{ib\lambda t^4} dt \quad (3.5)$$

We claim that λt^3 and λt^4 are small compared to 1 and will use a self consistency argument to demonstrate the point. Supposing they are small, $K(\lambda)$ can be written

$$\begin{aligned} K(\lambda) &= \int_{-\infty}^{\infty} e^{i\lambda t^2} \left[1 + i\lambda at^3 + i\lambda bt^4 - \frac{1}{2}\lambda^2 a^2 t^6 + \dots \right] dt \\ &= \sqrt{\frac{i\pi}{\lambda}} \left[1 - \frac{3ib}{4\lambda} + \frac{i15}{16} \frac{a^2}{\lambda} + \dots \right] \end{aligned} \quad (3.6)$$

*Dropping regions of t in which $f'(t) \neq 0$ can be justified as follows. If $f'(t) \neq 0$ for $\alpha < t < \beta$, then we can make the change of variables $z = f(t)$. Thus

$$F_{\alpha\beta} \equiv \int_{\alpha}^{\beta} \exp(i\lambda f(t)) dt = \int_{f(\alpha)}^{f(\beta)} \exp(i\lambda z) \varphi(z) dz$$

where

$$\varphi(z) = \frac{1}{f'(t)}$$

Assuming that φ is differentiable, do an integration by parts to yield

$$F_{\alpha\beta} = \left(\frac{1}{i\lambda} \right) [\varphi(z) e^{i\lambda z}] \Big|_{f(\alpha)}^{f(\beta)} - \left(\frac{1}{i\lambda} \right) \int_{f(\alpha)}^{f(\beta)} \exp(i\lambda z) \left(\frac{d\varphi}{dz} \right) dz$$

Hence $F_{\alpha\beta}$ goes to zero like $1/\lambda$; from (3.4) regions having $f' = 0$ give contributions that decrease like $1/\sqrt{\lambda}$ and therefore ultimately dominate.

where we have used integration formulas from the appendix to this section. Thus the higher powers of t yield terms that go to zero relative to the first term as $\lambda \rightarrow \infty$. Moreover, each factor t^2 becomes, after integration, $1/\lambda$. In a sense this means that the effective size of t —or range of t contributing to the integral—is $1/\sqrt{\lambda}$. With this way of counting, λt^2 is of order 1, λt^3 of order $\lambda^{-1/2}$, and so forth.*

To summarize, as $\lambda \rightarrow \infty$, the behavior of the integral is determined solely by the point t_0 where $f'(t_0) = 0$. Returning to the sum in (3.1), we recall that the correspondence limit is $\hbar \rightarrow 0$, so we have a parameter $\lambda = 1/\hbar$ multiplying the function $S[x(\cdot)]$. If results about the integral (3.2) mean anything with regard to (3.1), the behavior of the propagator will be dominated by that $\bar{x}(\cdot)$ for which $\delta S/\delta x(\cdot) = 0$. Note that now a functional derivative is in order. But this is precisely the statement that the action is stationary with respect to variation of the path. In this way classical mechanics is recovered and the path $\bar{x}(\cdot)$ is found to satisfy the Euler-Lagrange equations.

It is worth mentioning though that although we have found the classical path to be significant in the $\hbar \rightarrow 0$ limit, what is done with it is not the same as is done in classical mechanics. Specifically it is used in a phase factor, so that if there are two classical paths converging on the same region we can still get interference patterns, in contrast to the purely classical prediction.

Of course, to see interference patterns there must be *something* small compared to \hbar . Specifically, suppose there are two classical paths converging on each of the points ξ in some region. Let the associated classical actions for each ξ be denoted $S_1(\xi)$ and $S_2(\xi)$. The propagator then has the general form

$$G = e^{iS_1/\hbar} + e^{iS_2/\hbar} \quad (3.7)$$

(actually these terms have real coefficients, which may not be equal). Since the probability distribution is given by $|G|^2$, an interference pattern arises from the effect of the cross term in the product GG^* . The cross term is the real part of

$$e^{i[S_2(\xi) - S_1(\xi)]/\hbar}$$

Therefore, if the variation in $S_2 - S_1$ is large compared to \hbar , the pattern will be destroyed by the rapid oscillation of this factor.

*The elimination of cubic and higher order terms can also be accomplished through a change of variable [in (3.2)] $f(t) - f(t_0) = z^2$, where $f'(t_0) = 0$. This method, which I find less graphic than the expansion just outlined, can be more easily converted into a rigorous proof.

Exercise: Suppose a wave function ψ has expectation values for position and momentum x_1 and p_1 , and is sharply concentrated near these values, for example, in a minimum uncertainty wave packet. Let the system evolve under the Hamiltonian $H=p^2/2m+V$ from time t_1 to time t_2 . Show that for $\hbar \rightarrow 0$, at time t_2 the wave function will be concentrated at x_2, p_2 where these are the values of position and momentum the classical system would have reached at t_2 , starting at x_1, p_1 at t_1 .

Solution (very heuristic): Take

$$\psi_{x_i, p_i}(x) = (2\pi\Delta^2)^{-1/4} \exp\left[-\frac{(x-x_i)^2}{4\Delta^2} + \frac{ip_i x}{\hbar}\right], \quad i=1,2 \quad (3.8)$$

What's needed is the propagator for a time interval $T=t_2-t_1$ and we guess (an educated guess) by analogy with (3.4) that it will take the form

$$G = Ce^{iS_c/\hbar} \quad (3.9)$$

with S_c the value of the action along the path \bar{x} for which $\delta S=0$ and with the appropriate boundary conditions. Moreover, C should be related to the second derivative of S but its exact form is not needed for the purpose of this exercise. What must be shown is that the amplitude

$$A = \langle x_2 p_2 | e^{-iHT/\hbar} | x_1 p_1 \rangle$$

is zero except when x_2, p_2 is the position in phase space at time $t=t_1+T$ of a particle leaving x_1, p_1 at $t=t_1$. Then from (3.8) and (3.9)

$$\begin{aligned} A &= \int d\xi_2 d\xi_1 \psi_{x_2, p_2}^*(\xi_2) G(\xi_2, T; \xi_1) \psi_{x_1, p_1}(\xi_1) \\ &= (2\pi\Delta^2)^{-1/2} \int d\xi_1 d\xi_2 C \exp\left[-\frac{(\xi_2-x_2)^2}{4\Delta^2} - \frac{ip_2\xi_2}{\hbar}\right. \\ &\quad \left. + \frac{i}{\hbar} S_c(\xi_2, T; \xi_1) - \frac{(\xi_1-x_1)^2}{4\Delta^2} + \frac{ip_1\xi_1}{\hbar}\right] \end{aligned}$$

For uncertainties in x and p to be about the same size it is convenient to take $\Delta^2=O(\hbar)$. For the major contribution to the integral we look to the regions where $\partial/\partial\xi_1$ and $\partial/\partial\xi_2$ of the argument of the exponent vanish. Keeping ξ_1 and ξ_2 real, there are the separate requirements that

$$\xi_2 = x_2, \quad \xi_1 = x_1$$

and

$$-p_2 + \frac{\partial S_c}{\partial \xi_2} = 0, \quad p_1 + \frac{\partial S_c}{\partial \xi_1} = 0 \quad (3.10)$$

This says that (p_2, x_2) and (p_1, x_1) are related to each other by the canonical transformation generated by the function S_c . But S_c generates (classically) time evolution, and we have obtained the condition that the particle stick to the classical trajectory. Just how closely it sticks is also implicit in these integrals.

Closely related to the method of stationary phase is Laplace's method, which deals with integrals of the form

$$G(\lambda) = \int_{-\infty}^{\infty} dt \exp(-\lambda g(x)) \quad (3.11)$$

with λ and g real. For g smooth and bounded from below points where $g'(x)=0, g''(x)>0$ give the greatest contribution. Because of the smallness of the exponential for large λ and g away from its minimum it is easier to prove asymptotic properties for G of (3.11) than for the F of (3.2) as $\lambda \rightarrow \infty$. Laplace's method is dealt with in greater detail in Section 11.

APPENDIX: USEFUL INTEGRALS

$$\int_{-\infty}^{\infty} e^{-ay^2} dy = \sqrt{\frac{\pi}{a}} \quad (3.12)$$

$$\int_{-\infty}^{\infty} y^2 e^{-ay^2} dy = \sqrt{\frac{\pi}{a}} \left(\frac{1}{2a} \right) \quad (3.13)$$

$$\int_{-\infty}^{\infty} y^4 e^{-ay^2} dy = \sqrt{\frac{\pi}{a}} \left(\frac{3}{4a^2} \right) \quad (3.14)$$

$$\int_{-\infty}^{\infty} y^6 e^{-ay^2} dy = \sqrt{\frac{\pi}{a}} \left(\frac{15}{8a^3} \right) \quad (3.15)$$

$$\int_{-\infty}^{\infty} e^{-ay^2+by} dy = \sqrt{\frac{\pi}{a}} e^{b^2/4a} \quad (3.16)$$

$$\int_{-\infty}^{\infty} y e^{-ay^2+by} dy = \sqrt{\frac{\pi}{a}} \left(\frac{b}{2a} \right) e^{b^2/4a} \quad (3.17)$$

$$\int_{-\infty}^{\infty} y^2 e^{-ay^2+by} dy = \sqrt{\frac{\pi}{a}} \left(\frac{1}{2a} \right) \left(1 + \frac{b^2}{2a} \right) e^{b^2/4a} \quad (3.18)$$

For $\operatorname{Re} \alpha$ not positive these formulas can be interpreted by analytic continuation, the only complication being a branch point at $\alpha=0$.

NOTES

A standard source on asymptotic approximations is

A. Erdelyi, *Asymptotic Expansions*, Dover, New York, 1956

Additional references are given later in the book when we take up the subject of stationary phase and other asymptotic approximations in a more systematic way. At that stage we also discuss the extent to which the arguments given here can be rigorously applied to path integrals. The problem posed in the exercise of this section has been rigorously solved using methods closely related to the path integral, namely the Trotter product formula. See

G. A. Hagedorn, Semiclassical Quantum Mechanics I: The $\hbar \rightarrow 0$ Limit for Coherent States, *Commun. Math. Phys.* **71**, 77 (1980)

FOUR

Vector Potentials and Another Proof of the Path Integral Formula

In this section we extend the path integral to situations where a magnetic field is present. The extension has a surprising subtlety and gives us our first inkling of the intimate relation of path integrals to the theory of Brownian motion.

First a generalization to three dimensions is needed. Nothing more strenuous is demanded for this generalization than making boldface all the position variables in (1.24). Each integral $d\mathbf{x}_j$ is three dimensional and correspondingly the normalization factor becomes $(m/2\pi i\hbar)^{3N/2}$.

For convenience in manipulation, and as a concession to popular usage, we shall also rewrite the formal path integral of (1.29) as follows

$$G(\mathbf{x}, t; \mathbf{y}) = \int_{\mathbf{y}, 0}^{\mathbf{x}, t} d\mathbf{x}(\tau) \exp\left(\frac{iS[\mathbf{x}(\tau)]}{\hbar}\right) \quad (4.1)$$

In the absence of magnetic fields we have already ascertained that S has the form

$$S[\mathbf{x}(\tau)] = \int_0^t L\left(\mathbf{x}, \frac{d\mathbf{x}}{d\tau}\right) d\tau, \quad L = \frac{1}{2}m\left(\frac{d\mathbf{x}}{d\tau}\right)^2 - V(\mathbf{x}) \quad (4.2)$$

It will come as no surprise that in the presence of a magnetic field \mathbf{B} , derivable from a vector potential \mathbf{A} ($\mathbf{B} = \nabla \times \mathbf{A}$), the only change in (4.1) and (4.2) is that L becomes

$$L = \frac{1}{2}m\left(\frac{d\mathbf{x}}{d\tau}\right)^2 + \frac{e}{c} \frac{d\mathbf{x}}{d\tau} \cdot \mathbf{A} - V(\mathbf{x}) \quad (4.3)$$

In the action there will appear the additional term

$$\frac{e}{c} \int_0^t d\tau \frac{d\mathbf{x}}{d\tau} \cdot \mathbf{A} \quad (4.4)$$

which in any sensible theory should equal

$$\frac{e}{c} \int d\mathbf{x} \cdot \mathbf{A} \quad (4.5)$$

The way to check whether the proposed propagator with the Lagrangian (4.3) is correct is to look at the expression for G as a limit

$$G(\mathbf{x}, t; \mathbf{y}) = \lim_{N \rightarrow \infty} \int d^3x_1 \cdots d^3x_N \left(\frac{m}{2\pi i \hbar \epsilon} \right)^{3(N+1)/2} \times \exp \left\{ \frac{ie}{\hbar} \sum_{j=0}^N \left[\frac{m}{2} \left(\frac{\mathbf{x}_{j+1} - \mathbf{x}_j}{\epsilon} \right)^2 - V(\mathbf{x}_j) \right] + \text{vector potential term} \right\},$$

$$\epsilon = \frac{t}{N+1} \quad (4.6)$$

(In (4.6) N intermediate \mathbf{x} 's are taken, rather than $N-1$ as in (1.24).) From (4.5) it follows that the vector potential contribution is

$$\frac{ie}{\hbar c} \sum_{j=0}^N (\mathbf{x}_{j+1} - \mathbf{x}_j) \cdot \mathbf{A}(\mathbf{x}) \quad (4.7)$$

but it is not clear whether $\mathbf{A}(\mathbf{x})$ should be evaluated at \mathbf{x}_j , \mathbf{x}_{j+1} or somewhere in between, such as $\frac{1}{2}(\mathbf{x}_j + \mathbf{x}_{j+1})$. In fact, \mathbf{A} must be evaluated at the midpoint; either that or one must use $\frac{1}{2}[\mathbf{A}(\mathbf{x}_j) + \mathbf{A}(\mathbf{x}_{j+1})]$. We now prove that with the midpoint choice path integral theory gives the same result as Schrödinger's equation, and later come back to show what would have happened had the wrong choice been made. In passing it will become evident why in (1.24) no fuss was made about the fact that x_j , rather than x_{j+1} appeared as the argument of $V(x)$.

The way we check the validity of this method for handling magnetic fields is the same way Feynman first verified the entire formalism. Namely, take a wave function at time T , $\psi(\mathbf{y}, T)$, propagate it to time $T+\epsilon$ by means of the putative path integral propagator, and see if this evolution is the same as that given by the Schrödinger equation. For this purpose we take as propagator the quantity under the limit operation in (4.6) for the case $N=0$ (i.e., for a single infinitesimal time step) and include the vector

potential ($A(\frac{1}{2}\mathbf{x}_j + \frac{1}{2}\mathbf{x}_{j+1})$) contribution. This will be used to propagate ψ for a time ϵ and we shall eventually let ϵ go to zero to allow comparison with the $\partial/\partial t$ operation of Schrödinger's equation. If the theory is correct we should therefore have

$$\begin{aligned} \psi(\mathbf{x}, T+\epsilon) = & \int d^3y \left(\frac{m}{2\pi i\hbar\epsilon} \right)^{3/2} \exp \left\{ \frac{ie}{\hbar} \left[\frac{m}{2} \left(\frac{\mathbf{x}-\mathbf{y}}{\epsilon} \right)^2 - V(\mathbf{y}) \right] \right. \\ & \left. + \frac{ie}{\hbar c} (\mathbf{x}-\mathbf{y}) \cdot \mathbf{A} \left(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y} \right) \right\} \psi(\mathbf{y}, T) \end{aligned} \quad (4.8)$$

The next few steps involve a small nightmare of Taylor expansions and Gaussian integrals. Define the variable

$$\xi = \mathbf{y} - \mathbf{x}$$

Let this be the variable of integration in (4.8) and expand $\psi(\mathbf{y}, T)$, \mathbf{A} , and V about $\xi=0$. Equation 4.8 becomes

$$\begin{aligned} \psi(\mathbf{x}, T+\epsilon) = & \left(\frac{m}{2\pi i\hbar\epsilon} \right)^{3/2} \int d^3\xi \exp \left[\frac{im\xi^2}{2\epsilon\hbar} \right] \\ & \times \exp \left\{ -\frac{ie}{\hbar} V(\mathbf{x}) - \frac{ie}{\hbar} \nabla V(\mathbf{x}) \cdot \xi + \dots \right\} \\ & \times \exp \left\{ -\frac{ie}{\hbar c} \xi \cdot \left[\mathbf{A}(\mathbf{x}) + \frac{1}{2}(\xi \cdot \nabla) \mathbf{A}(\mathbf{x}) + \dots \right] \right\} \\ & \times \left\{ \psi(\mathbf{x}, T) + \xi \cdot \nabla \psi + \frac{1}{2} \sum_{m,n=1}^3 \xi_m \xi_n \frac{\partial^2 \psi}{\partial x_m \partial x_n} + \dots \right\} \end{aligned} \quad (4.9)$$

What makes the foregoing expression tractable is the fact that we are only interested in terms of first order in ϵ . Anything smaller (e.g., $\epsilon^{3/2}$) disappears, as in the discussion of the Trotter product formula. The factors ξ which appear inside the integral have an effective size in terms of ϵ , just as we found in the last section, $\Delta t \sim 1/\sqrt{\lambda}$. To be precise, we are considering the integral (4.9) in the limit $\epsilon \rightarrow 0$ and are keeping both the lowest terms (which turn out to be of order unity and cancel exactly) and the first term in ϵ (which will be of order ϵ), but no higher powers. The argument of the exponent, $im\xi^2/2\hbar\epsilon$, will be of order unity which is to say ξ is of order $\sqrt{\epsilon}$. This is demonstrated exactly as was done in the previous section, namely by doing Gaussian integrals.

First we note that the term $\exp(-ieV(\mathbf{x})/\hbar)$ can be factored from the integral—it does not depend on ξ —and can also be written as $(1 - ieV(\mathbf{x})/\hbar)$ to first order in ϵ .

Next the factor $\exp(-ie\nabla V \cdot \xi/\hbar)$ can simply be dropped. When the integration over ξ is performed this will be at most of order $\epsilon^{3/2}$. This, by the way, is the reason that it does not matter whether \mathbf{x}_j or \mathbf{x}_{j+1} is taken as the argument of V .

The other factors all remain, and we expand the exponential involving \mathbf{A} up to order ξ^2 , that is, up to ϵ . Equation 4.9 becomes

$$\begin{aligned}\psi(\mathbf{x}, T+\epsilon) = & \left(1 - \frac{i\epsilon V(\mathbf{x})}{\hbar}\right) \left(\frac{m}{2\pi i\hbar\epsilon}\right)^{3/2} \int d^3\xi \exp\left(\frac{im\xi^2}{2\epsilon\hbar}\right) \\ & \times \left[1 - \frac{ie}{\hbar c} \xi_n A_n - \frac{ie}{2\hbar c} \xi_n \xi_m \nabla_m A_n - \frac{1}{2} \frac{e^2}{\hbar^2 c^2} \xi_n A_n \xi_m A_m\right] \\ & \times \left[\psi + \xi_l \nabla_l \psi + \frac{1}{2} \xi_l \xi_k \nabla_l \nabla_k \psi\right] \\ & + \text{terms of higher order in } \epsilon\end{aligned}\quad (4.10)$$

It is understood in (4.10) that repeated indices k, l, m, n are summed from 1 to 3. By symmetry in ξ , all integrals over odd powers of ξ as well as all products of the form $\xi_n \xi_l$, $n \neq l$, give zero upon integration. Furthermore, each $\xi_n \xi_n$ (no summation) gives

$$\sqrt{\frac{2\pi\epsilon\hbar}{-im}} \left(\frac{2\epsilon\hbar}{2(-im)}\right) \quad (4.11)$$

by the appendix to Section 3 (3.13). The only cross term in (4.10) is thus from $\mathbf{A} \cdot \nabla \psi$ and integration yields

$$\begin{aligned}\psi(\mathbf{x}, T+\epsilon) = & \psi - \frac{i\epsilon}{\hbar} V\psi + \frac{\epsilon e}{2mc} \psi \nabla \cdot \mathbf{A} - \frac{i\epsilon e^2}{2\hbar mc^2} \mathbf{A}^2 \psi \\ & + \frac{i\epsilon\hbar}{2m} \nabla^2 \psi + \frac{\epsilon e}{mc} (\mathbf{A} \cdot \nabla) \psi\end{aligned}\quad (4.12)$$

The cancellation of the normalization factor is of course no accident. $\psi(\mathbf{x}, T)$ is brought to the left hand side, which becomes

$$\psi(\mathbf{x}, T+\epsilon) - \psi(\mathbf{x}, T) = \epsilon \frac{\partial \psi}{\partial t} \quad (4.13)$$

We divide the entire equation by ϵ and multiply by $\hbar i$. The result is

$$\begin{aligned}i\hbar \frac{\partial \psi}{\partial t} = & -\frac{\hbar^2}{2m} \nabla^2 \psi + \frac{i\hbar e}{mc} (\mathbf{A} \cdot \nabla) \psi \\ & + \frac{i\hbar e}{2mc} \psi \nabla \cdot \mathbf{A} + \frac{e^2}{2mc^2} \mathbf{A}^2 \psi + V\psi\end{aligned}\quad (4.14)$$

which is the same as

$$i\hbar \frac{\partial \psi}{\partial t} = \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A} \right)^2 \psi + V\psi \quad (4.15)$$

with \mathbf{p} the operator $-i\hbar\nabla$. Therefore wave functions propagated by the path integral have a time evolution satisfying Schrödinger's equation.

NOTES

This section follows Feynman's original paper [*Rev. Mod. Phys.* **20**, 367 (1948)] quite closely.

The Ito Integral and Gauge Transformations

Let us examine the consequences of an incorrect evaluation of

$$\int \mathbf{A} \cdot d\mathbf{x} \quad (5.1)$$

Suppose this were interpreted as

$$\sum_{j=0}^N (\mathbf{x}_{j+1} - \mathbf{x}_j) \cdot \mathbf{A}(\mathbf{x}_j) \quad (5.2)$$

This would have the effect in (4.9) of replacing $\frac{1}{2}(\xi \cdot \nabla) \mathbf{A}$ by $(\xi \cdot \nabla) \mathbf{A}$. Following the consequences of this replacement to (4.12), we find $(\epsilon e / mc) \psi \nabla \cdot \mathbf{A}$ instead of half that quantity. Finally this will have the effect of adding an additional term

$$\frac{i\hbar e}{2mc} \psi \nabla \cdot \mathbf{A} \quad (5.3)$$

to the Schrödinger equation, a term that does not belong there.

You should receive these remarks with the greatest alarm. The basic fact about Riemann integrals, the fact that makes them well defined objects, is the property that approximating sums

$$\sum_{j=0}^N (t_{j+1} - t_j) f(t_j^*) \quad (5.4)$$

all converge to the same number, called $\int_a^b f(t) dt$, so long as $\max_j (t_{j+1} - t_j) \rightarrow 0$, no matter what points t_j^* are selected (for $t_{j+1} \geq t_j^* \geq t_j$, $t_{N+1} = b$, $t_0 = a$). It follows that whatever sense we ultimately make out of the integral (5.1), it will not be a Riemann integral.

In the study of Brownian motion, K. Ito found similar problems in attempting to define an integral

$$\int_a^b f(x(t)) dx \quad (5.5)$$

where $x(t)$ is a Brownian motion path, that is, a particular trajectory followed by a particle undergoing Brownian motion. He chose to define this integral as

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} (x_{j+1} - x_j) f(x_j),$$

$$x_j = x\left(a + \frac{j(b-a)}{n}\right), \quad x_n = x(b) = x_b, x_0 = x(a) = x_a \quad (5.6)$$

He then found the remarkable result

$$\int_{x_a}^{x_b} \frac{d\varphi}{dx} dx = \varphi(x_b) - \varphi(x_a) - \frac{1}{2} \int_a^b \frac{d^2\varphi(x(t))}{dx^2} dt \quad (5.7)$$

The explanation of this strange formula lies in the behavior of Brownian motion paths. Wiener found that almost all of them are continuous but nowhere differentiable. Let $\Delta x = x(t+\Delta t) - x(t)$. Then it is known for Brownian motion paths that $(\Delta x)^2$ is of the order of Δt [the constant relating them is called the diffusion constant which is assumed to be $\frac{1}{2}$ in (5.7)]. This property holds for arbitrarily small Δt (in mathematics—not in physical Brownian motion). The velocity, or derivative of x , is $\Delta x / \Delta t$, but this is of order $1/\sqrt{\Delta t}$ which goes to infinity as $\Delta t \rightarrow 0$. In this light let us examine (5.7). Consider the identity

$$\varphi(x_b) - \varphi(x_a) = \sum_{j=0}^{n-1} [\varphi(x_{j+1}) - \varphi(x_j)] \quad (5.8)$$

with the same conventions as before for x_j , and so on. φ is assumed to be a well behaved function and (5.8) can be rewritten

$$\begin{aligned} \varphi(x_b) - \varphi(x_a) &= \sum_{j=0}^{n-1} [\varphi'(u_{\theta j})(x_{j+1} - x_j) + \frac{1}{2}\varphi''(u_{\theta j})(x_{j+1} - x_j)^2(1-2\theta) \\ &\quad + O((x_{j+1} - x_j)^3)] \end{aligned} \quad (5.9)$$

where $u_{\theta j}$ is defined as

$$u_{\theta j} = x_j + \theta(x_{j+1} - x_j) \quad (5.10)$$

and $1 > \theta > 0$. Suppose we let $\theta = 0$; then the first sum in (5.9) becomes, in the limit $n \rightarrow \infty$, the Ito integral. But the important thing to notice is that as $n \rightarrow \infty$ the second term in the sum does not vanish! Because $(\Delta x)^2$ is of order Δt the limit of $\sum \varphi''(\Delta x)^2$ is $\int \varphi'' dt$ and not zero. This becomes the extra integral in (5.7). Note that if one were to take $\theta = \frac{1}{2}$ —the midpoint rule—the anomalous term would not appear.

In quantum mechanics we have found the need for the midpoint rule and by now the reason for our difficulties should be evident. We found in Section 4 that the “size” of ξ is $\sqrt{\epsilon}$. Thus the broken line paths that we “sum” to get the path integral have the property that the characteristic size of $(x_{j+1} - x_j)^2$ is $\epsilon (= T/N)$. The paths are therefore like Brownian motion paths and due caution must be exercised in using them in integrals.

The complete expression for the path integral is

$$G(\mathbf{x}, t; \mathbf{y}) = \int_{(\mathbf{y}, 0)}^{(\mathbf{x}, t)} d\mathbf{x}(\tau) \exp \left\{ \frac{i}{\hbar} \int_0^t d\tau \left[\frac{1}{2} m \dot{\mathbf{x}}^2 - V(\mathbf{x}) + \frac{e}{c} \dot{\mathbf{x}} \cdot \mathbf{A} \right] \right\} \quad (5.11)$$

Let us perform a gauge transformation on \mathbf{A} , taking

$$\mathbf{A}' = \mathbf{A} + \nabla \varphi \quad (5.12)$$

where φ is some scalar function. This adds a term

$$\frac{ie}{\hbar c} \int_0^t \dot{\mathbf{x}} \cdot \nabla \varphi d\tau = \frac{ie}{\hbar c} \int_y^x \nabla \varphi \cdot d\mathbf{x} \quad (5.13)$$

to the argument of the exponential in (5.11). Since we are using the midpoint in evaluating the limit sums for the integral of (5.13), there will be no extra term ($\int \nabla^2 \varphi dt$) such as appears in the Ito formula, (5.7). If φ is a well behaved function on the Euclidean 3-space on which our problem is defined we will get

$$\int_y^x \nabla \varphi \cdot d\mathbf{x} = \varphi(\mathbf{x}) - \varphi(\mathbf{y}) \quad (5.14)$$

This quantity is the same for all paths. We have thus found that the gauge transformation (5.12) induces the transformation on G

$$G'(\mathbf{x}, t; \mathbf{y}) = e^{ie\varphi(\mathbf{x})/\hbar c} G(\mathbf{x}, t; \mathbf{y}) e^{-ie\varphi(\mathbf{y})/\hbar c} \quad (5.15)$$

But this only corresponds to the change in phase of the wave function

$$\psi'(\mathbf{x}) = e^{ie\varphi(\mathbf{x})/\hbar c} \psi(\mathbf{x}) \quad (5.16)$$

The foregoing amounts to a proof of gauge invariance in quantum mechanics. In path integral theory—once you know the midpoint rule—the proof is a matter of two lines. The proof from Schrödinger's equation, while mathematically equivalent, is not so obvious or simple.

Exercise: Prove gauge invariance for a gauge transformation on the potentials for both magnetic and electric fields.

NOTES

Equation 5.7 is actually a probabilistic statement and is true almost everywhere with respect to Wiener measure. It is false for functions $x(t)$ which are differentiable, but of course these form a set of measure zero among the continuous functions. The proof of (5.7) can be found in

K. Ito, *Mem. Am. Math. Soc.*, No. 4 (1951)

For some applications other than quantum mechanics the midpoint is also used in evaluating the integral [i.e., $\theta = \frac{1}{2}$ in (5.9)]. See

R. L. Stratonovich, *Conditional Markov Processes and their Application to the Theory of Optimal Control*, American Elsevier, New York, 1968.

Doing the Integral: Free Particle and Quadratic Lagrangians

The prescription for evaluating a path integral suggests that the task is formidable, but in some instances the calculation can be done in full.

For a free particle, $V=0$, and $A=0$. The three-dimensional propagator is simply the product of three one-dimensional propagators, so there is no point in cluttering our equations with vectors. We wish to evaluate

$$\begin{aligned} G(x, t; y) &= \int_{y_0}^{xt} dx(\tau) \exp \left[\frac{i}{\hbar} \int_0^t \frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 d\tau \right] \\ &= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi i \hbar \epsilon} \right)^{(N+1)/2} \int dx_1 \cdots dx_N \exp \left[\frac{im}{2\hbar \epsilon} \sum_{j=0}^N (x_{j+1} - x_j)^2 \right] \end{aligned} \quad (6.1)$$

(where $x_{N+1}=x$, $x_0=y$ and $\epsilon=t/(N+1)$). This integral turns out to be easy to do because of the following identity

$$\int_{-\infty}^{\infty} du \sqrt{\frac{a}{\pi}} e^{-a(x-u)^2} \sqrt{\frac{b}{\pi}} e^{-b(u-y)^2} = \sqrt{\frac{ab}{\pi(a+b)}} \exp \left[-\frac{ab}{a+b}(x-y)^2 \right] \quad (6.2)$$

This formula is proved by completing the square in the exponential. The fact that the expression can depend only on $(x-y)$ is obvious by a change of variables. Using (6.2) consider the integral on x_1 in (6.1).

Including two of the $(m/2\pi i\hbar\epsilon)^{1/2}$ factors this is

$$\left(\frac{m}{2\pi i\hbar\epsilon}\right) \int dx_1 \exp\left[\frac{im}{2\hbar\epsilon}(x_1 - x_0)^2 + \frac{im}{2\hbar\epsilon}(x_2 - x_1)^2\right] \quad (6.3)$$

which equals

$$\sqrt{\frac{m}{2\pi i\hbar(2\epsilon)}} \exp\left[\frac{im}{2\hbar(2\epsilon)}(x_2 - x_0)^2\right] \quad (6.4)$$

The effect of the x_1 integration is to change ϵ to 2ϵ (both in the square root and in the exponential) and to replace $(x_2 - x_1)^2 + (x_1 - x_0)^2$ by $(x_2 - x_0)^2$. The integral over x_2 changes 2ϵ to 3ϵ (both in the square root and in the exponential) and yields the term $(x_3 - x_0)^2$ in the exponential.

This is continued for all N integrals. The ϵ becomes $(N+1)\epsilon$, which is just t , and the distance squared appearing in the exponential becomes $(x_{N+1} - x_0)^2$, which is $(x - y)^2$. There is no longer any N dependence so the limit operation is trivial and we obtain the result

$$G(x, t; y) = \sqrt{\frac{m}{2\pi i\hbar t}} \exp\left[\frac{im}{2\hbar t}(x - y)^2\right] \quad (6.5)$$

For a free particle the classical path from $(y, 0)$ to (x, t) is

$$x(\tau) = y + \frac{\tau}{t}(x - y) \quad (6.6)$$

The classical action is

$$S = \int_0^t \frac{m}{2} \left(\frac{x - y}{t}\right)^2 d\tau = \frac{m}{2} \frac{(x - y)^2}{t} \quad (6.7)$$

Evidently the phase in (6.5) is just the classical action (divided by \hbar), and we find that the classical path gives the exact propagator for a free particle.

The second example to be studied is a quadratic Lagrangian, which includes as special cases the simple harmonic oscillator and the uniform force field. The Lagrangian is

$$L = \frac{1}{2}m\dot{x}^2 + b(t)x\dot{x} - \frac{1}{2}c(t)x^2 - e(t)x \quad (6.8)$$

and we wish to evaluate

$$G(x_b, t_b; x_a, t_a) = \int_{x_a t_a}^{x_b t_b} dx(\tau) \exp\left(\frac{i}{\hbar} \int_{t_a}^{t_b} L d\tau\right) \quad (6.9)$$

The change of variables that we now perform demands a certain amount of belief in the formal object in (6.9), but since the answer we get in the end can be checked by applying Schrödinger's equation there is no harm done if we proceed optimistically. Let $\bar{x}(t)$ be the solution of the Euler-Lagrange equation

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0 \quad (6.10)$$

or

$$m\ddot{x} + (c + \dot{b})x + e = 0 \quad (6.11)$$

with boundary conditions $x(t_a) = x_a$, $x(t_b) = x_b$. Let

$$y(\tau) = x(\tau) - \bar{x}(\tau) \quad (6.12)$$

be the new "integration variable" in (6.9). The "Jacobian" of this change of variables is 1. What this means is that if (6.9) were written out as a limit, each x_j , would be replaced by $y_j = x_j - \bar{x}(t_j)$, which is just a translation. The action term in (6.9) behaves simply. This is most easily seen by working formally

$$\begin{aligned} \int L dt &= S(x(\tau)) = S(y(\tau) + \bar{x}(\tau)) \\ &= S(\bar{x}(\tau)) + \frac{\delta S}{\delta x} \Big|_{\bar{x}} y(\tau) + \frac{1}{2} \frac{\delta^2 S}{\delta x^2} \Big|_{\bar{x}} y(\tau)^2 \end{aligned} \quad (6.13)$$

The Taylor expansion of S terminates after its second derivative, since S is a quadratic functional. Furthermore, the first derivative term is zero since this is just the definition of \bar{x} —the classical path is that for which the first variation of S vanishes. Finally, again because S is quadratic, $\frac{1}{2} \delta^2 S / \delta x^2$ is the same quadratic form as S except that the term that was linear in x does not appear. It follows that

$$S(\bar{x} + y) = S(\bar{x}) + \frac{1}{2} \delta^2 S y^2 \quad (6.14)$$

$S(\bar{x})$ can be factored out of the path integral over $dy(\tau)$. Taking into account the fact that the boundary conditions for each $y(\tau)$ are

$$\begin{aligned} y(t_a) &= 0 \\ y(t_b) &= 0 \end{aligned} \quad (6.15)$$

[by (6.12)], we obtain

$$G(x_b, t_b; x_a, t_a) = \exp \left[\frac{iS(x_b, t_b; x_a, t_a)}{\hbar} \right] \tilde{G}(0, t_b; 0, t_a) \quad (6.16)$$

where $S(x_b, t_b; x_a, t_a)$ is the action along the classical path \bar{x} from (x_a, t_a) to (x_b, t_b) and \tilde{G} is the Green's function for the Lagrangian $L = \frac{1}{2}m\dot{x}^2 + b\dot{x}\dot{x} - \frac{1}{2}\epsilon x^2$.

Evaluation of the factor $\tilde{G}(0, t_b; 0, t_a)$, a function of time only, gives us the opportunity to do the first respectable integral of the book. Taking the usual definitions we have

$$\begin{aligned} \tilde{G}(0, t_b; 0, t_a) &= \lim_{N \rightarrow \infty} \int dy_1 \cdots dy_N \left(\frac{m}{2\pi i \hbar \epsilon} \right)^{(N+1)/2} \\ &\times \exp \left\{ \frac{i}{\hbar} \sum_{j=0}^N \left[\frac{m}{2\epsilon} (y_{j+1} - y_j)^2 - \frac{1}{2}\epsilon c_j y_j^2 \right] \right\} \end{aligned} \quad (6.17)$$

where $c_j = c(t_j)$, $t_j = t_a + (j/N)(t_b - t_a)$ and where we have dropped the term $b\dot{x}\dot{x}$ because of its irrelevance in the one-dimensional problem. (In three dimensions an $\mathbf{x} \times \dot{\mathbf{x}}$ term has physical consequences and we leave as an exercise the development of the analogue of the results we shall obtain in one dimension.)

To deal with the multitude of integrals in (6.17) we define an N vector η

$$\eta = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \quad (6.18)$$

and note that the argument of the exponent in (6.17) can be written

$$-\eta^T \sigma \eta \quad (6.19)$$

where T signifies transpose and σ is the matrix

$$\sigma = \frac{m}{2\epsilon \hbar i} \begin{pmatrix} 2 & -1 & & & & & & & \\ -1 & 2 & -1 & & & & & & \\ & -1 & 2 & \ddots & & & & & \\ & & \ddots & \ddots & & & & & \\ & & & & 2 & -1 & & & \\ & & & & & -1 & 2 & & \\ & \ddots & & & & & & \ddots & \\ \textcircled{O} & & & & & & & & \textcircled{O} \\ & & & & & & & & c_N \end{pmatrix} + \frac{i\epsilon}{2\hbar} \begin{pmatrix} c_1 & & & & & & & \\ & \ddots & & & & & & & \\ & & \ddots & & & & & & \\ & & & \ddots & & & & & \\ & & & & \textcircled{O} & & & & \\ & & & & & \ddots & & & \\ & & & & & & \ddots & & \\ & & & & & & & \textcircled{O} & c_N \end{pmatrix} \quad (6.20)$$

Equation 6.17 can now be concisely written

$$\tilde{G} = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi i \hbar \epsilon} \right)^{(N+1)/2} \int d^N \eta \exp(-\eta^T \sigma \eta) \quad (6.21)$$

There is a variety of ways to evaluate the integral in (6.21). We use a straightforward approach. Consider

$$\int d^N \eta e^{-\eta^T \sigma \eta} \quad (6.22)$$

where σ is of the form $i\tilde{\sigma}$ with $\tilde{\sigma}$ real and Hermitian. Then σ can be diagonalized by a unitary matrix:

$$\sigma = U^\dagger \sigma_D U \quad (6.23)$$

where σ_D is the diagonal matrix of eigenvalues of σ (call these σ_α) and U is unitary. For real eigenvectors U is also real; let $\zeta = U\eta$. The vector ζ is real and $\{\zeta\}$ is an N dimensional vector space. Note that because $|\det U| = 1$ the Jacobian for going from $d^N \eta$ to $d^N \zeta$ is 1. Therefore

$$\int d^N \eta e^{-\eta^T \sigma \eta} = \int d^N \zeta e^{-\zeta^T \sigma_D \zeta} = \prod_{\alpha=1}^N \sqrt{\frac{\pi}{\sigma_\alpha}} = \frac{\pi^{N/2}}{\sqrt{\det \sigma}} \quad (6.24)$$

provided there are no zero eigenvalues. If there are any zero eigenvalues the integral simply does not exist.

The evaluation of $\det \sigma$ proceeds by a well-known technique. Our interest is in

$$\begin{aligned} \tilde{G}(0, t_b; 0, t_a) &= \lim_{N \rightarrow \infty} \left[\left(\frac{m}{2\pi i \hbar \epsilon} \right)^{N+1} \frac{\pi^N}{\det \sigma} \right]^{1/2} \\ &= \lim_{N \rightarrow \infty} \left[\frac{m}{2\pi i \hbar} \cdot \frac{1}{\epsilon} \cdot \frac{1}{\left(\frac{2i\hbar\epsilon}{m} \right)^N \det \sigma} \right]^{1/2} \end{aligned} \quad (6.25)$$

Define f as

$$f(t_b, t_a) = \lim_{N \rightarrow \infty} \left[\epsilon \left(\frac{2i\hbar\epsilon}{m} \right)^N \det \sigma \right] \quad (6.26)$$

From (6.20) it follows that

$$\left(\frac{2i\hbar\epsilon}{m} \right)^N \det \sigma = \det \left[\begin{array}{cccccc} 2 & -1 & & & & & \textcircled{O} \\ -1 & 2 & & & & & \\ & & \ddots & & & & \\ & & & \ddots & & & \\ & & & & -1 & 2 & \\ \textcircled{O} & & & & & & \\ & & & & & & \end{array} \right] - \frac{\epsilon^2}{m} \left[\begin{array}{cccccc} c_1 & & & & & & \textcircled{O} \\ & \ddots & & & & & \\ & & \ddots & & & & \\ & & & c_N & & & \end{array} \right] \equiv \det \sigma'_N \equiv p_N \quad (6.27)$$

where σ'_N is defined to be the matrix in the brackets. We first define truncated $j \times j$ matrices σ'_j —but ϵ remains $(t_b - t_a)/(N+1)$ —that consist of the first j rows and columns of σ'_N and whose determinants are p_j . By expanding σ'_{j+1} in minors, the following recursion formula for the p_j is easy to show:

$$p_{j+1} = \left(2 - \frac{\epsilon^2}{m} c_{j+1} \right) p_j - p_{j-1}, \quad j = 1, \dots, N \quad (6.28)$$

where $p_1 = 2 - \epsilon^2 c_1 / m$ and p_0 is defined to be +1. When (6.28) is rewritten in the form

$$\frac{p_{j+1} - 2p_j + p_{j-1}}{\epsilon^2} = -\frac{c_{j+1} p_j}{m} \quad (6.29)$$

it becomes evident that p_N will be obtained by solving a differential equation. If we now let

$$\varphi(t) = \epsilon p_j \quad (6.30)$$

for $t = t_a + j\epsilon$ we observe that (6.29) implies that in the limit $\epsilon \rightarrow 0$, $\varphi(t)$ satisfies the equation

$$\frac{d^2\varphi}{dt^2} = -\frac{c(t)}{m} \varphi \quad (6.31)$$

The initial values for φ follow from

$$\varphi(0) = \epsilon p_0 \rightarrow 0$$

$$\frac{d\varphi(0)}{dt} = \epsilon \left(\frac{p_1 - p_0}{\epsilon} \right) = 2 - \frac{\epsilon^2}{m} c_1 - 1 \rightarrow 1, \quad \text{as } N \rightarrow \infty \quad (6.32)$$

Consequently $f(t_b, t_a) = \varphi(t_b)$ is obtained by solving the differential equation

$$m \frac{\partial^2 f(t, t_a)}{\partial t^2} + c(t) f(t, t_a) = 0 \quad (6.33)$$

with the initial conditions

$$f(t_a, t_a) = 0, \quad \left. \frac{\partial f}{\partial t}(t, t_a) \right|_{t=t_a} = 1 \quad (6.34)$$

We have thus arrived at the following form for the Green's function for the quadratic Lagrangian given in (6.8):

$$G(x_b, t_b; x_a, t_a) = \sqrt{\frac{m}{2\pi i \hbar f(t_b, t_a)}} \exp\left[\frac{i}{\hbar} S_c(x_b, t_b; x_a, t_a)\right] \quad (6.35)$$

where S_c is the action along the classical path and f is given by (6.33) and (6.34).

Exercise: Under what circumstance can there be more than one classical path connecting the given end points?

Exercise: Show that (6.5) (the free particle) is a special case of (6.35).

We next specialize the foregoing results to the simple harmonic oscillator. For this system $e(t) \equiv 0$ and $c(t) \equiv m\omega^2 = \text{const}$. Equations 6.33 and 6.34 are easily solved to yield

$$f(t_b, t_a) = \frac{\sin \omega T}{\omega} \quad (6.36)$$

where $T = t_b - t_a$. The more tedious job of evaluating $S(x_b, t_b; x_a, t_a)$ along the classical path is left as an exercise.

Exercise: Show that for $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2x^2$ the action along the classical path from (x_a, t_a) to (x_b, t_b) , $T = t_b - t_a$, is

$$S(x_b, t_b; x_a, t_a) = \frac{m\omega}{2\sin\omega T} [(x_b^2 + x_a^2) \cos\omega T - 2x_a x_b] \quad (6.37)$$

The exact Green's function is therefore

$$G(x_b, t_b; x_a, t_a) = \sqrt{\frac{m\omega}{2\pi i\hbar \sin\omega T}} \times \exp\left\{\frac{im\omega}{2\hbar \sin\omega T} [(x_b^2 + x_a^2) \cos\omega T - 2x_a x_b]\right\}. \quad (6.38)$$

The particle in a uniform field is even simpler. The Lagrangian is

$$L = \frac{1}{2}m\dot{x}^2 - ex \quad (6.39)$$

[so that $e(t) \equiv e = \text{const.}$] and you can easily show that*

$$G(x_b, t_b; x_a, t_a) = \sqrt{\frac{m}{2\pi i\hbar T}} \exp\left\{\frac{i}{\hbar} \left[\frac{m(x_b - x_a)^2}{2T} - \frac{1}{2}eT(x_b + x_a) - \frac{e^2 T^3}{24m} \right]\right\} \quad (6.40)$$

Finally we quote the result for

$$L = \frac{1}{2}m\dot{x}^2 - \frac{m\omega^2}{2}x^2 - e(t)x \quad (6.41)$$

The function $f(T)$ is the same as that given in (6.36), since the function $e(t)$ has no effect on f , while the classical action is

$$\begin{aligned} S(x_b, t_b; x_a, t_a) &= \frac{m\omega}{2\sin\omega T} \left[(x_b^2 + x_a^2) \cos\omega T - 2x_a x_b \right. \\ &\quad - \frac{2x_b}{m\omega} \int_{t_a}^{t_b} e(t) \sin\omega(t-t_a) dt \\ &\quad - \frac{2x_a}{m\omega} \int_{t_a}^{t_b} e(t) \sin\omega(t_b-t) dt \\ &\quad \left. - \frac{2}{m^2\omega^2} \int_{t_a}^{t_b} \int_t^t e(t) e(s) \sin\omega(t_b-t) \sin\omega(s-t_a) ds dt \right]. \end{aligned} \quad (6.42)$$

*Beware of misprints in Feynman and Hibbs, Equation 3-62.

Exercise: Obtain a closed form expression for the (time dependent) Green's function for a charged particle in three dimensions under the following circumstances:

- (a) Uniform electric field.
- (b) Uniform magnetic field.

Solution to (b): For field $\mathbf{B} = B\hat{\mathbf{B}}$, mass m , charge e , $c = 1$, $\hbar = 1$, $\omega = eB/m$, $\mathbf{r} = \mathbf{x}_2 - \mathbf{x}_1$

$$G(\mathbf{x}_2, t; \mathbf{x}_1, 0) = \left(\frac{m}{2\pi i t} \right)^{3/2} \frac{\omega t}{2} \csc \frac{\omega t}{2} \cdot \exp \left\{ \frac{im}{2} \left[\frac{\omega}{2} \cot \frac{\omega t}{2} (\hat{\mathbf{B}} \times \mathbf{r})^2 - \omega \hat{\mathbf{B}} \cdot \mathbf{x}_2 \times \mathbf{x}_1 + (\hat{\mathbf{B}} \cdot \mathbf{r})^2 \right] \right\}$$

NOTES

Methods for handling the quadratic Lagrangian are legion and have been well developed since the earliest work on path integrals. Oddly enough papers on the subject continue to appear and may give some historian of science material for a case history on the nondiffusion of knowledge. Our treatment of the quadratic Lagrangian is similar to those of

I. M. Gelfand and A. M. Yaglom, *J. Math. Phys.* 1, 48 (1960)

E. W. Montroll, *Commun. Pure and Appl. Math.* 5, 415 (1952)

There is no shortage of other methods, and any document purporting to be an introduction to path integration will have some derivation of the results.

APPENDIX: EXACTNESS OF THE SUM OVER CLASSICAL PATHS

For all path integrals evaluated above, the result was expressible entirely in terms of the classical path. It turns out that every propagator that anyone has ever been able to evaluate exactly and in closed form is a sum over "classical paths" only. There has even appeared in the literature the claim that *all* propagators are given exactly as a sum over classical paths. Recall that we are discussing the time dependent propagator, so that quantum mechanical tunneling, which is a phenomenon at fixed energy, does not constitute an immediate contradiction to the foregoing claim. Nevertheless, we shall see (Sections 15 and 16) that the propagator near a caustic does not vanish and shall even get an expression for its value in the shadow region—that is, in the region where no classical paths penetrate. This absence of classical paths is not due to a fixing of the energy but is rather a consequence of the focusing effect of the potential responsible for the caustic. Other counterexamples have also been suggested.*

*E. Nelson gave the following counterexample. Let the potential V be zero except in a small region where it is smooth and small. For sufficiently small times and for two points near each other but far from the support of V there is no classical path between these points that "feels" V . Nevertheless, we know that for arbitrarily small times the propagator is influenced by V .

As seen in this section Lagrangians (or Hamiltonians) for which the sum over classical paths is exact include those that are at most quadratic in the coordinate and its derivative. This will also be evident in our development of the WKB method where a quadratic Lagrangian has the special property that the third variation of the action, $\delta^3 S$, and all higher variations, are automatically zero.

Another situation leading to the exact sum over classical paths is where the particle is confined in a simple geometry by perfectly reflecting walls. A trivial example of this is the free particle confined to the positive x axis (ignore y and z) by a potential $V(x) = +\infty$ for $x < 0$. The propagator is

$$G(b, T; a, 0) = \sqrt{\frac{1}{2\pi iT}} [e^{i(b-a)^2/2T} - e^{i(b+a)^2/2T}] \quad (6.43)$$

where we have taken $\hbar = m = 1$ for convenience. See the discussion following (9.38) for a proof of this assertion. Now the two terms in (6.43) can be associated with the two classical paths from a to b in time T : one direct path with action $(b-a)^2/2T$ and one reflected path with action $(b+a)^2/2T$. The only difficulty from the path integral point of view is the minus sign for the reflected path. If a turning point were involved, a relative phase of $\pi/2$ might appear, but there is no way for a phase of π to occur in one dimension. There is an *ad hoc* way to put in the sign, a way that has been employed in the more difficult problem of an obstacle in two or three dimensions. One pastes another copy of the positive real axis onto the existing one, calculates the classical path to the image of the final point in the new space ignoring what transpires where they are attached (in higher dimensions there is a creep wave), and adds the propagators associated with the two images of the final point. The relative sign with which they are taken is determined by the boundary condition at the obstacle. For vanishing wave function the two contributions must be subtracted. If one desired vanishing derivative of the wave function, the terms would be added.

This rule, while correct, is an embarrassment to the purist so long as it has not been derived in a purely path integral context. In the sum over paths the only change that has occurred is that integrals of the form

$$\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_N$$

have had the range of integration restricted to $x_i > 0$, $i = 1, \dots, N$. It is not clear if this is enough to recover the correct form easily, since for any given

mesh in time there may be ranges in x for which one needs both direct and reflected paths even for the “infinitesimal” propagator.

Despite the foregoing “embarrassment” the propagator does involve classical paths only. It is thus likely that the propagator is given exactly by such a sum in those geometries for which the “method of images” works in electrodynamics. In both cases one uses special properties of the geometry to satisfy the boundary conditions.

There is another, more subtle, situation where the sum over classical paths is exact. This is for Green’s functions for free motion on the group manifolds of Lie groups. This was first found to be true for $SU(2)$ and at the time conjectured to hold for the larger class. Subsequently, Dowker proved the result for a large class of group manifolds. However, the exactness of the sum is not generally true for homogeneous spaces such as $S^2 = SU(2)/U(1)$. Further discussion of curved space path integrals is given in Section 24. Here we only wish to remark that while other examples of exact sums may seem in a sense trivial, the curved space path integral with its plethora of $g_{\mu\nu}$ ’s and curvature terms looks anything but. Of course that may only mean that one simply needs the right coordinates to make the group manifold path integral look trivial too. It may also be possible to define harmonic oscillator like potentials on group manifolds and obtain closed form Green’s functions for them.

Finally we refer the reader to Section 31 for remarks on systems with solition modes and global invariants and on the possibility that such systems might yield exact semiclassical “approximations.”

Notes

An expression for the propagator in terms of classical paths only appears in

M. Clutton-Brock, *Proc. Camb. Phil. Soc.* **61**, 201 (1965)

The Green’s function in the presence of a reflecting obstacle has been calculated by

V. S. Buslaev, in *Topics in Mathematical Physics*, Sh. Birman, Ed., Consultants Bureau, New York, 1967, p. 67

Exactness of the sum over classical paths for $SU(2)$ is shown in

L. S. Schulman, *Phys. Rev.* **176**, 1558 (1968)

and for Lie group manifolds by

J. S. Dowker, *J. Phys. A* **3**, 451 (1970)

J. S. Dowker, *Ann. Phys.* **62**, 561 (1971)

SEVEN

Properties of Green's Functions; the Feynman-Kac Formula

In this section we develop further properties of Green's functions. This gives some perspective on the comprehensiveness of the information contained in a formula like (6.38), which is the explicit closed form Green's function for a particular dynamical system.

In operator form the time dependent Green's function (for time independent H) is

$$G(t) = \theta(t) e^{-iHt/\hbar} \quad (7.1)$$

The family $\{G(t) | t > 0\}$ satisfies

$$G(t+s) = G(t)G(s) \quad t, s > 0 \quad (7.2)$$

which in coordinate space takes the form

$$G(x, t+s; y, 0) = \int dz G(x, t; z, s)G(z, s; y, 0) \quad (7.3)$$

This innocuous equation is an important consistency requirement underlying the whole theory. When the formalism of path integration is carried over to Brownian motion it is known as the Chapman-Kolmogorov equation. (See Kac's book for elaboration.) In our work it is implicit in the definition of a propagator.

Suppose we have a complete set of energy eigenstates $|n\rangle$, $n=0, 1, \dots$ with

$$H|n\rangle = E_n |n\rangle \quad (7.4)$$

These correspond to wavefunctions

$$u_n(x) = \langle x | n \rangle \quad (7.5)$$

Clearly, for $t > 0$,

$$\begin{aligned} G(x, t; y, 0) &= \langle x | G(t) | y \rangle = \sum_n \langle x | n \rangle \langle n | e^{-iHt/\hbar} | y \rangle \\ &= \sum_n u_n(x) e^{-iE_n t/\hbar} u_n^*(y) \end{aligned} \quad (7.6)$$

This formula enables us to go from facts about the time dependent Green's function to the more traditional information about energy levels and wave functions.

Of interest also is the Fourier transform of (7.1), namely

$$\tilde{G}(z) = \int_{-\infty}^{\infty} e^{izt/\hbar} G(t) dt = \int_0^{\infty} e^{i(z-H)t/\hbar} dt \quad (7.7)$$

For $\text{Im } z > 0$ we have

$$\tilde{G}(z) = \frac{i\hbar}{z - H} \quad (7.8)$$

Singularities at $\text{Im } z = 0$ arise because the spectrum of H is real. Fourier transformation applied to (7.6) yields

$$\tilde{G}(x, y; z) = i\hbar \sum_n \frac{u_n(x) u_n^*(y)}{z - E_n} \quad (7.9)$$

For a discrete spectrum we thus identify the poles in the energy Green's function with the location of the bound states and the residues with the bound state wave functions. For the continuous spectrum of H , G has a cut. All this should be familiar from the usual formulation of quantum mechanics.

From (7.6) follows the Feynman-Kac formula, which relates the asymptotic ($t \rightarrow -i\infty$) behavior of the time dependent Green's function to the ground state energy level. In (7.6) we set $y = x$ and integrate over all x

(one or several dimensional integral, as the case may be). By the normalization of the energy eigenfunctions the result is

$$\int dx G(x, t; x, 0) = \sum_n e^{-iE_n t/\hbar} \int dx |u_n(x)|^2 = \sum_n e^{-iE_n t/\hbar} \quad (7.10)$$

We consider this equation for complex t , namely $it/\hbar = \tau$. For large (positive) τ the sum on the right of (7.10) will be dominated by the contribution from the lowest energy level, $\exp(-\tau E_0)$. We have therefore that

$$\lim_{\tau \rightarrow \infty} e^{E_0 \tau} \int G(x, -i\hbar\tau; x, 0) dx = 1 \quad (7.11)$$

for a nondegenerate ground state [for a degenerate ground state the degree of degeneracy appears on the right of (7.11)]. When there is no bound state (7.11) picks out the bottom of the continuum. If before taking the limit in τ we had taken logarithms, we would have instead of (7.11)

$$E_0 = - \lim_{\tau \rightarrow +\infty} \frac{1}{\tau} \log \int dx G(x, -i\hbar\tau; x, 0) = \lim_{\tau \rightarrow \infty} \frac{-1}{\tau} \log \text{Tr } e^{-\tau H} \quad (7.12)$$

This is the Feynman-Kac formula and has often shown itself to be an excellent way of computing ground state energy levels. One need not know all details of the Green's function, but only a certain asymptotic limit.

As an example of the foregoing formalism, we consider the harmonic oscillator for which we have found

$$G(x, t; y) = \sqrt{\frac{m\omega}{2\pi i\hbar \sin \omega t}} \exp \left\{ \frac{im\omega}{2\hbar \sin \omega t} [(x^2 + y^2) \cos \omega t - 2xy] \right\} \quad (7.13)$$

By (7.6) this must equal

$$G(x, t; y) = \sum_n e^{-iE_n t/\hbar} u_n(x) u_n^*(y) \quad (7.14)$$

with $u_n(x)$ normalized harmonic oscillator wave functions involving Hermite polynomials. It is easy to deduce the energy levels E_n from (7.10) in the following way. The trace of (7.13) involves a Gaussian integral

which can be performed immediately.

$$\begin{aligned} \sum_n \exp\left(-\frac{iE_n t}{\hbar}\right) &= \int dx G(x, t; x) \\ &= \int dx \sqrt{\frac{m\omega}{2\pi i\hbar \sin \omega t}} \exp\left\{\frac{im\omega}{2\hbar \sin \omega t} \cdot 2x^2(\cos \omega t - 1)\right\} \\ &= \frac{1}{2i} \left(\sin \frac{\omega t}{2}\right)^{-1} \end{aligned} \quad (7.15)$$

By writing the inverse of the sin as follows

$$\begin{aligned} \frac{1}{2i} \left(\sin \frac{\omega t}{2}\right)^{-1} &= \frac{\exp\left(-\frac{1}{2}i\omega t\right)}{1 - \exp(-i\omega t)} \\ &= \exp\left(-\frac{1}{2}i\omega t\right) \sum_{n=0}^{\infty} \exp(-in\omega t) \end{aligned}$$

(putting in appropriate small imaginary parts where needed) and comparing the extreme left term in (7.15) it is obvious that

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad n = 0, 1, \dots \quad (7.16)$$

Feynman and Hibbs in their book go through an expansion of the Green's function without taking a trace and get information on the wave functions too.

The free particle also illustrates some of the points made above. We have found

$$G(x, t; 0) = \sqrt{\frac{m}{2\pi i\hbar t}} \exp\left[\frac{imx^2}{2\hbar t}\right] \quad (7.17)$$

The Fourier transform is

$$\tilde{G}(E) = \int_0^\infty dt \sqrt{\frac{m}{2\pi i\hbar t}} \exp\left[\frac{i}{\hbar}\left(Et + \frac{mx^2}{2t}\right)\right] \quad (7.18)$$

To evaluate this we use the following identity (see the Appendix)

$$\int_0^\infty du e^{-(a/u^2) - bu^2} = \sqrt{\frac{\pi}{4b}} e^{-2\sqrt{ab}} \quad (7.19)$$

with the change of variables $u^2 = t$. Thus

$$\tilde{G}(E) = \sqrt{\frac{m}{2E}} e^{i|x|\sqrt{2mE}/\hbar} \quad (7.20)$$

which gives the well known branch point in the E plane at $E=0$. Just where the branch cut will run (for that branch point) will depend on how the cut is defined for (7.19). The derivation of (7.19) assumed a and b to be real and positive so the location of the cut is open.

Exercise 1: Use the Feynman-Kac formula to get the ground state energy and wave function for the simple harmonic oscillator.

The trace operation in (7.12) is not vital and one can use instead of $\text{Tr}G$, G itself evaluated at specific spatial positions, provided the ground state wave function does not vanish at those points. This freedom can also be used to advantage to get information on other levels. For a spinless particle in one dimension suppose $V(x) = V(-x)$; then the ground state (ψ_0) will be symmetric and the first excited state (ψ_1) antisymmetric. Consider

$$G(x, t; y, 0) - G(-x, t; y, 0) = \sum_{n=0}^{\infty} \exp\left(\frac{-iE_n t}{\hbar}\right) [\psi_n(x) - \psi_n(-x)] \psi_n^*(y)$$

For $it=\tau$, $\tau \rightarrow +\infty$, the largest contribution to the difference of Green's functions will come from E_1 .

Exercise 2: Use this method to find the energy of the first excited level of the harmonic oscillator.

Exercise 3: Show that the propagator for a particle of mass m in an infinite square well (walls at $x=0$ and $x=a$) is

$$G(x, t; y) = \sqrt{\frac{m}{2\pi i\hbar t}} \sum_{n=-\infty}^{\infty} \left[\exp\left(\frac{im(x-y+2na)^2}{2\hbar t}\right) - \exp\left(\frac{im(x+y+2na)^2}{2\hbar t}\right) \right]$$

Relate this to the eigenfunction expansion for G [Hint: see (23.11) and (23.13).]

APPENDIX: EVALUATION OF THE INTEGRAL

$$\int_0^\infty e^{-(a/u^2) - bu^2} du \quad (7.21)$$

Let $y = u\sqrt{b}$, $c = \sqrt{ab}$. Then (7.21) is equal to

$$f(c) = \frac{1}{\sqrt{b}} \int_0^\infty e^{-(c^2/y^2) - y^2} dy \quad (7.22)$$

Keeping b constant, take the derivative with respect to c :

$$\frac{\partial f}{\partial c} = \frac{1}{\sqrt{b}} \int_0^\infty dy \left(\frac{-2c}{y^2} \right) e^{-(c^2/y^2) - y^2} \quad (7.23)$$

Change variables to $z = c/y$ and note the additional sign change due to changed limits of integration:

$$\frac{\partial f}{\partial c} = -\frac{2}{\sqrt{b}} \int_0^\infty e^{-z^2 - (c^2/z^2)} dz = -2f(c) \quad (7.24)$$

It follows (assuming f to depend smoothly on c) that

$$f(c) = f(0)e^{-2c} \quad (7.25)$$

But $f(0)$ is a Gaussian and we get the result of (7.19).

NOTES

For the Chapman-Kolmogorov equation see Section IV.1-3 of

M. Kac, *Probability and Related Topics in the Physical Sciences*, Interscience, New York, 1959

For the Feynman-Kac formula in quantum field theory see

J. Glimm and A. Jaffe, Functional Integral Methods in Quantum Field Theory, in M. Levy and P. Mitter, Eds., *New Developments in Quantum Mechanics and Statistical Mechanics; Cargese 1976*, Plenum, New York, 1977

Although we have called (7.12) the Feynman-Kac formula, many authors mean by this formula the more general relation, namely the expression for the propagator as a functional integral.

In many, many applications one uses the Green's function to derive energy levels or wave functions by the methods described in this section. Some examples of this are:

M. J. Goovaerts and J. T. Devreese, *J. Math. Phys.* **13**, 1070 (1972) (for the Coulomb potential)

M. J. Goovaerts, A. Babcenko, and J. T. Devreese, *J. Math. Phys.* **14**, 554 (1973) (for the delta function potential)

M. C. Gutzwiller, *J. Math. Phys.* **8**, 1979 (1967) (for the Coulomb potential)

D. C. Khandekar and S. V. Lawande, *J. Math. Phys.* **16**, 384 (1975) (for $V = \frac{1}{2} \omega(t)^2 x^2 + g/x^2$)

Further instances will be found later in this book.

EIGHT

Functional Derivatives and Commutation Relations

In using path integrals it is possible to bring functional calculus to bear on quantum mechanics in a particularly convenient way.

First we define a functional derivative. Let $F[x(\tau)]$ be a functional. A “derivative” of F should say something about

$$F[x(\tau) + \eta(\tau)] - F[x(\tau)] \quad (8.1)$$

where $\eta(\tau)$ is a “small” function. If F were a function of N variables, say x_i , $i = 1, \dots, N$ (with similar definition for η_i) then we could write (8.1) as

$$F(x_1 + \eta_1, x_2 + \eta_2, \dots) - F(x_1, x_2, \dots) = \sum_i \frac{\partial F}{\partial x_i} \eta_i \quad (8.2)$$

This leads to the definition of functional derivative $\delta F / \delta x(\sigma)$ as that function of σ for which, for “small” η ,

$$F[x(\tau) + \eta(\tau)] - F[x(\tau)] = \int \frac{\delta F}{\delta x(\sigma)} \eta(\sigma) d\sigma \quad (8.3)$$

Because (8.3) is merely a generalization of (8.2) the usual rules for differentiation of products, integration by parts, and so on, can be derived for functional derivatives.

Although $\exp[iS[x(\tau)]/\hbar]$ is not a real valued weight function it is still possible to use it to define a kind of “expectation value.” For a functional

F , let

$$\langle F \rangle_S = \int dx(\tau) F[x(\tau)] \exp\left[\frac{iS[x(\tau)]}{\hbar}\right] \quad (8.4)$$

In (8.4) we note that we can translate $x(\tau)$, and because $dx(\tau) = d[x(\tau) + \eta(\tau)]$ the value of the integral remains the same. Thus

$$\begin{aligned} 0 &= \int dx(\tau) \{ F(x+\eta) e^{iS(x+\eta)/\hbar} - F(x) e^{iS(x)/\hbar} \} \\ &= \int dx(\tau) \int d\sigma \left[\frac{\delta F}{\delta x(\sigma)} + F \frac{i}{\hbar} \frac{\delta S}{\delta x(\sigma)} \right] e^{iS/\hbar} \eta(\sigma) + O(\eta^2) \end{aligned} \quad (8.5)$$

It follows from the arbitrariness of η that

$$\left\langle \frac{\delta F}{\delta x(\sigma)} \right\rangle_S = -\frac{i}{\hbar} \left\langle F \frac{\delta S}{\delta x(\sigma)} \right\rangle_S, \quad \text{for all } \sigma \quad (8.6)$$

This important result can also be written in discrete form, where $x(\tau)$ is the broken line connecting x_1, \dots, x_N , and the limit $N \rightarrow \infty$ in the definition of the path integral has not yet been taken. We have

$$\left\langle \frac{\partial F}{\partial x_k} \right\rangle_S = -\frac{i}{\hbar} \left\langle F \frac{\partial S}{\partial x_k} \right\rangle_S \quad (8.7)$$

Note that the expectation value—the path integral—will now involve only a finite number of integrations. Writing

$$S = \int \left[\frac{1}{2} m \dot{x}^2 - V(x) \right] d\tau \rightarrow \sum_{k=1}^N \left[\frac{1}{2} \frac{mN}{t} (x_{k+1} - x_k)^2 - \frac{V(x_k)t}{N} \right] \quad (8.8)$$

Equation 8.7 becomes ($\epsilon = t/N$)

$$\left\langle \frac{\partial F}{\partial x_k} \right\rangle_S = \frac{ie}{\hbar} \left\langle F \left\{ m \left[\frac{x_{k+1} - 2x_k + x_{k-1}}{\epsilon^2} \right] + \frac{\partial V}{\partial x}(x_k) \right\} \right\rangle_S \quad (8.9)$$

A number of interesting results can be obtained by making various choices for F . Let $F \equiv 1$. The left side of (8.9) vanishes and we obtain (returning to continuum notation)

$$\langle m \ddot{x} \rangle_S = - \left\langle \frac{\partial V}{\partial x} \right\rangle_S \quad (8.10)$$

for all times (since k is arbitrary). This is a version of the Ehrenfest theorem.

Exercise: In comparing (8.10) to the standard quantum mechanical result, what sense can you make of \dot{x} ? Can you justify the relation between the expression $\langle(x_{k+1} - x_k)/\epsilon\rangle_s$ and the equation $\dot{x} = i[H, x]/\hbar$?

HINT: Recall that $\langle x_k \rangle_s = \langle x(t_k) \rangle_s = \langle x_{\text{final}} | \exp[-i(t-t_k)H/\hbar] x_{\text{operator}} \times \exp[-it_k H/\hbar] | x_{\text{initial}} \rangle$.

Now let F be x_k for some particular k . Then (8.9) implies

$$\langle 1 \rangle_s = \frac{i}{\hbar} \left\langle mx_k \left[\frac{x_{k+1} - x_k}{\epsilon} - \frac{x_k - x_{k-1}}{\epsilon} \right] + \epsilon x_k V'(x_k) \right\rangle_s \quad (8.11)$$

For $\epsilon \rightarrow 0$ the term involving V' vanishes, and we have the result that the expectation value

$$m \left\langle x_k \frac{(x_{k+1} - x_k)}{\epsilon} \right\rangle_s \quad (8.12)$$

is different from

$$m \left\langle x_k \frac{(x_k - x_{k-1})}{\epsilon} \right\rangle_s \quad (8.13)$$

They differ only in the times at which x and $m\Delta x/\epsilon = m(x_{k+1} - x_k)/\epsilon$ are to be evaluated. Furthermore, the fact that they differ by 1 suggests that $m(x_{k+1} - x_k)/\epsilon$ be identified with the momentum operator and that (8.11) be interpreted as a commutation relation. If operators are ordered so that that with the earliest time argument operates first, then (8.11) is the familiar relation $-i\hbar = [p, x]$.

The calculational rules for working with this “momentum operator” can be tricky, as illustrated by the following example. We expect kinetic energy to be given by

$$\left\langle \frac{p^2}{2m} \right\rangle_s \quad (8.14)$$

However, it will not do to evaluate

$$\frac{1}{2}m \left\langle \left(\frac{x_{k+1} - x_k}{\epsilon} \right)^2 \right\rangle_s \quad (8.15)$$

We already have experience with such integrals and know that $\langle (x_{k+1} - x_k)^2 \rangle_S$ will give an ϵ , so that dividing by ϵ^2 in (8.15) will give an infinite result. The trick is to identify kinetic energy with

$$\lim_{\epsilon \downarrow 0} \frac{1}{2m} \langle p(t+\epsilon)p(t) \rangle_S \quad (8.16)$$

This will lead to consideration of

$$\frac{1}{2}m \left\langle \left(\frac{x_{k+1} - x_k}{\epsilon} \right) \left(\frac{x_k - x_{k-1}}{\epsilon} \right) \right\rangle_S \quad (8.17)$$

which is in fact the correct form for kinetic energy.

Apropos of the present discussion it is worth mentioning that some people have felt that the Feynman path integral formalism will solve the ordering problem for operators. For example, given the classical quantity p^5x^3 a number of quantum mechanical operators can be associated with it. But one can turn to the path integral and observe that because p^5x^3 is a well defined classical quantity, when it is put into the path integral some definite choice of operators will be made. Various schemes of this sort have been proposed but since they all founder on the same basic point, there is no harm done if we look at a particularly naive proposal. Thus let us try

$$\langle p^5x^3 \rangle_S \quad (8.18)$$

and ask if this is a unique object obtained from the classical mechanics. In fact, we know this is not unique, since we have seen that

$$\lim_{\epsilon \downarrow 0} \langle p(t+\epsilon)x(t) \rangle_S \neq \lim_{\epsilon \downarrow 0} \langle x(t+\epsilon)p(t) \rangle_S \quad (8.19)$$

Moreover, the example of kinetic energy shows the arguments in (8.18) cannot all be taken at the same time. Therefore the particular choice that is made for the time sequencing of the p 's and x 's will determine the operator; different choices of sequence will lead to different operators. In the notes to Sections 27 and 31 are references on the use of path integrals in the operator ordering problem.

Brownian Motion and the Wiener Integral; Kac's Proof

The connection between path integration and Brownian motion is so close that they are nearly indistinguishable. Unfortunately though, like a body and its mirror image, the sum over paths for Brownian motion is a theory having substance, while its path integral image exists mainly in the eye of the beholder.

We shall go through a quick review of Brownian motion, taking an elementary approach. This will lead us to the Wiener integral which is a bona fide integral on the space of paths. As an added bonus we shall get some insight into the connection between probability and potential theory, a subject of current mathematical interest.

Consider a discrete random walk on a line in which all steps have equal length Δ , take place at time intervals ϵ , and have probability p of going to the right, probability q of going to the left, $p+q=1$. Let $x=j\Delta$, $t=N\epsilon$ and $u(j, N)$ be the probability of the system getting to x (j) at time t (N), having started at position 0 at time 0. The familiar line of reasoning is as follows: To get a net distance of j steps to the right means taking μ steps to the right and ν to the left where

$$\begin{aligned}\mu - \nu &= j \\ \mu + \nu &= N\end{aligned}\tag{9.1}$$

This can occur in $\binom{N}{\mu}$ ways, so that for $N-j$ even

$$u(j, N) = \binom{N}{\mu} p^\mu q^\nu\tag{9.2}$$

while for $N-j$ odd, u is zero. Using Stirling's approximation

$$n! = \int_0^\infty e^{-t} t^n dt \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (9.3)$$

a straightforward calculation leads to

$$u(j, N) = \sqrt{\frac{2}{\pi N}} \exp\left[-\frac{(j-\alpha)^2}{2N}\right] \quad (9.4)$$

where $\alpha = N(p-q)$ and j and N are assumed to be large. Consider now $\rho(x, t)$, the density of particles per unit x interval. This is just $u(j, N)/2\Delta$ (the "2" comes because $u(j, N)$ is zero for $N-j$ odd). Equation 9.4 can thus be written in terms of x and t as

$$\rho(x, t) = \sqrt{\frac{\epsilon}{2\pi t \Delta^2}} \exp\left[-\frac{1}{2t} \left(\frac{\epsilon}{\Delta^2}\right)(x-\alpha\Delta)^2\right] \quad (9.5)$$

Our interest is in describing continuous time Brownian motion (an idealization closer to the physical situation), and we wish to let $\Delta \rightarrow 0$ and $\epsilon \rightarrow 0$ while holding x and t finite. The only way this can be done so as to keep ρ finite and nontrivial is by keeping the ratio ϵ/Δ^2 finite. Define, therefore, the diffusion coefficient D to be

$$D = \frac{\Delta^2}{2\epsilon} \quad (9.6)$$

with this ratio to be held constant. Furthermore, since

$$\alpha\Delta = t \frac{\Delta}{\epsilon} (p-q) \quad (9.7)$$

the difference $p-q$ must also go to zero if a finite result is demanded. Defining

$$c = \frac{\Delta}{\epsilon} (p-q) \quad (9.8)$$

the final result is

$$\rho(x, t) = \sqrt{\frac{1}{4\pi D t}} \exp\left[-\frac{(x-ct)^2}{4Dt}\right] \quad (9.9)$$

The particle is thus undergoing a diffusion process whose mean position moves with drift velocity c . It is also easy to verify that ρ satisfies the partial differential equation

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} - c \frac{\partial \rho}{\partial x} \quad (9.10)$$

For further work we drop the drift (set $c=0$) and use the density

$$\rho(x, t) = \sqrt{\frac{1}{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad (9.11)$$

The next step is to describe the sum over paths approach to Brownian motion as developed by Wiener. If $x=0$ at $t=0$, then at time t the probability that the particle's position is between a and b is

$$\int_a^b \rho(x, t) dx \quad (9.12)$$

Similarly, the probability that a particle starting at $(0, 0)$ is between a_1 and b_1 at time t_1 , is between a_2 and b_2 at time t_2, \dots , and is between a_N and b_N at time t_N ($a_i < b_i, t_i < t_{i+1}$) is

$$\int_{a_1}^{b_1} dx_1 \cdots \int_{a_N}^{b_N} dx_N \rho(x_1, t_1) \rho(x_2 - x_1, t_2 - t_1) \cdots \rho(x_N - x_{N-1}, t_N - t_{N-1}) \quad (9.13)$$

The appearance of the arguments $x_i - x_{i-1}$ and $t_i - t_{i-1}$ reflects of course the homogeneity in space and in time of this diffusion process.

At this stage we shall part company from the mathematicians and drop the integration over the last variable x_N . For us the quantity of interest is the conditional probability that the particle satisfies $a_i < x(t_i) < b_i$ at $t_i, i=1, \dots, N-1$ given that it arrives at $(x, t) \equiv (x_N, t_N)$ and that it started at $(0, 0)$. This conditional probability (density) is given by (9.13) but without the integral over x_N and divided by $\rho(x_N, t_N)$. (Recall that the conditional probability of an event A given an event B , $P(A|B)$, is given by $P(A|B) = P(A \text{ and } B)/P(B)$. In the paper of Gel'fand and Yaglom will be found the more usual object of mathematical concern, namely that in which the final position is not fixed.) Figure 9.1 shows a graphical interpretation of the quantity. The integral (9.13) (without $\int dx_N$) is the probability that the path of the particle passed through each of the given portals at the associated times.

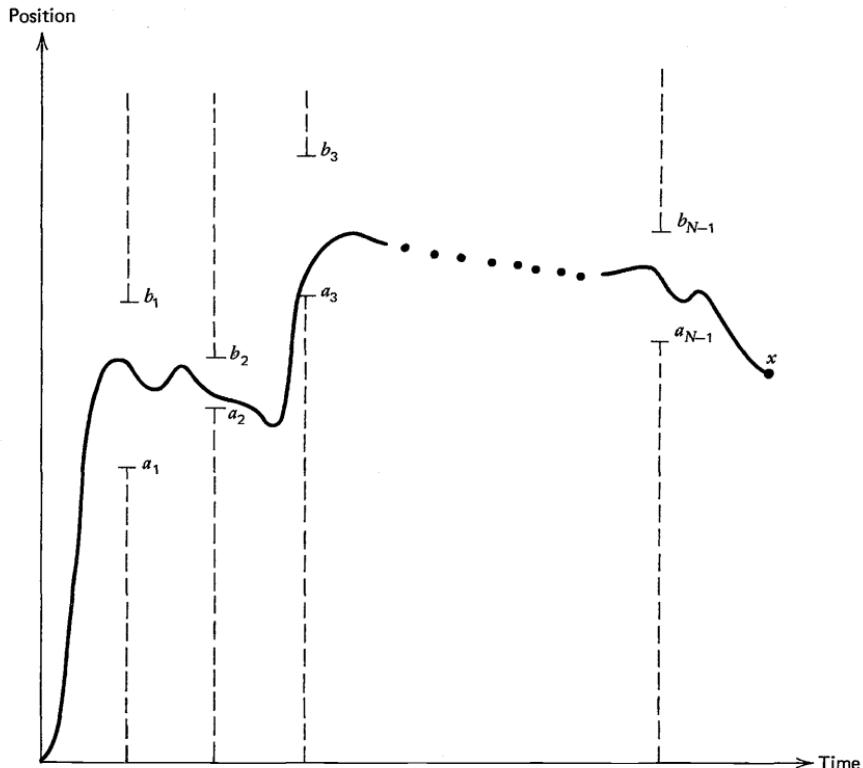


Fig. 9.1 A trajectory satisfying the conditions $x(0)=0$, $a_i < x(t_i) < b_i$, $i=1,\dots,N-1$, $x(t)=x$.

One can imagine that as we increase the number of times at which a specification of (a, b) is made and as $|b - a| \rightarrow 0$ the path of the particle is pinned down more and more precisely. Of course, if this intuition is to have any chance of being implemented it is important to restrict attention to continuous trajectories. Thus for $N \rightarrow \infty$ we obtain a measure on the space of paths and this is called Wiener measure. With this measure is defined the Wiener integral.

To connect (9.13) and Fig. 9.1 to the usual notions of measure theory it is necessary to identify (1) on what space the measure is defined, (2) what is the σ -algebra of measurable sets [recall that for Lebesgue measure on \mathbb{R} these are just countable unions and finite intersections of intervals (c, d)], and (3) what is the measure of a set in that σ -algebra [for Lebesgue

measure on \mathbb{R} , $\mu_2((c, d)) = d - c$ for $d \geq c$. The space on which we impose the proposed measure is that of continuous functions $x(\cdot)$ with $x(0) = 0$ and either $x(t)$ arbitrary or $x(t) = x$, the latter yielding conditional Wiener measure. The σ -algebra is built of "cylinder sets" defined as follows:

$$K[a_1, b_1, t_1; a_2, b_2, t_2; \dots; a_n, b_n, t_n] \\ = \{x(\cdot) | a_i < x(t_i) < b_i, \quad i=1, \dots, N\}$$

The measure of such a set is just the number appearing in (9.13). (For the conditional measure one uses

$$\frac{\int_{a_1}^{b_1} dx_1 \cdots \int_{a_{N-1}}^{b_{N-1}} dx_N \rho(x_1, t_1) \cdots \rho(x_N - x_{N-1}, t_N - t_{N-1})}{\int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_N \rho(x_1, t_1) \cdots \rho(x_N - x_{N-1}, t_N - t_{N-1})} \quad (9.13a)$$

and the definition of the cylinder set is also modified appropriately.) To obtain the σ -algebra we form countable unions and intersections of cylinder sets and the measure is defined as in the usual measure theory. Various general measure theoretic theorems can be invoked to prove that the resulting object, what we have called the Wiener measure, is indeed a measure. A fact that has been of considerable relevance in many of our calculations can now be stated as: almost all (in the sense of Wiener measure) continuous functions are nowhere differentiable.

We could write the Wiener integral as $\int dx(\tau)$ as we have been doing for the path integral. Alternatively, one might wish to write $\int d\mu(x(\tau))$. In fact we will adopt a probabilistic point of view and look at the integral as an expectation value, E . Thus suppose Q is some function of the paths $x(\tau)$. Then

$$E_{xt}(Q) = \sum_q q P'_w\{x(\tau) \in \Omega | Q[x(\tau)] = q\} \quad (9.14)$$

where Ω is the space of paths from $(0, 0)$ to (x, t) , $P'_w\{x \in \Omega | Q(x) = q\}$ is the unnormalized Wiener measure of the set of paths on which Q takes the value q [i.e., one uses only the numerator in (9.13a)]. The subscripts x and t on E refer to the fact that an unnormalized and conditional expectation is being evaluated. In practice (9.14) would be evaluated in a manner identical to that used for the Feynman integral. Let $Q(x_1, \dots, x_{N-1})$ be the value Q takes on the broken line path from $(0, 0)$ to (x_1, t_1) to \dots to (x_{N-1}, t_{N-1}) .

t_{N-1}) to (x, t) with $t_j = jt/N$. Then

$$\begin{aligned} E_{xt}(Q) &= \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \cdots \int_{-\infty}^{\infty} dx_{N-1} \rho(x_1, t_1) \rho(x_2 - x_1, t_2 - t_1) \cdots \\ &\quad \times \rho(x - x_{N-1}, t - t_{N-1}) Q(x_1, \dots, x_{N-1}) \end{aligned} \quad (9.15)$$

E_{xt} 's lack of normalization arises from the absence of an x_N integration in (9.15), and this circumstance is precisely expressed through the following equation:

$$E_{xt}(1) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad (9.16)$$

Correcting the normalization would be easy, but we prefer to work with the object E_{xt} .

Let $U(x)$ be some function of x . We have seen (Section 4) how Feynman showed that his path integral propagator satisfied the Schrödinger equation; now we give a proof by Kac of a very similar result about the quantity

$$W(x, t) \equiv E_{xt} \left[\exp\left(-\int_0^t U(x(\tau)) d\tau\right) \right] \quad (9.17)$$

A basic fact about expectation values (or integrals) is their linearity, so that

$$W(x, t) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} E_{xt} \left[\left(\int_0^t U(x(\tau)) d\tau \right)^j \right] \quad (9.18)$$

Kac takes pains to work first with a bounded function U so as to ensure the validity of the operations being performed. We shall not be so careful, but refer the interested reader to his book for details of the rigorous proof.

Consider the various terms in the sum over j . For $j=0$ we have

$$E_{xt}(1) = \frac{e^{-x^2/4Dt}}{\sqrt{4\pi Dt}} \equiv W_0(x, t) \quad (= \rho(x, t)) \quad (9.19)$$

where we have defined the function W_0 . For $j=1$ we must consider

$$W_1(x, t) \equiv E_{xt} \left[\int_0^t U(x(\tau)) d\tau \right] = \int_0^t E_{xt}[U(x(\tau))] d\tau \quad (9.20)$$

The expected value of $U(x(\tau))$ is just the probability of $x(\tau)$ taking some

value ξ times $U(\xi)$, summed over ξ . Thus

$$\begin{aligned} W_1(x, t) &= \int_0^t d\tau \int_{-\infty}^{\infty} d\xi \rho(x - \xi, t - \tau) U(\xi) \rho(\xi, \tau) \\ &= \int_0^t d\tau \int d\xi \rho(x - \xi, t - \tau) U(\xi) W_0(\xi, \tau) \end{aligned} \quad (9.21)$$

For $j=2$ there are two integrals, and it is convenient to define

$$W_2(x, t) \equiv \frac{1}{2} E_{xt} \left[\int_0^t U(x(\tau_1)) d\tau_1 \int_0^t U(x(\tau_2)) d\tau_2 \right] \quad (9.22)$$

We next invoke the following well known formula:

LEMMA

$$\left[\int_0^t f(\tau) d\tau \right]^N = N! \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{N-1}} d\tau_N f(\tau_1) f(\tau_2) \cdots f(\tau_N) \quad (9.23)$$

Proof is by induction on N . For $N=1$, (9.23) is obviously true. Let

$$g_N(t) = \left[\int_0^t f(\tau) d\tau \right]^N \quad (9.24)$$

Then $dg_N(s)/ds = N g_{N-1}(s) f(s)$ for all s . When s is integrated from 0 to t and the inductive hypothesis for $(N-1)$ used, (9.23) follows.

Equation 9.22 thus becomes

$$W_2(x, t) = \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 E_{xt} [U(x(\tau_1)) U(x(\tau_2))] \quad (9.25)$$

Noting that $\tau_1 \geq \tau_2$ and using the same reasoning that led to (9.21) we obtain

$$\begin{aligned} W_2(x, t) &= \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \int \int d\xi d\eta \rho(x - \xi, t - \tau_1) U(\xi) \rho(\xi - \eta, \tau_1 - \tau_2) \\ &\quad \times U(\eta) \rho(\eta, \tau_2) \\ &= \int_0^t d\tau_1 \int d\xi \rho(x - \xi, t - \tau_1) U(\xi) W_1(\xi, \tau_1) \end{aligned} \quad (9.26)$$

It should now be clear that if we let

$$W_j(x, t) \equiv \frac{1}{j!} E_{xt} \left[\left(\int_0^t U(x(\tau)) d\tau \right)^j \right] \quad (9.27)$$

it will follow that

$$W_j(x, t) = \int_0^t d\tau \int d\xi \rho(x - \xi, t - \tau) U(\xi) W_{j-1}(\xi, \tau) \quad j = 1, 2, \dots \quad (9.28)$$

Returning to (9.17), W can be written

$$\begin{aligned} W(x, t) &= \sum_{j=0}^{\infty} (-1)^j W_j(x, t) \\ &= W_0(x, t) - \sum_{j=1}^{\infty} (-1)^{j-1} \int_0^t d\tau \int d\xi W_0(x - \xi, t - \tau) U(\xi) W_{j-1}(\xi, \tau) \\ &= W_0(x, t) - \int_0^t d\tau \int d\xi W_0(x - \xi, t - \tau) U(\xi) W(\xi, \tau) \end{aligned} \quad (9.29)$$

Recalling that $(\partial/\partial t - D\partial^2/\partial x^2)W_0 = 0$ and that for $t \downarrow 0$ $W_0(x, t)$ is a delta function, this integral equation for W implies that W satisfies the differential equation:

$$\frac{\partial W(x, t)}{\partial t} = D \frac{\partial^2 W(x, t)}{\partial x^2} - U(x) W(x, t) \quad (9.30)$$

The parallel structures of path integration and Wiener integration should, with (9.30), be overwhelmingly manifest. Equation 9.29 can be recognized as a familiar formula from time dependent perturbation theory and as being a Fourier transform of the Lippmann-Schwinger equation (with $t \rightarrow -it$).

The formulas of this section establish ties between probability and potential theory. Consider for example the problem of a point particle undergoing Brownian motion in one dimension (without drift). Suppose it starts at $x=0$ at $t=0$ and we wish to know its “first passage time” for reaching $x=+a$. Phrased slightly differently, we wish to calculate the probability that its trajectory $x(\tau)$ satisfies $x(\tau) \leq a$ for all τ , $0 \leq \tau \leq t$. Call this probability $p(t)$.

For this problem it is easier to work with the unconditional Wiener measure which is gotten by integrating over the variable x in E_{xt} . This

integral, call it E_t , is the sum over paths that start at $(0,0)$ and reach anywhere at time t . Thus

$$E_t(Q) = \int_{-\infty}^{\infty} dx E_{xt}(Q) \quad (9.31)$$

From (9.16) it is evident that E_t is correctly normalized, so that $E_t(1)=1$. Define

$$Q(x(\tau)) = \exp\left(-\int_0^t V_a(x(\tau)) d\tau\right) \quad \text{with} \quad V_a(x) = \begin{cases} 0 & x < a \\ +\infty & x \geq a \end{cases} \quad (9.32)$$

The functional Q vanishes for any path that goes to the right of $x=a$ and is 1 otherwise. Therefore $p(t)$, the probability of the particle never having gone to the right of a , is simply

$$p(t) = E_t(Q) = \int_{-\infty}^{\infty} dx E_{xt}\left[\exp\left(-\int_0^t V_a(x(\tau)) d\tau\right)\right] \quad (9.33)$$

However, we also know that the expectation

$$E_{xt}\left(\exp\left(-\int V_a d\tau\right)\right) \equiv W(x, t) \quad (9.34)$$

satisfies (9.30) with $U=V_a$. Now the Green's function for Schrödinger's equation, with the same potential V_a , satisfies ($\hbar=1$)

$$i \frac{\partial G(x, s; 0)}{\partial s} = -D \frac{\partial^2 G}{\partial x^2} + V_a G \quad (9.35)$$

and both G and W approach delta functions as $t \rightarrow 0$. Comparing (9.35) and (9.30) it follows that

$$G(x, -it; 0) = W(x, t) \quad (9.36)$$

To calculate G one could expand in terms of eigenfunctions. However, we can use path integration and the method of images to get a more explicit result. Path integral experience suggests looking for classical paths that start at $(0,0)$ and reach (x, t) with a potential that is zero for $x < a$ and an infinite hard wall at $x = a$. There are two such paths, one (a) going directly with velocity (x/t) and the other (b) bouncing off the wall and

having velocity $(2a-x)/t$. The classical action for each is (the particle's effective mass is $1/2D$)

$$S_a = \frac{x^2}{4Dt}, \quad S_b = \frac{(2a-x)^2}{4Dt} \quad (9.37)$$

Should it happen that the WKB approximation is exact then a reasonable guess for G would be

$$G(x, t; 0) = \sqrt{\frac{1}{4\pi D t}} \left[\exp\left(\frac{ix^2}{4Dt}\right) - \exp\left(i\frac{(2a-x)^2}{4Dt}\right) \right] \quad (9.38)$$

(for $x > a$, G is zero). The only feature of (9.38) so far unmotivated is the minus sign for the second exponential term. The reason for the minus is that the infinite potential at $x=a$ forces the wave functions (hence the Green's function) to vanish there, that is, there are Dirichlet boundary conditions at $x=a$. The relative sign guarantees that $G(a, t; 0)=0$. But now we claim that G of (9.38) is in fact the exact propagator. This is because its two summands clearly satisfy (9.35) individually, together they satisfy the boundary condition at a , and finally for $t \rightarrow 0$ the first term becomes a delta function and the second term drops out (since $x=2a$ is impossible). Equation 9.38 is essentially an application of the method of images.

Using (9.36) we substitute (9.38) in (9.33) to get finally

$$\begin{aligned} p(t) &= \int_{-\infty}^a dx \sqrt{\frac{1}{4\pi D t}} \left[\exp\left(-\frac{x^2}{4Dt}\right) - \exp\left(-\frac{(2a-x)^2}{4Dt}\right) \right] \\ &= 2\sqrt{\frac{1}{4\pi D t}} \int_0^a \exp\left(-\frac{x^2}{4Dt}\right) dx \end{aligned} \quad (9.39)$$

Obviously the method used here could be generalized to deal with all sorts of questions about confinement in various regions. In Section 28 we shall apply it to a problem involving bosons with random impurities.

Returning to the comparison of the path integral and the Wiener integral, we consider the following question. Can the Wiener integral which, as has been pointed out, provides a bona fide measure on the space of paths, be used to further the cause of making an honest integral out of the Feynman path integral? Can the path integral be taken beyond its derivation as a glorified use of the Trotter product formula?

First let us see how we might try to go from the Wiener integral to the Feynman integral. Let the function U above be pure imaginary and write $U(x) = iV(x)$. This costs nothing. Now consider $W(x, t)$ as a function of D

(the diffusion coefficient) also, with D allowed to become complex. We have found that for D real and positive the Wiener measure is fully well defined. In particular, expectations are evaluated as in (9.15) with the appropriate D appearing in

$$\rho(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}$$

Now let D become complex. A great deal can be gotten from Wiener integral formulas by analytic continuation, in particular the fact that the analytically continued object satisfies Schrödinger's equation. Unfortunately, though, once D ceases to be pure real the continued object ceases to be a functional integral.

In recent years there have been efforts to develop more general or weaker kinds of measure so as to encompass the path integral (i.e., imaginary D). C. DeWitt-Morette has been exploring various kinds of "pseudomeasures" in order to justify the very useful picture of the path integral as a sum over paths. S. Albeverio and R. Hoegh-Krohn and A. Truman have also provided rigorous approaches to the path integral. Truman in particular defines a Hilbert space of paths which in some sense are the polygonal paths to be summed over according to Feynman's original conception of the path integral. We shall not go into the details of this work as it is outside the main thrust of this book. However, if you find yourself uncomfortable with the balancing act done throughout the book — that is, our attempts to draw firm conclusions from not so firm a theory — you may be interested in pursuing these matters. Some references are given below.

Exercise: A particle undergoing Brownian motion in one dimension with diffusion coefficient D is somewhere in the interval $[0, 1]$, at $t=0$. At that time it is equally likely to be anywhere in the interval. On the average, how much time will elapse before it first leaves the interval?

HINT: Exercise 3, Section 7.

NOTES

Much of the material of this section is taken from

M. Kac, *Probability and Related Topics in the Physical Sciences*, Interscience, New York, 1959

A popular presentation of the connection between probability and potential theory appears in

R. Hersh and R. J. Griego, "Brownian Motion and Potential Theory," *Sci. Amer.* **220**, No. 3, 67 (March 1967)

An important general reference which includes material touching on the topics of this section is

I. M. Gel'fand and A. M. Yaglom, *J. Math. Phys.* **1**, 48 (1960)

The analytic continuation of the Wiener integral formulas to the Feynman integral is done by Nelson in

E. Nelson, *J. Math. Phys.* **5**, 332 (1964)

The fact that the analytically continued propagator is not a functional integral is shown in

R. H. Cameron, *J. Math. and Phys.* **39**, 126 (1960)

Papers on pseudomeasures and similar matters are

C. Morette DeWitt, *Commun. Math. Phys.* **28**, 47 (1972)

C. DeWitt-Morette, *Commun. Math. Phys.* **37**, 63 (1974)

Other approaches to a rigorous path integral theory are those of

L. L. Lee, Continuum Calculus and Feynman's Path Integrals, *J. Math. Phys.* **17**, 1988 (1976)

S. A. Albeverio and R. J. Hoegh-Krohn, *Mathematical Theory of Feynman Path Integrals*, Lecture Notes in Mathematics No. 523, Springer Verlag, Berlin, 1976

A. Truman, The Feynman Maps and the Wiener Integral, *J. Math. Phys.* **19**, 1742 (1978)

Truman is particularly concerned with retaining Feynman's polygonal path picture within a rigorous context.

Perturbation Theory and Feynman Diagrams

“Feynman diagrams” provide a graphic, intuitive, and useful way of expressing perturbation theory. These diagrams emerge naturally from the path integral expression for the propagator through techniques similar to those used by Kac in his proof that the conditional expectation satisfies the diffusion equation.

The mathematics to be used in obtaining the Feynman diagrams is almost identical to that used in the preceding section, except that now we deal with the image—the path integral—rather than the substance—the Wiener measure. As in (9.17) and (9.18) we expand the Green’s function

$$\begin{aligned} G(x, t; y, 0) &= \int_{y, 0}^{x, t} dx(\tau) \exp \left[\frac{i}{\hbar} \int_0^t \left(\frac{m\dot{x}^2}{2} - V(x) \right) d\tau \right] \\ &= \int_{y, 0}^{x, t} dx(\tau) \exp \left[\frac{i}{\hbar} \int_0^t \frac{m\dot{x}^2}{2} \right] \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{-i}{\hbar} \right)^j \left(\int_0^t V(x(\tau)) d\tau \right)^j \end{aligned} \quad (10.1)$$

In (10.1) $\exp[i \int_0^t (m\dot{x}^2/2\hbar) d\tau]$ appears as part of the integrand. In the notation of Section 9 (which is the usual notation in the mathematical literature) the factor $\rho(x, t) = (4\pi Dt)^{-1/2} \exp(-x^2/4Dt)$ was absorbed into the measure and did not explicitly appear when an expectation was written.

The first term ($j=0$) in the expansion (10.1) is

$$\begin{aligned} G_0(x, t; y, 0) &\equiv \int_{y, 0}^{x, t} dx(\tau) \exp \left[\frac{i}{\hbar} \int_0^t \frac{m\dot{x}^2}{2} d\tau \right] \\ &= \sqrt{\frac{m}{2\pi i\hbar t}} \exp \left[\frac{im}{2\hbar t} (x-y)^2 \right] \end{aligned} \quad (10.2)$$

which is the free particle propagator. The $j=1$ term contains the expression

$$G_1(x, t; y, 0) \equiv \int_{y, 0}^{x, t} dx(\tau) \exp \left[\frac{im}{2\hbar} \int_0^t \dot{x}^2 d\tau \right] \int_0^t V(x(\tau)) d\tau \quad (10.3)$$

As in the Kac proof we interchange the order of integration and do the path integration prior to the τ integration. At this stage in the Kac proof we argued probabilistically to get (9.21), but now we use slightly different reasoning. We appeal to the definition of the path integral as a limit of a multiple integral; let τ (dummy variable after order of integration has been changed in (10.3)) be t_j , and let $x(\tau)$ be x_j . This gives

$$\begin{aligned} G_1(x, t; y, 0) &= \sum_j (\Delta t) \int \prod_{i=1}^N dx_i \left(\frac{m}{2\pi i \hbar \Delta t} \right)^{(N+1)/2} \\ &\times \exp \left[\frac{im}{2\hbar \Delta t} \sum_{k=0}^N (x_{k+1} - x_k)^2 \right] V(x_j) \end{aligned} \quad (10.4)$$

The integrals for $k=1, \dots, j-1$ and $k=j+1, \dots, N$ are just free particle integrals. If x_j is called ξ we have, as in Section 9,

$$G_1(x, t; y, 0) = \int_0^t d\tau \int d\xi G_0(x - \xi, t - \tau) V(\xi) G_0(\xi - y, \tau) \quad (10.5)$$

Define

$$G_k(x, t; y, 0) \equiv \frac{1}{k!} \int_{y, 0}^{x, t} dx(\tau) \exp \left[\frac{im}{2\hbar} \int_0^t \dot{x}^2 d\tau \right] \left[\int_0^t V(x(\tau)) d\tau \right]^k \quad (10.6)$$

and as in Section 9, it follows that

$$\begin{aligned} G_k(x, t; y, 0) &= \int_{y, 0}^{x, t} dx(\tau) \exp \left[\frac{im}{2\hbar} \int_0^t \dot{x}^2 d\tau \right] \\ &\times \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{k-1}} d\tau_k V(x(\tau_1)) \cdots V(x(\tau_k)) \\ &= \int_0^t d\tau \int d\xi G_0(x, t; \xi, \tau) V(\xi) G_{k-1}(\xi, \tau; y, 0) \end{aligned} \quad (10.7)$$

As before, we can immediately establish that

$$G(x, t; y, 0) = G_0(x, t; y, 0) - \frac{i}{\hbar} \int_0^t d\tau \int d\xi G_0(x, t; \xi, \tau) V(\xi) G(\xi, \tau; y, 0) \quad (10.8)$$

which is the formulation of time dependent perturbation theory. [Incidentally, in this derivation V could have an explicit t dependence, in which case $V(\xi, \tau)$ would appear in (10.8).] Now, however, we examine and interpret individual terms in the expansion

$$G(x, t; y, 0) = \sum_{j=0}^{\infty} \left(\frac{-i}{\hbar} \right)^j G_j(x, t; y, 0) \quad (10.9)$$

The first interpretive step is to consider each summand in (10.9) to be an amplitude for a specific way of getting from $(y, 0)$ to (x, t) —inasmuch as the propagator $G(x, t; y, 0)$ itself is the total amplitude for getting there. $G_0(x, t; y, 0)$ is the amplitude for going from $(y, 0)$ to (x, t) as a free particle and can thus be interpreted as the amplitude that the particle makes the trip completely unaffected by the potential V . G_1 , given in (10.5), is itself an integral over τ and ξ and can be considered the sum of amplitudes for various ways of making the trip. Namely [see (10.5)] the particle propagates freely to (ξ, τ) with amplitude $G_0(\xi, \tau; y, 0)$, then is hit

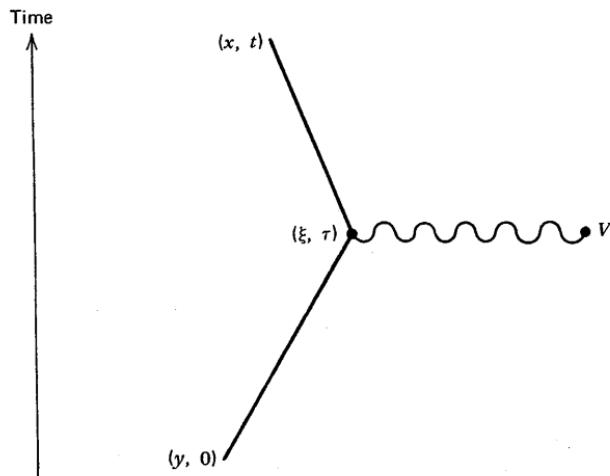


Fig. 10.1 Feynman diagram for contribution of first order perturbation theory.

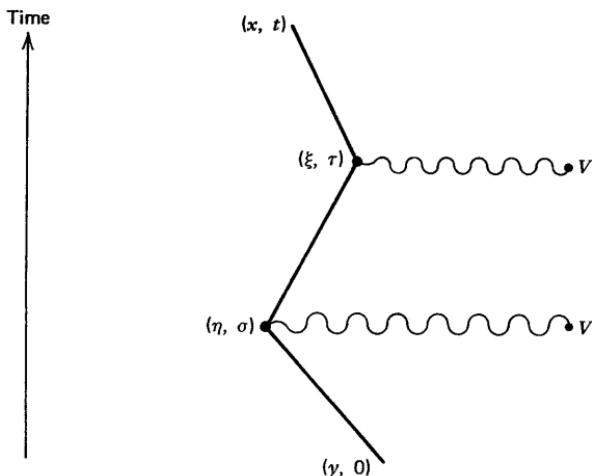


Fig. 10.2 Feynman diagram for contribution of second order perturbation theory.

by the potential $V(\xi)$, and finally goes from (ξ, τ) to (x, t) freely. Then this amplitude is added for each possible position ξ and time τ . One of Feynman's big steps was to provide a pictorial representation of this sequence of events. Figure 10.1 is a schematic diagram (time running upwards) with a heavy straight line segment representing free propagation. The wiggly line from the symbol "V" to the break in the free propagation (i.e., the vertex of the heavy lines) represents the interaction with the potential, and in the field theory version of perturbation theory will actually correspond to the free propagation of some other particle, such as the photon. The term $G_2(x, t; y, 0)$ can similarly be written out as successive stretches of free propagation and (two) interactions with the potential. The Feynman diagram for G_2 is given in Fig. 10.2. Again, there will be integrals over η , σ , ξ , and τ to include all possible times and places that the interaction could have occurred. It should be obvious how successive G_j , $j = 3, 4, \dots$, are represented by diagrams.

It is worth mentioning that other perturbation expansions besides that given in (10.1) are easily realized. In particular, if V is written as $(V - U) + U$, the expansion (10.1) takes the form

$$G(x, t; y, 0) = \int dx(\tau) \exp \left[\frac{i}{\hbar} \int_0^t \left(\frac{m\dot{x}^2}{2} - U(x(\tau)) \right) d\tau \right] \\ \times \sum_{j=0}^{\infty} \left(\frac{-i}{\hbar} \right)^j \frac{1}{j!} \left[\int_0^t (V - U) d\tau \right]^j \quad (10.10)$$

The entire foregoing development goes through, but G_0 should now be defined as

$$G_0(x, t; y, 0) \equiv \int_{y,0}^{x,t} dx(\tau) \exp \left[\frac{i}{\hbar} \int_0^t \left(\frac{1}{2} m \dot{x}^2 - U(x(\tau)) \right) d\tau \right] \quad (10.11)$$

and of course may not be amenable to explicit evaluation, as in (10.2). In (10.8) the V would be replaced by $(V - U)$ and there would also be corresponding modification of the Feynman diagram.

PART TWO

Selected Applications of the Path Integral

ELEVEN

Asymptotic Analysis

In the next few sections we will get deep into the WKB approximation. To do this properly we shall first review some ideas from the theory of asymptotic approximations (touched upon in Section 3) and also topics in classical mechanics and the calculus of variations that are not generally familiar to physicists.

First we establish the vocabulary of asymptotics. For two functions $f(x)$ and $g(x)$ consider limiting values as $x \rightarrow c$ for c some finite or infinite number. The small o-symbol is defined as follows:

If $f(x)/g(x) \rightarrow 0$ as $x \rightarrow c$, we write

$$f(x) = o(g(x)) \quad (x \rightarrow c) \quad (11.1)$$

The large O-symbol allows f to be relatively larger. It is defined as:

If $|f(x)/g(x)|$ is bounded as $x \rightarrow c$, we write

$$f(x) = O(g(x)) \quad (x \rightarrow c) \quad (11.2)$$

When $|f(x)/g(x)|$ is bounded on an interval $[a, c]$ we write $f = O(g)$, $x \in [a, c]$. In that case there exists a bound $M (= \sup |f/g|)$ such that

$$|f(x)| \leq M |g(x)| \quad x \in [a, c] \quad (11.3)$$

The equal signs in (11.1) and (11.2) take the meanings assigned to them by the respective definitions and differ from ordinary equal signs. For example, while it is true that $o(\varphi) + o(\varphi) = o(\varphi)$ and $o(\varphi) = O(\varphi)$, it is not true that $O(\varphi) = o(\varphi)$.

The concept of asymptotic series is due to Poincaré and we give a somewhat restricted definition:

The formal series

$$\sum_{n=0}^{\infty} a_n x^n \quad (11.4)$$

is said to be an asymptotic series for the function $f(x)$, as $x \rightarrow 0$ if for each N

$$|f(x) - \sum_{j=0}^N a_j x^j| = o(x^N) \quad (11.5)$$

In this case we write

$$f(x) \sim \sum_{n=0}^{\infty} a_n x^n \quad (11.6)$$

More generally asymptotic series are defined for any sequence of functions $\phi_n(x)$ such that $\phi_{n+1}(x) = o(\phi_n(x))$, $x \rightarrow c$ for some c .

There is a fundamental distinction between asymptotic series and convergent series. Convergence for a series like (11.4) means: given an ϵ and an x there is an N such that $|f(x) - \sum_{n=0}^M a_n x^n| < \epsilon$ for all $M > N$. For an asymptotic series it is ϵ and N that are given. Then if the series is asymptotic, there is a ξ such that for $x < \xi$, $|f(x) - \sum_{n=0}^N a_n x^n| < \epsilon x^N$.

An example of a nonconvergent asymptotic series will illustrate the above point. Let

$$F(t) = \int_0^\infty dx \exp(-x^2 - tx^4) \quad (11.7)$$

First observe that this is related to the parabolic cylinder function whose integral representation is

$$U(a, z) = \frac{\exp(-\frac{1}{4}z^2)}{\Gamma(\frac{1}{2} + a)} \int_0^\infty e^{-sz - (1/2)s^2} s^{a-1/2} ds \quad (11.8)$$

With some changes of variable it turns out that

$$F(t) = \frac{\sqrt{\pi} \exp(1/8t)}{2(2t)^{1/4}} U\left(0, \frac{1}{\sqrt{2t}}\right) \quad (11.9)$$

Consider $F(t)$ for $t \rightarrow 0$. In (11.7) we expand the part of the exponential containing t , and interchange summation and integration

$$F(t) \sim \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \int_0^{\infty} e^{-x^2} x^{4n} dx \quad (11.10)$$

The gamma function is

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt \quad (11.11)$$

and we obtain

$$\int_0^{\infty} e^{-x^2} x^{4n} dx = \frac{1}{2} \Gamma\left(2n + \frac{1}{2}\right) = \frac{(4n-1)! \sqrt{\pi}}{(2n-1)! 2^{4n}} \quad (11.12)$$

Thus

$$F(t) \sim \sum_{n=0}^{\infty} \frac{1}{2} \frac{1}{n!} \Gamma\left(2n + \frac{1}{2}\right) (-t)^n \quad (11.13)$$

Now $\Gamma(2n + \frac{1}{2})$ is roughly $(2n)!$ so that division by $n!$ still leaves a term of order $2^n n!$. The series in (11.13) therefore has zero radius of convergence and we have indeed been well advised to forgo the equality symbol in (11.10) and (11.13). We shall not prove that the series is asymptotic which would involve error estimates for the partial sums. Instead we shall use the series as a calculational tool to show its usefulness despite the divergence. The first few terms of (11.13) are

$$F(t) = \sqrt{\pi} \left[\frac{1}{2} - \frac{3t}{8} + \frac{105}{32} t^2 - \frac{10395}{128} t^3 + \dots \right] \quad (11.14)$$

which already give a good approximation for $t = 1/50$ (from five digit tables $U/\sqrt{\pi} = 0.49307$ and disagreement is in the fifth digit).

(As an aside, we note that this demonstration, worked in the opposite direction, provides an example of Borel resummation.)

A simpler example is

$$G(t) = \int_0^{\infty} \frac{e^{-xt}}{1+x} dx \quad (11.15)$$

The factor $(1+x)^{-1}$ is expanded as $1-x+x^2-\dots$ and these integrals

done to yield

$$G(t) \sim \frac{1}{t} - \frac{1!}{t^2} + \frac{2!}{t^3} - \frac{3!}{t^4} + \dots \quad (11.16)$$

This series obviously diverges, but let us consider its "predictions" for $t=10$:

$$G(10) = 0.1 - 0.01 + 0.002 - 0.0006 + \dots \quad (11.17)$$

The sum up to four terms is thus 0.0914. The correct value, known from other methods is 0.09156..., so our expansion has given quite reasonable results.

Exercise: Show that (11.16) is an asymptotic series (for G of (11.15)) in powers of $1/t$.

Exercise: By definition $1 - \text{erf}(x) = (2/\sqrt{\pi}) \int_x^\infty \exp(-t^2) dt$. Using integration by parts show that the following is an asymptotic series

$$\int_x^\infty e^{-t^2} dt = e^{-x^2} \left[\frac{1}{2x} - \frac{1}{2^2 x^3} + \frac{1 \cdot 3}{2^3 x^5} - \frac{1 \cdot 3 \cdot 5}{2^4 x^7} + \dots \right]$$

The moral is that if you do not take too many terms, an asymptotic series can give useful results.

Much is known about the validity of adding, multiplying, integrating, and differentiating asymptotic series. See the references given below.

One last observation is relevant to our later applications. Because $x^{-n} e^{-1/x} \rightarrow 0$ for any n as $x \rightarrow 0^+$ the asymptotic expansion for $e^{-1/x}$ is identically zero.

Using the vocabulary of the theory of asymptotic approximation let us look again at some of our previous results on the Laplace method and the method of stationary phase (see Section 3). Consider the integral

$$F(\lambda) = 2\sqrt{\frac{\lambda}{\pi}} \int_0^\infty e^{-\lambda(x^2 + f(x))} dx \quad (11.18)$$

for $\lambda \rightarrow \infty$ and where $f(x) = o(x^2)$, $f'(x) \geq 0$. The stationary point is $x=0$ and by the Laplace method we should have $F(\lambda) - 1 = o(1)$, ($\lambda \rightarrow \infty$). Let us check the result. Consider

$$\begin{aligned} \frac{1}{2}(1 - F(\lambda)) &= \sqrt{\frac{\lambda}{\pi}} \int_0^\infty dx e^{-\lambda x^2} (1 - e^{-\lambda f(x)}) \\ &= \int_0^a + \int_a^\infty \equiv \Delta_1 + \Delta_2 \end{aligned} \quad (11.19)$$

As indicated by the second equation of (11.19) we will make some judicious selection of a so as to allow convenient estimates. For positive $f(x)$,

$$0 < 1 - e^{-\lambda f(x)} \leq \lambda f(x) \quad (11.20)$$

(Proof: $\varphi(u) \equiv \exp(-u) - 1 + u = \int_0^u dx \int_0^x dy \exp(-y)$, so that $\varphi \geq 0$). It follows that

$$\Delta_1 \leq \sqrt{\frac{\lambda}{\pi}} \int_0^a \lambda f(x) e^{-\lambda x^2} dx \leq \frac{\lambda^{3/2}}{\sqrt{\pi}} \left(\int_0^\infty e^{-\lambda t^2} dt \right) \max_{0 < x < a} f(x) \quad (11.21)$$

Since $f(x) = o(x^2)$, there is a function η such that $\max_{0 < x < a} f(x) = a^2 \eta(a)$ with $\eta(a) = o(1)$, $a \rightarrow 0$. We have the following bound on Δ_1 :

$$\Delta_1 \leq \lambda a^2 \eta(a) \quad (11.22)$$

For Δ_2 we make the gross estimate of discarding $e^{-\lambda f(x)}$ and find that

$$\begin{aligned} \Delta_2 &\leq \sqrt{\frac{\lambda}{\pi}} \int_a^\infty e^{-\lambda x^2} dx = \sqrt{\frac{\lambda}{\pi}} \int_a^\infty \left(\frac{-1}{2\lambda x} \right) \frac{d}{dx} e^{-\lambda x^2} dx \\ &= -\sqrt{\frac{\lambda}{\pi}} \frac{1}{2\lambda} \frac{\exp(-\lambda x^2)}{x} \Big|_a^\infty - \sqrt{\frac{\lambda}{\pi}} \int_a^\infty e^{-\lambda x^2} \frac{1}{2\lambda x^2} dx \\ &\leq \frac{e^{-\lambda a^2}}{\sqrt{4\pi\lambda a^2}} \end{aligned} \quad (11.23)$$

We wish to make $\Delta_1 + \Delta_2 \rightarrow 0$ as $\lambda \rightarrow \infty$. Clearly a must go to zero a bit more slowly than $1/\sqrt{\lambda}$ (hence $\lambda a^2 \rightarrow \infty$) in order to make Δ_2 go to zero, while it must go fast enough for $\eta(a)\lambda a^2 \rightarrow 0$ also. Suppose $f(x) = x^4$, so that $\eta(a) = \text{const} \cdot a^2$. Clearly $a = (\log \lambda)/\sqrt{\lambda}$ will do the job.

The proof just given, while not lacking in rigor, is somewhat less elegant than those involving the change of variables

$$y^2 = x^2 + f(x) \quad (11.24)$$

The assumptions on the growth of f are enough to guarantee that this transformation is well defined (in some interval) and this neater method of proof is what is found in textbooks. The purpose of going through the estimates given above was to develop a feeling for which regions of integration are important in fixing the value of the integral.

NOTES

The following are good references for various aspects of asymptotic analysis:

- E. T. Copson, *Asymptotic Analysis*, Cambridge University Press, London, 1965
- N. G. De Bruijn, *Asymptotic Methods in Analysis*, North Holland, Amsterdam, 1961
- A. Erdelyi, *Asymptotic Expansions*, Dover, New York, 1956
- H. Jeffreys, *Asymptotic Approximations*, Oxford University Press, London, 1962

An excellent, more recent work is

- F. W. J. Olver, *Asymptotics and Special Functions*, Academic Press, New York, 1974
- There is also a shortened paperback version of this book that can be recommended entitled, *Introduction to Asymptotics and Special Functions*.
- For parabolic cylinder functions see Chapter 19 of

- M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, 1954

TWELVE

The Calculus of Variations

This section is devoted to topics in classical mechanics and the calculus of variations. The material we cover here is seldom studied by physicists, and in the mathematical literature the simpler facts are often hidden in the recesses of manifolds and tangent bundles. Specifically, our inquiry centers on the question of when a classical path is a minimum of the action and when it is merely an extremum.

The classical action is

$$S = S[x(t)] = \int_0^T L dt, \quad L = L(x, \dot{x}) \quad (12.1)$$

evaluated along the path $x(t)$ with the Lagrangian L which depends on x and $\dot{x} = dx/dt$. It is specified that $x(0) = a$, $x(T) = b$, and we seek paths with these boundary values that make S an extremum. Let $\eta(t)$ be a function on $[0, T]$ with $\eta(0) = \eta(T) = 0$ and consider the expansion of S about some fixed $x(t)$:

$$\begin{aligned} S(x + \epsilon\eta) &= S(x) + \epsilon \int_0^T \left(\frac{\partial L}{\partial \dot{x}} \dot{\eta} + \frac{\partial L}{\partial x} \eta \right) dt \\ &\quad + \frac{\epsilon^2}{2!} \int_0^T \left[\frac{\partial^2 L}{\partial \dot{x}^2} \dot{\eta}^2 + 2 \frac{\partial^2 L}{\partial x \partial \dot{x}} \dot{\eta} \eta + \frac{\partial^2 L}{\partial x^2} \eta^2 \right] dt + \dots \\ &\equiv S + \epsilon \delta S + \frac{\epsilon^2}{2!} \delta^2 S + \dots \end{aligned} \quad (12.2)$$

The extremum property of S is expressed by $\delta S = 0$, and an integration by parts in δS leads to the Euler-Lagrange equations for x :

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0 \quad (12.3)$$

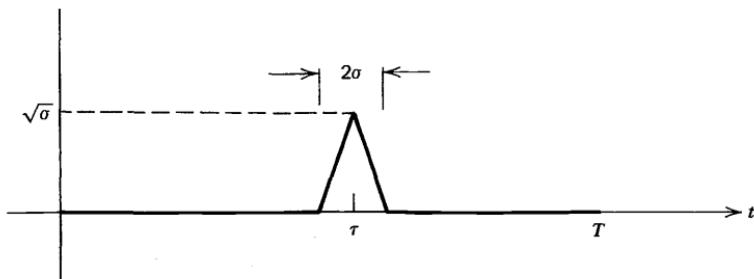


Fig. 12.1 The function $\eta_{\sigma\tau}(t)$.

However, for information on minimizing properties one must go to higher derivatives, starting with $\delta^2 S$.

An obvious necessary condition for $x(t)$ to be a minimum of S is that the quadratic form $\delta^2 S$ be nonnegative. To get an easy consequence of this requirement consider the function (see Fig. 12.1)

$$\eta_{\sigma\tau}(t) \equiv \begin{cases} 0 & |t-\tau| > \sigma \\ \sqrt{\sigma} \left(1 - \frac{|t-\tau|}{\sigma}\right) & |t-\tau| \leq \sigma \end{cases} \quad (12.4)$$

If the derivatives of L along $x(t)$ are well behaved the dominant contribution to $\delta^2 S$ comes from $\dot{\eta}^2$, for

$$\int \eta_{\sigma\tau}^2 = O(\sigma^2), \quad \int \eta_{\sigma\tau} \dot{\eta}_{\sigma\tau} = O(\sigma), \quad \int \dot{\eta}_{\sigma\tau}^2 = O(1) \quad (\sigma \rightarrow 0) \quad (12.5)$$

Since this must be true for any τ and for arbitrarily small σ we immediately have the Legendre condition:

$$\frac{\partial^2 L}{\partial \dot{x}^2} \geq 0 \quad (12.6)$$

For improvements on this sort of estimate see Courant and Hilbert's chapter on the calculus of variations. Equation 12.6 for Lagrangians of the form $\frac{1}{2} m \dot{x}^2 - V(x)$ means $m \geq 0$. For Riemannian geometry it means that the metric must not have negative eigenvalues.

The question of positive definiteness of $\delta^2 S$ can be phrased in terms of the variational problem associated with $\delta^2 S$ itself. Again Courant and Hilbert should be consulted if there is a gap in your education concerning the variational properties of eigenfunctions.

What we wish to know is whether there is *any* function $\eta(t)$ for which δ^2S is negative. To this end we seek the function $\varphi(t)$ which minimizes δ^2S and if for that function δ^2S is positive then we shall know δ^2S is positive for all functions. The function φ must be normalized, since its homogeneous boundary conditions set no scale for the problem. We therefore seek a function φ satisfying

$$\int_0^T |\varphi(t)|^2 dt = 1 \quad \text{and} \quad \varphi(0) = \varphi(T) = 0 \quad (12.7)$$

which minimizes

$$\delta^2S = \int \tilde{L} dt \quad \tilde{L}(\varphi) = \frac{\partial^2 L}{\partial \dot{x}^2} \dot{\varphi}^2 + 2 \frac{\partial^2 L}{\partial x \partial \dot{x}} \varphi \dot{\varphi} + \frac{\partial^2 L}{\partial x^2} \varphi^2 \quad (12.8)$$

where $\partial^2 L / \partial x^2$, etc., are evaluated along the path $x(t)$. The normalization constraint is handled by the standard method of Lagrange multipliers and it follows that φ satisfies the Euler-Lagrange equation derived from \tilde{L}

$$\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{x}^2} \dot{\varphi} \right) + \left[\frac{d}{dt} \left(\frac{\partial^2 L}{\partial x \partial \dot{x}} \right) - \frac{\partial^2 L}{\partial x^2} \right] \varphi + \lambda \varphi = 0 \quad (12.9)$$

where λ is the Lagrange multiplier. Moreover, for φ a solution of (12.9) a simple integration by parts shows that

$$\begin{aligned} \delta^2S[\varphi] &\equiv \int_0^T dt \left[\frac{\partial^2 L}{\partial \dot{x}^2} \dot{\varphi}^2 + 2 \frac{\partial^2 L}{\partial x \partial \dot{x}} \varphi \dot{\varphi} + \frac{\partial^2 L}{\partial x^2} \varphi^2 \right] \\ &= \lambda \int_0^T \varphi^2 dt = \lambda \end{aligned} \quad (12.10)$$

Hence the Lagrange multiplier has the significance of being the value of δ^2S that we seek.

So far we have looked for the smallest value of δ^2S . However, (12.9) with the conditions (12.7) has a significance well beyond that. Equation 12.9 in fact defines a Sturm-Liouville problem whose eigenfunctions and eigenvalues are those of δ^2S where δ^2S is considered a quadratic form (for this to be true the positivity of $\partial^2 L / \partial \dot{x}^2$ is also required). Eigenfunctions φ for eigenvalues λ higher than the lowest do not minimize δ^2S , but they do make it stationary and also satisfy (12.10).

From the properties of Sturm-Liouville equations we know that (12.9) has an infinity of eigenvalues and eigenvectors, which we designate

$$\lambda_n, \quad \varphi_n \quad n = 1, 2, \dots \quad (\lambda_1 < \lambda_2 < \dots) \quad (12.11)$$

The $\{\varphi_n\}$ form a complete orthonormal set. The coefficients $\partial^2 L / \partial x^2$, and so on are evaluated along the stationary path $x(t)$ and depend on it and its endpoints. Hence λ_n and φ_n are functions of $b = x(T)$, $a = x(0)$, T and, if these endpoints define a number of extrema, some label to distinguish among extrema. Any function $\eta(t)$ vanishing at $t = 0$ and $t = T$ can be written

$$\eta(t) = \sum_{n=1}^{\infty} a_n \varphi_n(t) \quad (12.12)$$

and we immediately have

$$\delta^2 S[\eta] = \sum_{n=1}^{\infty} \lambda_n a_n^2 \quad (12.13)$$

The question of positive definiteness of $\delta^2 S$ reduces to that of the positivity of all its eigenvalues λ_n . Hence $x(t)$ is a minimum if

$$\lambda_n > 0 \quad \text{for all } n \quad (12.14)$$

Conversely it is obvious that $x(t)$ is *not* a minimum if for any n , $\lambda_n < 0$. In particular, for sufficiently small ϵ we would have

$$S(x + \epsilon \varphi_n) = S(x) + \epsilon^2 \lambda_n + O(\epsilon^3) < S(x) \quad \text{for } \lambda_n < 0 \quad (12.15)$$

Let us consider the Lagrangian

$$L = \frac{1}{2} m \dot{x}^2 - V(x) \quad (12.16)$$

although our results are considerably more general. For sufficiently small T the kinetic energy term dominates and if V is not singular the only path will be given approximately by

$$x(t) = a + \frac{(b-a)t}{T} \quad (12.17)$$

(up to corrections due to V). For this path and for this Lagrangian, (12.9) becomes

$$m \ddot{\varphi} + \left. \frac{\partial^2 V}{\partial x^2} \right|_{x(t)} \varphi + \lambda \varphi = 0, \quad \varphi(0) = \varphi(T) = 0 \quad (12.18)$$

The Dirichlet boundary conditions of (12.18) make that equation equivalent to a Schrödinger equation with infinite hard walls. As the walls come

together ($T \rightarrow 0$) the ground state energy gets large (essentially this is just the uncertainty principle) so that once again in the $T \rightarrow 0$ limit the term arising from the potential can be neglected. In this approximation the solutions to (12.18) are

$$\lambda_n = m \left(\frac{n\pi}{T} \right)^2, \quad n=1, 2, \dots$$

$$\varphi_n(t) = \sqrt{\frac{2}{T}} \sin\left(\frac{n\pi t}{T}\right) \quad (12.19)$$

It follows that for T small enough $x(t)$, the classical trajectory given approximately by (12.17), is not just an extremum but is actually a minimum of the action.

Obviously for larger T the proof above does not hold. Hence when a trajectory which is a minimum is extended in such a way that $x(t)$ continues to satisfy the Euler-Lagrange equations it may at some stage lose the minimum property. It is this phenomenon we now wish to examine.

Notice that our viewpoint has changed slightly. We are no longer looking at the boundary value problem ($x(0)=a, x(T)=b$) but are concentrating instead on some particular trajectory out of a at time zero (a solution of the initial value problem $x(0)=a, \dot{x}(0)=v$, a and v given). For that specific trajectory (call it \bar{x}) we can also pick one of its points, say $b=\bar{x}(T)$ for $t=T$, and consider the boundary value problem for the pair ($x=a, t=0$) and ($x=b, t=T$). The conclusion of the paragraphs above is that for T small enough the curve $\bar{x}(\tau)$, $0 \leq \tau \leq T$, minimizes the action for that boundary value problem. With a larger T , \bar{x} may only be an extremum. However, for each T the boundary value problem defines a Sturm-Liouville equation and a corresponding set of eigenvalues which are in effect functions of T . The foregoing results about the small T minimizing property of \bar{x} can be restated

$$\lambda_n(T) > 0, \quad \text{for all } n \text{ for sufficiently small } T \quad (12.20)$$

As T increases the collection of eigenvalues generally decreases numerically and $x(t)$ (we drop the bar) will cease to be a minimum for that T for which

$$\lambda_1(T) = 0 \quad (12.21)$$

The fact of this decrease for small T can be seen in (12.19), but for larger T there is no guarantee of this, nor is there any guarantee that λ_1 will ever reach zero. The point $b=x(T)$ at which (12.21) holds is of particular

significance and is known as a *conjugate point* of the point a at which $x(t)$ began. We find therefore that if b is conjugate to a along a trajectory $x(t)$, then there exists a solution to the differential equation

$$\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{x}^2} \dot{\varphi} \right) + \left[\frac{d}{dt} \left(\frac{\partial^2 L}{\partial x \partial \dot{x}} \right) - \frac{\partial^2 L}{\partial x^2} \right] \varphi = 0 \quad (12.22)$$

with boundary conditions $\varphi(0) = \varphi(T) = 0$ and with the derivatives of L evaluated along the extremum $x(t)$ from a (at $t=0$) to b (at $t=T$). For the Lagrangian $L = \frac{1}{2} m \dot{x}^2 - V(x)$ of (12.6), (12.22) takes the simpler form

$$m \ddot{\varphi} + \frac{\partial^2 V}{\partial x^2} \Big|_{x(t)} \varphi = 0 \quad (12.23)$$

Equation 12.23 has appeared earlier in this book [see (6.31)] in connection with the path integral for the quadratic Lagrangian, and we shall soon see that the reappearance here is not accidental. This equation is known as the Jacobi equation as it was he who demonstrated its significance with regard to the variational problem. The solution of the Jacobi equation is known as the Jacobi field. (In quantum field theory this mode corresponds to the Goldstone boson.)

Once $x(t)$ has passed through its first conjugate point it is no longer a minimum. Presumably this could be shown by doing perturbation theory on the eigenvalue $\lambda_1(T+\epsilon)$, but instead we shall give a direct demonstration of a path for which $\delta^2 S$ is negative.

Let T be the time at which $x(t)$ passes through b , the conjugate point of $a = x(0)$. Let $\varphi(t), 0 \leq t \leq T$ be the associated Jacobi field. Consider $x(t)$ for $0 \leq t \leq T' = T + \epsilon$ (for a fixed positive ϵ), and let $\xi(t)$ be a variation about this path. Thus $\xi(0) = 0$ and $\xi(T+\epsilon) = 0$. We shall now produce a path $y(t)$ for which $\delta^2 S < 0$. For a given c (later to be specified) let

$$y(t) \equiv \begin{cases} \frac{1}{c} \varphi(t) - c \xi(t) & 0 \leq t \leq T \\ -c \xi(t) & T \leq t \leq T + \epsilon \end{cases} \quad (12.24)$$

Then

$$\delta^2 S[y] = \int_0^{T+\epsilon} dt \left[\frac{\partial^2 L}{\partial \dot{x}^2} \dot{y}^2 + \dots \right] = \int_0^T dt \left[\frac{\partial^2 L}{\partial \dot{x}^2} \dot{y}^2 + \dots \right] + O(c^2) \quad (12.25)$$

When y , as defined by (12.24), is inserted in the integral from 0 to T , the term $\delta^2 S[\varphi]$ drops out by the definition of the Jacobi field. Thus to order c only cross terms involving ξ and φ remain. We do integration by parts on the cross terms and use the fact that φ satisfies the Jacobi equation. The result is

$$\delta^2 S[y] = -2\xi(T) \frac{\partial^2 L}{\partial \dot{x}^2} \Big|_T \dot{\varphi}(T) + O(c^2) \quad (12.26)$$

So far ξ has been arbitrary; we now select ξ to be such that

$$\xi(T) = \dot{\varphi}(T) \quad (12.27)$$

It follows that for sufficiently small c ,

$$\delta^2 S < 0 \quad (12.28)$$

It should be pointed out that although at some stages in our discussion we seemed to be restricted to one dimension (for the position, x), in fact the results hold generally, with appropriate notational extensions; for example, φ becomes a vector function of t , $\partial^2 V / \partial x^2$ a matrix.

There is another way to arrive at the Jacobi equation which will give us additional perspective and assist us greatly in making the connection to quantum mechanics. Consider families of (extremum) paths, all leaving a at time 0 (without specification of endpoint). Label these by their initial momenta p , so that a trajectory is designated

$$x(p, t) \quad (12.29)$$

and

$$x(p, 0) = a, \quad \text{all } p \quad (12.30)$$

We wish to study the way in which two nearby extrema deviate from one another and to this end consider the quantity

$$J(p, t) = \frac{\partial x(p, t)}{\partial p} \quad (12.31)$$

The difference between two trajectories is therefore

$$x(p + \varepsilon, t) - x(p, t) = \varepsilon J(p, t) + O(\varepsilon^2) \quad (12.32)$$

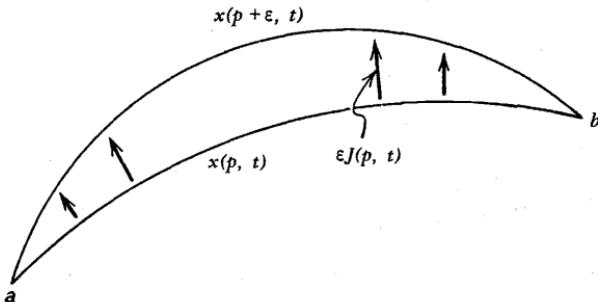


Fig. 12.2 Paths $x(p, t)$ and $x(p + \epsilon, t)$ out of a which (for $\epsilon \rightarrow 0$) meet at the focal point b . Both $x(p, t)$ and $x(p + \epsilon, t)$ satisfy the Euler-Lagrange equation. The vector connecting $x(p, t)$ and $x(p + \epsilon, t)$ at the same time t is proportional to the Jacobi field $J(p, t)$. At the focal point $J(p, t)$ vanishes.

See Fig. 12.2. Since $x(p, t)$ satisfies

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0 \quad (12.33)$$

we merely differentiate (12.33) with respect to p to discover that J satisfies

$$\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{x}^2} J \right) + \left[\frac{d}{dt} \left(\frac{\partial^2 L}{\partial x \partial \dot{x}} \right) - \frac{\partial^2 L}{\partial x^2} \right] J = 0 \quad (12.34)$$

which is the Jacobi equation. By (12.30) we have that

$$J(p, 0) = 0 \quad (12.35)$$

Furthermore, $\dot{x}(p, 0)$ and p are related by $p = m\dot{x}(p, 0)$ for the simplest Lagrangian. With an external magnetic field there is in addition a vector potential. But in either situation

$$\frac{\partial \dot{x}(p, 0)}{\partial p} = \frac{1}{m}, \quad \text{implying} \quad \frac{\partial J(p, 0)}{\partial t} = \frac{1}{m} \quad (12.36)$$

Thus J starts out zero, but with nonzero derivatives. Conjugate points occur when for some T ,

$$J(p, T) = 0 \quad (12.37)$$

In terms of deviations between nearby trajectories we form the following picture: Along the momentum direction defined by that $J(p, T)$ which vanishes* we have a family of trajectories spreading out initially; but

*The term "momentum direction" is appropriate for the many dimensional version defined immediately below. In one dimension there is of course only one direction.

eventually since $J(p, T) = 0$, they must reconverge at the point b . For this reason a conjugate point is often known as a *focal point*.

In n dimensions x is indexed $x_i, i = 1, \dots, n$. There is an n parameter family of extrema leaving the initial point, and these parameters can be labeled $p_k, k = 1, \dots, n$. The analogues of the foregoing one-dimensional equations are

$$x_i(p, 0) = a_i, \quad \text{all } p; \quad J_{ik}(p, t) = \frac{\partial x_i(p, t)}{\partial p_k} \quad (12.38)$$

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{x}_i \partial \dot{x}_l} j_{lk} \right) + \left(\frac{\partial^2 L}{\partial \dot{x}_i \partial x_l} - \frac{\partial^2 L}{\partial x_i \partial \dot{x}_l} \right) j_{lk} \\ + \left[\frac{d}{dt} \left(\frac{\partial^2 L}{\partial \dot{x}_i \partial x_l} \right) - \frac{\partial^2 L}{\partial x_i \partial \dot{x}_l} \right] J_{lk} = 0 \end{aligned} \quad (12.39)$$

$$J_{ik}(p, 0) = 0 \quad \frac{\partial J_{ik}(p, 0)}{\partial t} = \left(\frac{1}{m} \right) \delta_{ik} \quad (12.40)$$

$$\det J_{ik}(p, T) = 0 \text{ when } x_i(T) \text{ is a focal point} \quad (12.41)$$

Equation 12.39 corresponds to the equation of geodesic deviation familiar from general relativity.

For each t and $x(t)$ the quantities J_{ik} can be thought of as a collection of n vectors, labeled by k . Each points in the direction reached at time t by a vector pointing in the k direction at time 0. At focal points the determinant of J_{ik} vanishes, which is the condition that the vectors no longer span the tangent space at $x(t)$. Since $\det J$ is zero, it has an eigenvector with zero eigenvalue, say u_k , so that

$$\sum_k J_{ik}(T) u_k = 0 \quad (12.42)$$

the u_k 's being simply numbers. Consider the function

$$\varphi_i(t) = \sum J_{ik}(t) u_k \quad (12.43)$$

By (12.40) and (12.42) it vanishes at $t=0$ and T and by the linearity of (12.39) it satisfies the Jacobi equation. Consequently in n ($n > 1$) dimensions focal points have an associated direction (and in fact in general are part of geometric structures known as caustic surfaces).

It can happen (and generically, does happen) that J has more than one zero eigenvalue at a point. Such higher order focal points can often be located by symmetry arguments.

The quantity J can also be written in terms of the action. Consider endpoints a, b and times $t=0$ and $t=T$. Pick some (extremum) classical path between the set of endpoints, and as the endpoints vary slightly the path will also vary slightly (generally speaking). The action computed along each of these classical paths will be denoted $S(b, T; a)$. For one dimension, the initial momentum $p(0)$ is given by

$$p(0) = - \left. \frac{\partial S}{\partial a} \right|_{b, T} \quad (12.44)$$

Usually one thinks of the momentum as $\partial L / \partial \dot{x}$, but anyone who has gotten to the later chapters of Goldstein (see references) will find (12.44) familiar [see also (3.10)]. Of course $p(0)$ is what we have been calling p , but because in (12.44) its value is fixed by the boundary value problem it is here thought of as a function of b and T also. Equation 12.44 is differentiated with respect to b and we obtain

$$\frac{1}{J} = \frac{\partial p}{\partial b} = - \left. \frac{\partial^2 S}{\partial a \partial b} \right|_{b, T} \quad (12.45)$$

Similarly for n ($n > 1$) dimensions the n -dimensional version of (12.44) can be used to show that J_{ik} and $\partial^2 S / \partial b_i \partial a_k$ are inverse matrices, with b the final position and a the initial position.

Define D as

$$D = \det \left[- \left. \frac{\partial^2 S}{\partial a_k \partial b_l} \right|_{b, T} \right] \quad (12.46)$$

D will appear prominently in the WKB approximation and is interpreted as a density of paths. [From (12.39) and our remarks above we already know it to be the inverse of the determinant of the Jacobi fields.] This interpretation is justified by the following continuity equation:

$$\frac{\partial D}{\partial t} + \sum_k \frac{\partial}{\partial b_k} (v_k D) = 0 \quad (12.47)$$

where v_k is the velocity at time T . To prove (12.47) we introduce the Hamiltonian $H(x, p)$, related to the Lagrangian in the usual way. Then the action satisfies

$$\frac{\partial S(x, t; a)}{\partial t} + H \left(x, \frac{\partial S(x, t; a)}{\partial x} \right) = 0 \quad (12.48)$$

In terms of H we also have

$$v_k = \frac{\partial H}{\partial p_k} \quad (12.49)$$

so that v is also considered a function of x and p . Equation 12.48 is differentiated twice, once with respect to a_j and once with respect to x_k . Next set $x=b$, and after numerous manipulations one can obtain (12.47).

At a focal point, D is infinite (and J is zero). Thus many (at least a one-dimensional family of) paths have come together again. For example, for the harmonic oscillator the action is

$$S = \frac{m\omega}{2 \sin \omega T} [(a^2 + b^2) \cos \omega T - 2ab]$$

Then

$$\frac{\partial^2 S}{\partial a \partial b} = - \frac{m\omega}{\sin \omega T}$$

and this is infinite at each half period. Thus the “wave function” becomes a δ -function at each half period.

Exercise: For the Lagrangian $\frac{1}{2} \dot{x}^2 - V(x)$, $V(x) = \lambda e^x$, $\lambda > 0$ find the solutions of the Euler-Lagrange equations. Under what circumstances does the boundary value problem $x(0) = x(T) = a$ have 0, 1, or 2 solutions? The Jacobi equation can be solved in closed form. What do the results of the boundary value problem above say about zeros of the solution of the Jacobi differential equation?

We close this section by quoting without proof (but with commentary) some additional results from the calculus of variations.

First, there is the famous Morse index theorem which relates the conjugate points along a classical path to the negative eigenvalues of $\delta^2 S$. Parts of this relation have been developed above.

Definition: The *index* of a bilinear functional is the dimension of the space on which it is negative definite.

In our case, the index of $\delta^2 S$ is the number of eigenvalues λ_n , with $\lambda_n < 0$. As we move along a classical path (an extremum) the index of $\delta^2 S$ can change and in particular can only change at points conjugate to the initial point because that is where one of the eigenvalues goes through zero. The Morse index theorem makes the following stronger assertion:

THEOREM (MORSE): Let $x(t)$, $0 \leq t \leq T$, be an extremum of S . The index of $\delta^2 S$ is equal to the number of conjugate points to $x(0)$ along the curve $x(t)$ ($0 \leq t \leq T$). Each such conjugate point is counted with its multiplicity.

In (12.24) and the equations that follow an example was shown of the fact that once a curve passes through a conjugate point, $\delta^2 S$ picks up a negative eigenvalue. In fact that demonstration is one of the components of the proof of the index theorem.

References for Morse theory are works by Milnor and Morse. Some years ago other aspects of Morse theory made an appearance in the physics literature. Van Hove was able to draw conclusions about singularities in the phonon spectrum in a solid by using results of Morse theory relating the space on which a function is defined to the number of saddle points necessarily possessed by that function.

Just for the record, we include some results from the calculus of variations on sufficient conditions. The history of this field of mathematics is rich in failures to appreciate the distinction between necessity and sufficiency. (See Bliss's monograph, referenced below.) In the following example, the term "minimum" means, as usual, a local minimum.

Consider the action S , $S = \int_0^T L(x, v, t) dt$, $v = dx/dt$ and an extremal $x(t)$ that satisfies $d(\partial L/\partial v)/dt - \partial L/\partial x = 0$. Define the Weierstrass E -function as follows:

$$E(t, x, v, w) \equiv L(x, w, t) - L(x, v, t) - (w - v) \frac{\partial L}{\partial v}(x, v, t) \quad (12.50)$$

Then it is possible to give sufficient conditions for $x(t)$ to be a strong minimum:

THEOREM: $x(t)$ is a strong minimum among curves $\xi(t)$ satisfying $\xi(0) = x(0)$ and $\xi(T) = x(T)$ if

- (1) There is no conjugate point of $x(0)$ along $x(t)$, $0 \leq t \leq T$.
- (2) $\partial^2 L/\partial v_i \partial v_j$ is positive definite.
- (3) $E(\tau, x, v, w) > 0$ for all triples (τ, x, v) sufficiently near $(t, x, dx/dt)$ and all $v \neq w$.

Note that for $L = \frac{1}{2} v^2 - U(x)$ conditions 2 and 3 are automatically satisfied and one need only examine conjugate points.

NOTES

An excellent introduction to the calculus of variations can be found in the famous text

R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, Interscience, New York, 1953

Another famous text that treats the calculus of variations as well as properties of the action used later in this section is

H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, Mass., 1959

For Morse theory see the following

J. Milnor, *Morse Theory*, Princeton University Press, Princeton, N. J., 1963

M. Morse, *Variational Analysis*, Wiley, New York, 1973

Another good reference, and a relatively more readable one, is

M. M. Postnikov, *The Variational Theory of Geodesics*, Saunders, Philadelphia, 1967

Application to the singularities of the phonon spectrum in solids is given by

L. Van Hove, The Occurrence of Singularities in the Elastic Frequency Distribution of a Crystal, *Phys. Rev.* **89**, 1189 (1953)

More on the Weierstrass *E*-function can be found in

G. A. Bliss, *Calculus of Variations*, Mathematical Association of America, Open Court, La Salle, Ill., 1925

The WKB Approximation and Its Application to the Anharmonic Oscillator

Derivation of the WKB approximation is mainly a question of justifying replacement of an arbitrary Lagrangian by a certain quadratic Lagrangian and then doing the path integral for the quadratic Lagrangian according to the results of Section 6. The “justification” involves an asymptotic approximation in the limit $\hbar \rightarrow 0$, and although there will be holes in our proofs we try to keep these gaps as small as possible. Our development is not the simplest, but it has the advantage that certain generalizations (the study of caustics, in particular) are easy.

The Lagrangian is taken to be of the form

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 + \frac{e}{c}\dot{\mathbf{x}} \cdot \mathbf{A} - V(\mathbf{x}) \quad (13.1)$$

and we wish to compute $G(\mathbf{b}, T; \mathbf{a})$. Suppose for the moment that there is just one classical path $\bar{\mathbf{x}}(\tau)$ from $(\mathbf{a}, 0)$ to (\mathbf{b}, T) . Then roughly speaking what we are about to do is change variables in the path integral from $d\mathbf{x}(\tau)$ to $dy(\tau)$ where $y(\tau) = \mathbf{x}(\tau) - \bar{\mathbf{x}}(\tau)$. The action will be rewritten in terms of y , just as was done in (6.13) for the quadratic Lagrangian. Now, however, $\delta^3 S$ and higher powers of y do not vanish and the path integral

takes the form

$$G(\mathbf{b}, T; \mathbf{a}) = e^{iS(\bar{\mathbf{x}})/\hbar} \int_{(0,0)}^{(0,T)} d\mathbf{y}(\tau) \times \exp \left\{ \frac{i}{\hbar} \left[\frac{1}{2!} \delta^2 S(\mathbf{y}) + \frac{1}{3!} \delta^3 S[\mathbf{y}] + \dots \right] \right\} \quad (13.2)$$

with

$$\begin{aligned} \delta^2 S[\mathbf{y}] &= \int_0^T d\tau \left\{ m\dot{\mathbf{y}}^2 + 2\frac{e}{c} \frac{\partial A_k}{\partial x_i} y_i \dot{y}_k + \frac{\partial^2}{\partial x_j \partial x_k} [A_i \dot{x}_i - V] y_j y_k \right\} \\ \delta^3 S[\mathbf{y}] &= \int_0^T d\tau \left\{ \frac{3e}{c} \frac{\partial^2 A_i}{\partial x_j \partial x_k} y_j y_k \dot{y}_i + \frac{\partial^3}{\partial x_j \partial x_k \partial x_l} [A_i \dot{x}_i - V] y_j y_k y_l \right\} \end{aligned} \quad (13.3)$$

(summation convention), and so forth. At this stage one would like to use stationary phase arguments to show that as $\hbar \rightarrow 0$ terms that are cubic (or higher) in y , that is, $\delta^3 S$, and so on, are of smaller order in \hbar . Specifically we have something of the form of (11.18) and would like to say that the leading contribution, asymptotically as $\hbar \rightarrow 0$, depends only on the quadratic term in the exponent. Of course it would be nice to go further and develop an asymptotic series in \hbar involving $\delta^3 S$, and so on, but here we set our sights lower.

Continuing to argue heuristically we note that if indeed it is the quadratic term that dominates, then we have as in Section 3 [see the discussion following (3.6)]

$$\frac{\delta^2 S[\mathbf{y}]}{\hbar} = O(1) \quad (\hbar \rightarrow 0) \quad (13.4)$$

This means that \mathbf{y} is on the order of $\sqrt{\hbar}$, a fact of some importance for semi-classical considerations. It follows that quantum effects will be significant at a distance proportional to $\sqrt{\hbar}$ from the classical trajectory. There are, of course, several other dimensional quantities, so \hbar alone does not set the scale of distance for quantum effects. If we now assume that

$$\mathbf{y} = O(\sqrt{\hbar}) \quad (\hbar \rightarrow 0) \quad (13.5)$$

then $\delta^3 S$, which goes like y^3 , has an additional power of $\sqrt{\hbar}$ and

$$\frac{\delta^3 S}{\hbar} = 0(\sqrt{\hbar}) = o(1) \quad (\hbar \rightarrow 0) \quad (13.6)$$

Higher variations of S are correspondingly smaller, and to leading power of \hbar only $\delta^2 S$ contributes to the WKB approximation. But if only $\delta^2 S$ contributes, the calculation of the path integral in $dy(\tau)$ is exactly what has already been done for the quadratic Lagrangian (except that we are now carrying along a vector potential A , but this does not change the earlier results).

From (6.35) the path integral for the quadratic Lagrangian in one dimension is

$$G(b, t_b; a, t_a) = \sqrt{\frac{m}{2\pi i \hbar f(t_b, t_a)}} \exp\left(\frac{i}{\hbar} S(b, t_b; a, t_a)\right) \quad (13.7)$$

where $S(b, t_b; a, t_a)$ is the action along the classical path, f satisfies

$$\begin{aligned} m \frac{\partial^2 f(t, t_a)}{\partial t^2} + c(t)f(t, t_a) &= 0 \\ f(t_a, t_a) &= 0, \quad \left. \frac{\partial f(t, t_a)}{\partial t} \right|_{t=t_a} = 1 \end{aligned} \quad (13.8)$$

and $L = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} c x^2 + ex$. If (13.8) is compared to (12.34) through (12.36) it is seen that f and $J = \partial x(p, t)/\partial p$ satisfy exactly the same linear equation and have proportional boundary conditions. Moreover, since by (12.45) J is just $-1/(\partial^2 S/\partial b \partial a)$ it follows that

$$G(b, t_b; a, t_a) = \sqrt{\frac{i}{2\pi \hbar} \frac{\partial^2 S}{\partial b \partial a}} \exp\left(\frac{iS}{\hbar}\right) \quad (13.9)$$

In n dimensions this formula generalizes to

$$G(\mathbf{b}, t_b; \mathbf{a}, t_a) = \sqrt{\det\left(\frac{i}{2\pi \hbar} \frac{\partial^2 S}{\partial \mathbf{b} \partial \mathbf{a}}\right)} \exp\left(\frac{iS}{\hbar}\right) \quad (13.10)$$

That (13.10) is the correct n -dimensional generalization of (13.9) can be verified in a straightforward way. The G of (13.10) is supposed to be an approximation to the propagator for the system whose Lagrangian is given

by (13.1). However, it is the exact propagator for the quadratic action $\delta^2 S$ of (13.3). [The last assertion can be proved simply by applying the Hamiltonian corresponding to the quadratic Lagrangian (of $\delta^2 S$) directly to G .] Therefore, since G is exact for the truncation to $\delta^2 S$, it is formally the leading approximation (in powers of \hbar) to the exact propagator.

From another point of view (13.10) is a remarkable result, relating the determinant of a quadratic form to the solution of an ordinary differential equation. For $n > 1$ dimensions, as for one dimension, the object within the square root symbol in (13.10) is the determinant of $\delta^2 S$, with infinities appropriately defused (as a result of $e^{N/2}$ and similar terms appearing in the "infinite normalization constant" of the path integral). But we also know from the results of Section 12 that that object is the inverse of the determinant of n fields each obeying the Jacobi equation. For $n > 1$ dimensions the method of proof used in Section 6, namely going from a recursion relation for truncated versions of the determinant to a differential equation, is somewhat more complicated although it has been done by Papadopoulos. A less direct proof which generalizes quite easily to many dimensions has been provided by Levit and Smilansky.

It follows therefore that (13.10) can also be written

$$G(\mathbf{b}, t_b; \mathbf{a}, t_a) = \left(\frac{1}{2\pi i\hbar} \right)^{n/2} \sqrt{D} e^{iS/\hbar} = \left(\frac{1}{2\pi i\hbar} \right)^{n/2} (\det J_{ij})^{-1/2} e^{iS/\hbar} \quad (13.10a)$$

with D given by (12.46) and J_{ij} the Jacobi fields satisfying (12.38) to (12.40).

Finally an extremely neat—if a bit indirect—proof has been given by Coleman which makes it all a statement about complex functions. Define the functions

$$f(z) = \frac{\det(-\partial^2/\partial t^2 + W - z)}{\det(-\partial^2/\partial t^2 - z)}$$

$$g(z) = \frac{\varphi_{W-z}(T)}{\varphi_{-z}(T)}$$

where we have introduced quite a bit of new notation. By $\det(-\partial^2/\partial t^2 + U)$ we mean the product of the eigenvalues of the Schrödinger Hamiltonian $-\partial^2/\partial t^2 + U$ with (it is understood) Dirichlet boundary conditions at 0 and T . The function φ_U (up to normalization) is the solution of

$$\frac{-\partial^2 \varphi_U}{\partial t^2} + U \varphi_U = 0$$

with boundary conditions $\varphi_U(0)=0$, $d\varphi_U(0)/dt=1$. The theorem we are talking about is simply that $f(0)=g(0)$. But Coleman observed that the two functions must in fact be the same for all z . First their poles and zeros are the same, since each $\varphi_U(t)$ that vanishes at $t=T$ is in fact an eigenvector of the corresponding Dirichlet boundary value problem with zero eigenvalue. Moreover, for $|z|\rightarrow\infty$, except possibly along the line $\text{Re } z>0$, $\text{Im } z=0$, both f and g approach unity. Hence f and g must be the same function.

The determinant in (13.10) is known as the Van Vleck determinant, and in the 1920s he found it to be useful in semiclassical approximations. Early in the development of path integrals C. Morette (now DeWitt-Morette) discovered the need to include the Van Vleck determinant in certain propagators and later Pauli and B. DeWitt considered other contexts in which this determinant played a role.

So far we have considered the case of a single extremum. If there are several, label them x_α, x_β, \dots ; then if $(x_\alpha - x_\beta)^2 \delta^2 S \gg \hbar$ the Green's function is simply the sum of individual contributions (cf. the one-dimensional stationary phase approximation). If this inequality does not hold, the paths are near a focus, and we discuss this separately (Sections 15 and 16).

Rigorous proofs of analogous asymptotic formulas for the Wiener integral have been given in work of Schilder, Varadhan, and Donsker. Besides the fact that the Wiener integral involves a well defined measure, asymptotic approximations in the Wiener integral context have the advantage of involving Laplace's method (real exponential, $e^{-\lambda x^2}$), whereas we have been struggling with the method of stationary phase ($e^{i\lambda x^2}$).

Before giving a pseudoproof of (13.9) we present an application of it. The application is to the anharmonic oscillator, a system that has been extensively investigated as a sort of simplest nonlinear problem. But it is not really all that simple. The work described below is due to Lam. Later in the book, when looking at statistical mechanics applications, we study methods introduced by Langer and applied to the anharmonic oscillator, among other things. The techniques there involve "critical droplets" alias "instantons." In this section, however, those concepts are not introduced.

Consider then the Lagrangian

$$L = \frac{1}{2}\dot{x}^2 - \frac{1}{2}\omega^2 x^2 - \frac{1}{4}\lambda x^4, \quad \omega^2, \lambda > 0 \quad (13.11)$$

Our interest is in $\lambda \rightarrow 0$. We study (13.11) and a simpler analogue of it in four ways: perturbation theory, stationary phase, effective field approximation, and complex scaling.

First we examine an integral of the sort seen in Section 11—but not exactly the same. Let

$$I = \int_0^\infty dx e^{i(x^2 - \lambda x^4)} \quad (13.12)$$

As observed earlier, this is related to the parabolic cylinder function. First look at “perturbation theory” in the “coupling constant” λ :

$$\begin{aligned} I &= \int_0^\infty dx e^{ix^2} \sum_{n=0}^{\infty} \frac{(-i\lambda)^n}{n!} x^{4n} \\ &\sim \frac{1}{2}\sqrt{i} \sum_{n=0}^{\infty} \frac{(i\lambda)^n}{n!} \Gamma(2n + \frac{1}{2}) \end{aligned} \quad (13.13)$$

We have seen this before. It is an asymptotic series with zero radius of convergence.

Now consider the stationary phase approximation. Let $x=y/\sqrt{\lambda}$. I becomes

$$I = \frac{1}{\sqrt{\lambda}} \int_0^\infty \exp\left[\frac{i}{\lambda}(y^2 - y^4)\right] dy \quad \frac{1}{\lambda} \rightarrow \infty \quad (13.14)$$

The critical points occur at

$$y=0, \quad \frac{1}{\sqrt{2}} \quad (13.15)$$

The $y=0$ term yields the asymptotic series given above. For the $y=1/\sqrt{2}$ critical point we expand the argument of the exponent about this point. Let $u=y-1/\sqrt{2}$ and a be some number between 0 and $1/\sqrt{2}$ and at least $\sqrt{\lambda}$ away from either. Then calling I' the contribution from this critical point

$$\begin{aligned} I' &= \frac{1}{\sqrt{\lambda}} \int_{-a}^\infty du \exp\left\{\frac{i}{\lambda}\left[\frac{1}{4} - 2u^2 - u^3(2\sqrt{2} + u)\right]\right\} \\ &= e^{i/4\lambda} \frac{1}{\sqrt{\lambda}} \int_{-a}^\infty du e^{-2iu^2/\lambda} \sum_{n=0}^{\infty} \left(-\frac{i}{\lambda}\right)^n \frac{[u^3(2\sqrt{2} + u)]^n}{n!} \end{aligned} \quad (13.16)$$

This too leads to an asymptotic expansion similar to (13.13). The sum of the leading contributions to I is

$$I \sim \frac{1}{2}\sqrt{i\pi} (1 + i\sqrt{2} e^{i/4\lambda}) \quad (13.17)$$

The notable feature of (13.16) and (13.17) is the essential singularity $\exp(i/4\lambda)$, a singularity that is missed by the perturbation theory result (13.13). People have been thinking about essential singularities in ostensibly well behaved perturbation expansions for a long time. An early paper is that of Dyson in which some physically appealing arguments about the catastrophe to be expected from changing the sign of α (=fine structure constant, the expansion parameter for QED) are mustered to suggest that the perturbation expansion for QED is divergent. We find below indications of an essential singularity in the anharmonic oscillator expansion; there the dramatic consequences of changing the sign of the quartic term are clear.

As a third approach to evaluating the integral (13.12) we try something analogous to the “effective field” approximation of statistical mechanics (this is getting to be quite a bit of machinery for a nineteenth century integral, but I think the analogies to be drawn are worthwhile). The term x^4 is replaced by μx^2 and then a value of μ is sought which is effectively equal to some expected value of x^2 . With the replacement,

$$I = \int_0^\infty dx e^{ix^2(1-\lambda\mu)} = \frac{1}{2} \sqrt{\frac{i\pi}{1-\lambda\mu}} \quad (13.18)$$

To select an effective value μ we can use the approximate integral I itself to define a kind of expectation value (in statistical mechanics this would be a natural step):

$$\langle f(x) \rangle = \left| \frac{\int dx f(x) e^{ix^2(1-\lambda\mu)}}{\int dx e^{ix^2(1-\lambda\mu)}} \right|$$

for a real function f . In this way the demand

$$\mu = \langle x^2 \rangle \quad (13.19)$$

becomes

$$\mu = \frac{1}{2(1-\lambda\mu)}$$

which for small λ has the approximate solutions

$$\mu = \frac{1}{2}, \quad \frac{1}{\lambda} - \frac{1}{2}$$

The first of these roots when inserted in (13.18) yields a convergent perturbation expansion. The second does yield a singularity, but a rather weak $I \sim 1/\sqrt{\lambda}$. Hence while this method reveals a bit more structure than the most naive perturbation expansion it still misses a lot. There is also an ambiguity in the criterion for selecting μ : instead of (13.19) a better condition might be $\langle \mu x^2 \rangle = \langle x^4 \rangle$, which would change some of the constants.

Finally, we study the integral using “complex scaling,” which will help with the analytic continuation although it will not pin down the exact nature of the singularity at $\lambda=0$. As a function of λ , the integral $I(\lambda)$ is obviously analytic so long as $\text{Re}(i\lambda) > 0$, that is, $\text{Im}\lambda < 0$. With a new variable of integration $y = x/\alpha$

$$I(\lambda) = \alpha \int_0^\infty dy \exp(i\alpha^2 y^2 - i\lambda\alpha^4 y^4) \quad (13.20)$$

nothing has changed so long as α is real. Now suppose $\alpha = \exp(i\varphi)$. The integral exists, provided $\text{Im}(\lambda \exp(4i\varphi)) < 0$. For small enough φ there are values of λ for which the integral is defined both before and after the transformation, and because the integral along the contour $z = R \exp(i\psi)$, $-4\varphi \leq \psi \leq 0$, vanishes for $R \rightarrow \infty$ the value of I is the same. However, $I(\lambda)$ is now defined for values of λ extending beyond $\text{Im}\lambda < 0$ and the function $I(\lambda)$ has been analytically continued. We use this freedom to undo any phase in the y^4 term and also to rescale magnitudes. Specifically let

$$\alpha = (i\lambda)^{-1/4}$$

The integral becomes

$$I(\lambda) = (i\lambda)^{-1/4} \int_0^\infty dy \exp[i(i\lambda)^{-1/2} y^2 - y^4]$$

(y real). $I(\lambda)$ therefore lives on a four-sheeted Riemann surface and is analytic everywhere away from $\lambda=0$. If we define a new variable $z = -i(i\lambda)^{1/2}$ and a function

$$f(z) = (i\lambda)^{1/4} I(\lambda) = \int_0^\infty dy \exp\left(-y^4 + \frac{y^2}{z}\right)$$

we see that the singularities of $I(\lambda)$ have two sources: first, the fourth-order branch point and, second, the singularity associated with the function $f(z)$ which was what was being picked up in (13.17).

Now after all that warmup we turn to the path integral

$$G(b, T; a) = \int_{(a, 0)}^{(b, T)} dx(\tau) \exp \left\{ \frac{i}{\hbar} \int_0^T \left[\frac{1}{2} \dot{x}^2 - \frac{1}{2} \omega^2 x^2 - \frac{1}{4} \lambda x^4 \right] d\tau \right\} \quad (13.21)$$

By a change of variables $y = x \sqrt{\lambda}$ (13.21) is rewritten as

$$G(b, T; a) = \int_{(a\sqrt{\lambda}, 0)}^{(b\sqrt{\lambda}, T)} dy(\tau) \exp \left\{ \frac{i}{\hbar\lambda} \int_0^T \left[\frac{1}{2} \dot{y}^2 - \frac{1}{2} \omega^2 y^2 - \frac{1}{4} y^4 \right] d\tau \right\} \quad (13.22)$$

The normalizing factor in $dy(\tau)$ will of course be of the usual form $[m/2\pi i\hbar\epsilon]^{1/2}$, but will contain $\lambda\hbar$ instead of just \hbar . Equation 13.22 brings out the interesting fact that the stationary phase approximation for (13.21) involves as large parameter not $1/\hbar$ but $1/\lambda\hbar$. Therefore although our primary interest in this problem is $\lambda \rightarrow 0$, and not what is traditionally called the WKB approximation, that is, $\hbar \rightarrow 0$, the fact that only the product $\lambda\hbar$ appears in G means that all our WKB results can be used in studying the $\lambda \rightarrow 0$ limit.

Next we return to the form (13.21) to get a clearer idea of what the limit $\lambda \rightarrow 0$ does to the classical mechanics. We are interested in solutions of

$$\ddot{x} + \omega^2 x + \lambda x^3 = 0 \quad (13.23)$$

with boundary condition

$$x(0) = a, \quad x(T) = b \quad (13.24)$$

That the potential has stiff walls ($V(x) \rightarrow \infty$ rapidly for $x \rightarrow \infty$) implies that for given time T there are many classical paths. Besides the one that is close to the harmonic oscillator path there are those of high energy that go to large x and come back quickly. Label the classical paths \bar{x}_α . Then

$$G \sim \sum_\alpha \sqrt{\frac{\partial^2 S_\alpha}{\partial b \partial a}} e^{iS_\alpha/\hbar}, \quad S_\alpha \equiv S(\bar{x}_\alpha) \quad (13.25)$$

The term in (13.25) from the path approximating that of the harmonic oscillator is what we ordinarily call perturbation theory. This can be seen by considering the expansion of (13.21) in powers of λ :

$$G \sim \int dx(\tau) \exp \left\{ \frac{i}{\hbar} \int \left[\frac{1}{2} \dot{x}^2 - \frac{1}{2} \omega^2 x^2 \right] d\tau \right\} \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i\lambda}{4\hbar} \right)^n \left[\int_0^T x^4(\tau) d\tau \right]^n \quad (13.26)$$

The $n=0$ term is just what we called G_0 in our study of perturbation theory, and we have already seen how to supply Feynman diagrams for the higher n terms. But G_0 is obtained by looking only at the harmonic oscillator path from $(a, 0)$ to $(b, 0)$ —because the WKB approximation is exact for the harmonic oscillator.

Of interest here is the contribution to G from other paths, differing significantly from the harmonic one. To find the action for these paths we consider the Hamilton-Jacobi equation for the function $S_\alpha(x, t; a) = [\text{action along the classical path } x_\alpha \text{ from } (a, 0) \text{ to } (x, t)]$. S satisfies

$$\frac{1}{2} \left(\frac{\partial S}{\partial x} \right)^2 + \frac{1}{2} \omega^2 x^2 + \frac{1}{4} \lambda x^4 + \frac{\partial S}{\partial t} = 0 \quad (13.27)$$

For classical paths $\partial S / \partial t$ is a constant related to the energy:

$$\frac{\partial S}{\partial t} = -E \quad (13.28)$$

Moreover, at any time E is given by

$$E = \frac{1}{2} \dot{x}^2 + \frac{1}{2} \omega^2 x^2 + \frac{1}{4} \lambda x^4 \quad (13.29)$$

Let the maximum value in x reached by the particle be designated M . Since $\dot{x}=0$ at that point,

$$E = \frac{1}{2} \omega^2 M^2 + \frac{1}{4} \lambda M^4 \quad (13.30)$$

Now the data given in our problem are a , b and T , not E . Equation 13.29 is used to relate E to T in the usual way:

$$T = \int_a^b dx [2(E - V(x))]^{-1/2} \quad (13.31)$$

$$V(x) = \frac{1}{2} \omega^2 x^2 + \frac{1}{4} \lambda x^4 \quad (13.32)$$

The integral in (13.31) requires some care in case the particle reaches $x = \pm M$ several times. In particular if $a = b = 0$,

$$T = 2n \int_0^M dx [2(E - V)]^{-1/2} \quad (13.33)$$

where n is the total number of turning point ($\pm M$) impacts. For the action S , (13.27) gives a similar elliptic integral, and again looking at the case

$a = b = 0$ we have

$$T = \frac{2n}{M\sqrt{\lambda}} \int_0^1 \frac{du}{\sqrt{1-u^4 + (2\omega^2/\lambda M^2)(1-u^2)}}$$

$$S = 2nM^3\sqrt{\lambda} \int_0^1 du \sqrt{1-u^4 + (2\omega^2/\lambda M^2)(1-u^2)} \quad (13.34)$$

S as a function of T is given implicitly through the dependence on M . The many solutions of these equations reflect the many classical paths. We wish to get an idea of the dependence of S on λ for paths of large amplitude. Such a path should be characterized by the fact that at the turning point the quartic term dominates the quadratic one and the term in the integrands of (13.34) containing ω^2 can be neglected. For self consistency we should later return to check if indeed

$$\frac{1}{4}\lambda M^4 \gg \frac{1}{2}\omega^2 M^2 \quad (13.35)$$

for the value of M we shall obtain. When $\omega^2/\lambda M^2$ is neglected in (13.34) we see that $T^3 S$ is independent of M and

$$S = \frac{n^4 K}{\lambda T^3} \quad (13.36)$$

for some constant K . This is the desired result. The point to notice is that S is inversely proportional to λ so that the Green's function of (13.21) has a singularity at $\lambda = 0$. Let us check (13.35) by replacing M by T . Using (13.34) with the neglect of $\omega^2/\lambda M^2$ (it is self consistency we are after) we obtain

$$M = \frac{nc}{T\sqrt{\lambda}} \quad c = \text{const.} \quad (13.37)$$

so that (13.35) becomes

$$n^2 c' \gg \omega^2 T^2, \quad c' = \text{const.} \quad (13.38)$$

The validity of (13.35) is therefore not related to the size of λ , but holds for sufficiently small T . Moreover, by taking a path that bounces back and forth a sufficient number (n) of times one can get (13.35) to hold to arbitrary accuracy.

The origin of the singularity here in the Green's function is the old story: one does not expect a convergent perturbation expansion in λ , since

as λ changes sign all bound states cease to exist. Hence there should be no power series in λ for the energy.* It is of interest though to ask why perturbation theory works so well—and where can one expect it to break down. Let us consider $\tilde{G}(E)$, the Fourier transform of $G(t)$ (see Section 7):

$$\tilde{G}(E) = \int_0^\infty e^{i(E+ie)t/\hbar} G(t) dt, \quad \epsilon > 0, \epsilon \rightarrow 0 \quad (13.39)$$

We have found that G is a sum of terms $\exp(iS(x_\alpha)/\hbar)$ so that each $S(x_\alpha)$ is a function $S_\alpha(b, t; a)$. To evaluate (13.39) we make a stationary phase approximation for the t integration. Substituting $\exp(iS_\alpha/\hbar)$ for G and requiring the vanishing of the derivative with respect to t of the argument of the exponent in (13.39) yields

$$E + \frac{\partial S_\alpha}{\partial t} = 0 \quad (13.40)$$

[cf. (13.28)] and we see that given E [which is what one is given when looking at $\tilde{G}(E)$], (13.40) will pick out the time t which gives the greatest contribution to the integral (13.39). Let the endpoints be $a=0$ and $b=0$ and consider T as a function of E . For E small enough the anharmonic term introduces no large amplitude paths and only affects the small amplitude paths (those approximating pure harmonic oscillator paths) slightly. Consequently for such small E the potentially significant effect of the singularity is lost.

By contrast, for large E the large amplitude paths do contribute stationary points on the real positive t axis in the integral in (13.39), and a breakdown in perturbation theory may be expected. The location of the stationary points may be deduced by using some of our earlier results on the large amplitude paths to yield

$$E \sim \frac{1}{4} \lambda M^4$$

$$T \sim \frac{n}{M \sqrt{\lambda}} \quad (13.41)$$

*The expectation expressed here is called by Simon (reference given below) a “folk theorem” and he shows it to be unjustified using the example $H(\lambda) = p^2 - (1/r) + \lambda(1/r)$, for which the energies are $E_n(\lambda) = E_n(0)(\lambda^2 - 2\lambda + 1)$. For $\lambda > 1$ there are no bound states but the (three term) power series for $E_n(\lambda)$ is perfectly healthy. Dyson’s argument, mentioned after (13.17), has been criticized on similar grounds. Despite this counterexample, the folk theorem has intuitive appeal and very likely a considerable domain of validity. Although it may be inadequate as a foundation for beliefs, I think it can serve in educating one’s guesses.

so that

$$T = \frac{c'' n}{\lambda^{1/4} E^{1/4}} \quad c'' = \text{const.} \quad (13.42)$$

For purposes of locating stationary points in integrals like (13.39), that is, solutions of (13.40), it is often useful to plot T as a function of E , a solution curve in the $T-E$ plane being a set of paths having the same number of turning points. In Fig. 13.1 is shown the T vs. E curve for trajectories having 1 and 2 turning points. To solve (13.40) one draws a vertical line in Fig. 13.1 at the E in question. For E below the regime of validity of (13.42) one may expect answers close to those for the harmonic oscillator. The dilemma in calculating $G(t)$ is that E is not fixed and the line one would draw in Fig. 13.1 is a horizontal one. For any T , if one goes far enough to the right he will always intersect an infinity of paths (for large n possibly) which have been profoundly affected by the potential $\frac{1}{4}\lambda x^4$.

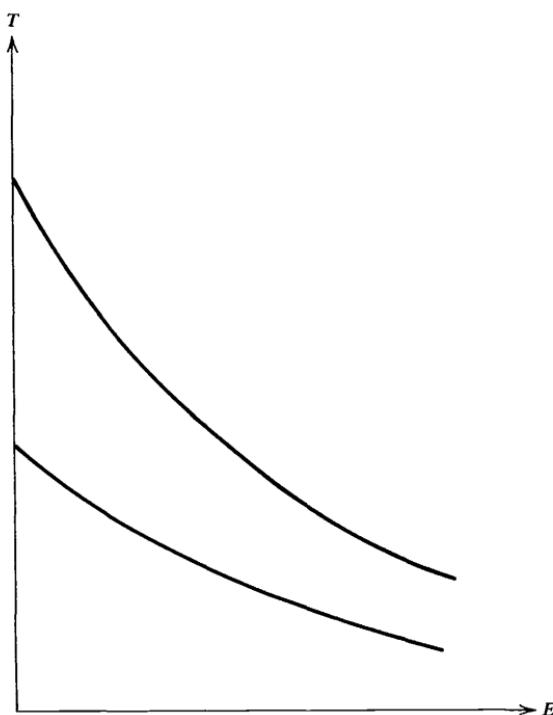


Fig. 13.1 Time elapsed for going between fixed endpoints, for varying energies. Paths having one and two turning points are shown.

Exercise: Draw $T(E)$ for various potentials including that considered in the exercise in Section 12 ($V(x)=\lambda e^x$). What properties of this curve characterize focusing of the system (i.e., the initial and final points are conjugate for the given T)?

As a final note, we mention that complex scaling can also be used for analytic continuation of the energies. The Hamiltonian for the Lagrangian (13.11) is

$$H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2x^2 + \frac{1}{4}\lambda x^4 \quad (13.43)$$

Let its energy eigenvalues be designated

$$E_n(\omega^2, \lambda) \quad n=0, 1, \dots \quad (13.44)$$

For real, positive λ this is an analytic function of ω^2 in some neighborhood of the origin (branch points appear when ω^2 is large enough and far enough from the real axis for level crossings to occur). Consider the change of variable $x \rightarrow x/\alpha$. H becomes

$$H \rightarrow \frac{1}{\alpha^2} \left[\frac{1}{2}p^2 + \frac{1}{2}\alpha^4\omega^2x^2 + \frac{1}{4}\lambda\alpha^6x^4 \right] \quad (13.45)$$

For real α this scale change is a unitary transformation and the eigenvalues are unchanged. From (13.43) to (13.45) it follows that

$$E_n(\omega^2, \lambda) = \frac{1}{\alpha^2} E_n(\alpha^4\omega^2, \alpha^6\lambda) \quad (13.46)$$

is an identity in α . The great power of complex scaling is that this identity remains true for complex α for discrete energy levels [on the other hand, the continuous spectrum is shifted. We shall not go into this nor into the exact conditions for validity of (13.46)]. Letting $\alpha = \lambda^{-1/6}$ we get

$$E_n(\omega^2, \lambda) = \lambda^{-1/3} E_n(\omega^2\lambda^{-2/3}, 1) \quad (13.47)$$

In this way a three-sheeted structure for E_n appears, and outside some neighborhood of the origin in λ (i.e., for sufficiently small $|\omega^2\lambda^{-2/3}|$) E_n is analytic.

Just as we found earlier in applying complex scaling to the integral $I(\lambda)$ the scaling has helped describe part of the singularity, but here too a good deal of singular behavior remains. Aside from the three-sheet structure, the scaling has transformed the singularity in $E_n(\omega^2, \lambda)$ at $\lambda=0$ into a

singularity in $E_n(z, l)$ at large z . The singularity at ∞ cannot be an isolated singularity, and there is an infinity of other singularities approaching it. Presumably it was these that appeared in our steepest descent calculations.

NOTES

The Van Vleck determinant and its appearance in the path integral can be found in the following papers:

J. H. Van Vleck, *Proc. Nat. Acad. U. S. Sci.* **14**, 178 (1928)

C. Morette (DeWitt), On the definition and approximation of Feynman's path integrals, *Phys. Rev.* **81**, 848 (1952)

W. Pauli, "Feldquantisierung," lecture notes, Zurich, 1951, published as Vol. 6, *Pauli Lectures on Physics, Selected Topics in Field Quantization*, C. P. Enz, Ed., MIT Press, Cambridge, Mass., 1973

B. DeWitt, *Rev. Mod. Phys.* **29**, 377 (1957)

As indicated in the text, several proofs are available relating determinants of infinite matrices (i.e., $\delta^2 S$) to solutions of an associated differential equation (i.e., the Jacobi equation). The n -dimensional version of the proof given in Section 6 is found in

G. J. Papadopoulos, *Phys. Rev. D* **11**, 2870 (1975)

A different sort of proof can be found in

S. Levit and U. Smilansky, *Proc. Amer. Math. Soc.* **65**, 299 (1977)

A proof using complex function theory is given in an appendix of

S. Coleman, The Uses of Instantons, in *The Whys of Subnuclear Physics*, Erice 1977, A. Zichichi, Ed., Plenum, New York, 1979.

A generalization to higher-order differential equations is given by

T. Dreyfus and H. Dym, *Duke Math. J.* **45**, 15 (1978)

Asymptotic approximations for the Wiener integral are developed in

M. Schilder, *Trans. Amer. Math. Soc.* **125**, 63 (1966)

S. R. S. Varadhan, *Comm. Pure and Appl. Math.* **19**, 261 (1966)

The treatment of the anharmonic oscillator given in this section follows that of

C. S. Lam, *Nuovo Cimento* **47A**, 451 (1966); **50A**, 504 (1967)

Dyson's paper on the perturbation expansion in QED, is

F. Dyson, *Phys. Rev.* **85**, 631 (1952)

Diagrams of the form of Fig. 13.1 were used by

D. W. McLaughlin, *J. Math. Phys.* **13**, 1099 (1972)

The use of complex scaling for the anharmonic oscillator was developed by

B. Simon, *Ann. Phys.* **58**, 76 (1970)

Divergence of perturbation theory for the anharmonic oscillator was shown by

W. M. Frank, *J. Math. Phys.* **8**, 1121 (1967)

Regarding the indications we found for the existence of essential singularities in various functions, it should be noted that technically the appearance of a term $\exp(ic/\lambda)$ is not necessarily an essential singularity according to the strict definition. Rather, this quantity may appear as a discontinuity across a branch cut. Recent usage in physics has, however, to some extent condoned this terminology.

FOURTEEN

Detailed Presentation of the WKB Approximation

In this section we look a bit more closely at the WKB approximation, taking care of the details where that is feasible and trying to locate with more precision where the gaps remain. To some extent we are elaborating calculations made in Section 6, except that now we allow nonquadratic Lagrangians and try to justify their replacement by quadratic Lagrangians. The notation and methods also serve as preparation for the subsequent section on caustics, which takes up the problem of expressing the propagator in regions where the second variation of the action vanishes.

We work with a finite mesh— G defined as a limit of iterations—and let the mesh go to zero.

Define

$$G^{(N)}(b, T; a) = \left(\frac{m}{2\pi i\hbar\epsilon} \right)^{(N+1)/2} \int dx_1 \cdots dx_N \exp \left[\frac{iS^{(N)}(x_1, \dots, x_N)}{\hbar} \right] \quad (14.1)$$

$$S^{(N)}(x_1, \dots, x_N) = \sum_{j=0}^N \epsilon \left[\frac{1}{2} m \left(\frac{x_{j+1} - x_j}{\epsilon} \right)^2 - V(x_j) \right] \quad (14.2)$$

with

$$x_{N+1} = b, \quad x_0 = a, \quad \epsilon = \frac{T}{N+1} \quad (14.3)$$

By our earlier work

$$G(b, T; a) = \lim_{N \rightarrow \infty} G^{(N)}(b, T; a)$$

We have also already seen the n -dimensional generalization of the equation above (\mathbf{x} an n vector) and from time to time in the present section we display the n -dimensional version of various formulas.

For fixed N we perform a stationary phase approximation in the variables x_1, \dots, x_N and wish to obtain the leading contribution to G for $\hbar \rightarrow 0$. The equations

$$\frac{\partial S^{(N)}}{\partial x_i} = 0 \quad i=1, \dots, N \quad (14.4)$$

lead to the finite difference equations

$$m \left(\frac{x_{i+1} - 2x_i + x_{i-1}}{\epsilon^2} \right) = - \frac{\partial V(x_i)}{\partial x} \quad i=1, \dots, N \quad (14.5)$$

which are obviously approximations to the Euler-Lagrange equations of the continuous problem. However, when we imitate the heuristic proof of Section 13 and pick an extremum path about which to expand, it is the solution of (14.5) that we use, not the exact classical trajectory. Let a solution to (14.5) be denoted $\bar{x}_i, i=1, \dots, N$ (temporarily ignoring the question of multiple solutions). As usual, we expand S about \bar{x} , letting

$$y_i = x_i - \bar{x}_i \quad (y_0 = 0, y_{N+1} = 0) \quad (14.6)$$

The expansion assumes smooth behavior of V , and we obtain

$$S^{(N)}(x) = S^{(N)}(\bar{x} + y) = \bar{S}^{(N)} + \delta S + \frac{1}{2} \delta^2 S + \sum_{k=3}^{\infty} \frac{1}{k!} \delta^k S \quad (14.7)$$

The term $\bar{S}^{(N)}$ is just $S^{(N)}(\bar{x})$ and for $N \rightarrow \infty$ becomes the action along the classical path. The first variation is identically zero (no higher-order terms in ϵ neglected) because $\bar{x}_i, i=1, \dots, N$ satisfy (14.5). Had we used the exact classical path we would now have to worry about errors. In proving $\delta S = 0$ the integration by parts usually encountered in deriving the Euler-Lagrange equations becomes, for finite mesh, simply a change in dummy index of summation making use of the fact that $y_0 = y_{N+1} = 0$. The next term is

$$\delta^2 S[y] = \sum_{j=0}^N \left[\frac{m}{2\epsilon} (y_{j+1} - y_j)^2 - \epsilon W_j y_j^2 \right] \quad (14.8)$$

where

$$W_j = \frac{\partial^2 V(\bar{x}_j)}{\partial x^2} \quad (14.9)$$

For n space dimensions, W is a matrix, and the contribution to $\delta^2 S$ is

$$-\epsilon \sum_{j=0}^N \sum_{\alpha, \beta=1}^n \frac{\partial^2 V(\mathbf{x}_j)}{\partial x_\alpha \partial x_\beta} y_{j\alpha} y_{j\beta} \quad (14.10)$$

Finally, the higher terms in S are

$$\delta^m S[y] = -\epsilon \sum_{j=0}^N \frac{\partial^m V(\bar{x}_j)}{\partial x^m} y_j^m \quad (14.11)$$

with corresponding generalization for increased space dimension.

Recalling the results for the quadratic Lagrangian, we take up the diagonalization of $\delta^2 S$ and expect to find that the determinant of this quadratic form emerges in a straightforward way. Let the quantities $y_i, i=1, \dots, N$ be represented by a column vector η and define the matrix σ by

$$\sigma = \frac{m}{2\epsilon^2} \begin{pmatrix} 2 & -1 & 0 & \cdots & & & \textcircled{O} \\ -1 & 2 & -1 & 0 & \cdots & & \\ 0 & -1 & 2 & -1 & 0 & \cdots & \\ \vdots & & & \ddots & \ddots & \ddots & \\ & & & & & -1 & \\ \textcircled{O} & & & & & -1 & \\ & & & & & & 2 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} W_1 & & & & & \\ & \ddots & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ \textcircled{O} & & & & & W_N \end{pmatrix} \quad (14.12)$$

essentially as in Section 6. Then (T signifies transpose)

$$\frac{1}{2} \delta^2 S[y] = \epsilon \eta^T \sigma \eta \quad (14.13)$$

In Section 6 we were content to evaluate the determinant of σ , but now we wish to have more details and consider also its diagonalization.

First we quote some well known results about Jacobi matrices, that is, matrices of the form

$$A = \begin{pmatrix} a & b & & & & & \textcircled{O} \\ c & a & b & & & & \\ & c & & \ddots & & & \\ & & \ddots & \ddots & & & \\ & & & & \ddots & & \\ \textcircled{O} & & & & & c & b \\ & & & & & & a \end{pmatrix} \quad N \times N \quad (14.14)$$

The eigenvalues are

$$\lambda_k = a - 2\sqrt{bc} \cos k\theta, \quad k=1, \dots, N, \quad \theta = \frac{\pi}{N+1} \quad (14.15)$$

The k th eigenvector is

$$(\alpha_k)_j = \text{const.} \cdot \left(\frac{c}{b} \right)^{(j-1)/2} \sin(jk\theta) \quad (14.16)$$

For $a=2$, $b=c=-1$ (our case) these take the values

$$\lambda_k = 4 \sin^2 \left(\frac{k\theta}{2} \right) \quad (14.17)$$

$$(\alpha_k)_j = \sqrt{\frac{2}{N+1}} \sin(jk\theta) \quad (14.18)$$

The eigenvalue problem for σ is

$$\sigma \xi = l \xi \quad (14.19)$$

which is a discrete form of the Sturm-Liouville problem, (12.9). But (14.19) is just an $N \times N$ matrix eigenvalue problem for a real symmetric matrix, so no esoteric theorems are needed to conclude that there are N eigenvalues and eigenvectors; call them

$$\begin{aligned} l_i, \xi_i &\quad i=1, \dots, N \\ l_1 < l_2 < \dots < l_N \end{aligned} \quad (14.20)$$

where the ξ_i are column vectors with components $(\xi_i)_j$, $j=1, \dots, N$.

For j large enough for the influence of the potential to be small (but j not necessarily of order N) the l 's are simply related to the λ 's of (14.17), the main difference being the scaling with N due to the ϵ 's in the definition of σ . Thus

$$l_j = \frac{2m}{\epsilon^2} \sin^2 \left(\frac{j\theta}{2} \right) + (\text{errors due to } W) \quad \theta = \frac{\pi}{N+1}, \quad \epsilon = \frac{T}{N+1} \quad (14.21)$$

It follows that for j large, but still small compared to N ,

$$l_j \sim \frac{m}{2T^2} j^2 \pi^2 \quad (14.22)$$

a result reminiscent of the infinite square well Schrödinger equation. The ξ 's are orthogonal and normalized so that they are orthonormal in the $N \rightarrow \infty$ limit, that is,

$$\epsilon \xi_i^T \xi_k = \epsilon \sum_{j=1}^N (\xi_i)_j (\xi_k)_j = \delta_{ik} \quad (14.23)$$

[comparing (14.18), note that each $(\xi_i)_j$ is $O(1)$]. Let X be the $N \times N$ matrix whose (i, j) th component X_{ij} is $(\xi_j)_i$. X satisfies $XX^T = X^T X = 1/\epsilon$. Any vector η can be expanded in terms of the $\{\xi_i\}$:

$$\eta_i = \sum_{j=1}^N (\xi_j)_i c_j \quad \text{or} \quad \eta = Xc \quad (14.24)$$

where in the second equality the quantities c_i are thought of as a column vector. From (14.24) Parseval's relation takes the form

$$\epsilon \eta^T \eta = c^T c \quad (14.25)$$

(Definitions have been made so that the ϵ appears in the sum over η_i , since the index i is running over a variable that becomes continuous in the $N \rightarrow \infty$ limit. The index of c is an eigenstate label, which because of our "infinite potential" at $t=0$ and T remains discrete.) The quantity $\delta^2 S[y]$ now takes the simple form

$$\frac{1}{2} \delta^2 S = \epsilon \eta^T \sigma \eta = \epsilon c^T X^T \sigma X c = \sum_{j=1}^N c_j^2 l_j \quad (14.26)$$

The higher variations $\delta^3 S$, and so on, can be similarly written in terms of c , for example

$$\delta^3 S = - \sum_{p, q, r} c_p c_q c_r \left[\epsilon \sum_j \frac{\partial^3 V(\bar{x}_j)}{\partial x^3} (\xi_p)_j (\xi_q)_j (\xi_r)_j \right] \quad (14.27)$$

Changing variables from $dy_1 \cdots dy_N = d(\eta)_1 \cdots d(\eta)_N$ to integration over $dc_1 \cdots dc_N$ is trivial and introduces only factors of ϵ as can be seen by evaluating the Jacobian of the transformation:

$$\frac{\partial(\eta_1 \cdots \eta_N)}{\partial(c_1 \cdots c_N)} = |\det X| = \sqrt{\det(X^T X)} = \sqrt{\det(1/\epsilon)} = \epsilon^{-N/2} \quad (14.28)$$

The Green's function therefore has the form

$$G^{(N)}(b, T; a) = \left(\frac{m}{2\pi i\hbar} \right)^{(N+1)/2} e^{-N-1/2} \exp\left(\frac{i\bar{S}^{(N)}}{\hbar} \right) \\ \times \int dc_1 \cdots dc_N \exp\left[\frac{i}{\hbar} \sum l_i c_i^2 + \frac{i}{\hbar} \sum_{m=3}^{\infty} \frac{1}{m!} \delta^m S \right] \quad (14.29)$$

At this point it is clear what the stationary phase approximation is supposed to do—get rid of all $\delta^m S$ terms, with $m \geq 3$. Let us go through some quick reminiscing on what happens if that drastic step can be justified. Let us call $G^{(N)}$ without $\delta^m S$, $m \geq 3$, $G_2^{(N)}$. Then this truncated propagator is

$$G_2^{(N)}(b, T; a) \equiv \left(\frac{m}{2\pi i\hbar} \right)^{(N+1)/2} e^{-N-1/2} \\ \times \exp\left[\frac{i}{\hbar} \bar{S}^{(N)} \right] \int dc_1 \cdots dc_N \exp\left(\frac{i}{\hbar} \sum l_i c_i^2 \right) \quad (14.30)$$

The Gaussian integral gives

$$G_2^{(N)}(b, T; a) = \left(\frac{m}{2\pi i\hbar} \right)^{(N+1)/2} e^{-N-1/2} e^{i\bar{S}^{(N)}/\hbar} \prod_{k=1}^N \left(\frac{\pi i\hbar}{l_k} \right)^{1/2} \quad (14.31)$$

We already know from Section 6 what becomes of this in the limit $N \rightarrow \infty$, since $\prod l_k = \det \sigma$. In particular

$$G_2^{(N)}(b, T; a) = \sqrt{\frac{m}{2\pi i\hbar f(T)}} e^{iS(\bar{x})/\hbar} \quad (14.32)$$

where \bar{x} is the classical path arising (in the limit $N \rightarrow \infty$) from \bar{x}_k and f is the function satisfying

$$m \frac{\partial^2 f}{\partial t^2} + W(t)f(t) = 0, \quad f(0) = 0, \quad \frac{\partial f(0)}{\partial t} = 1 \quad (14.33)$$

and $W(t)$ is the limit of W_k .

Let us return to the problem of the terms $\delta^m S[y]$, $m \geq 3$. Because we are looking at the path integral and not the Wiener integral it is the

stationary phase approximation to which we ought to appeal. However, in our arguments favoring the dropping of the higher-order terms we shall pretend that we have the easier problem of the Wiener integral and content ourselves with justifying the Laplace method. Consider therefore

$$G^{(N)} = \left(\frac{m}{2\pi i\hbar} \right)^{(N+1)/2} e^{-N-1/2} \int dc_1 \cdots dc_N \times \exp \left[-\frac{1}{\mu} \sum I_i c_i^2 + \frac{1}{6\mu} \sum_{j_1 j_2 j_3} U_{j_1 j_2 j_3} c_{j_1} c_{j_2} c_{j_3} + \cdots \right] \quad (14.34)$$

with

$$U_{j_1 j_2 j_3} = -\varepsilon \sum_k \frac{\partial^3 V(\bar{x}_k)}{\partial x^3} (\xi_{j_1})_k (\xi_{j_2})_k (\xi_{j_3})_k \quad (14.35)$$

and where μ has replaced $(i\hbar)$ but will be considered a positive real parameter whose asymptotic limit $\mu \rightarrow 0$ is being studied.

Our goal is to generalize arguments of the sort given in Section eleven. There we showed that for $f(x) = x^2 o(1)$ the function

$$F(\lambda) = 2\sqrt{\frac{\lambda}{\pi}} \int_0^\infty e^{-\lambda(x^2 + f(x))} dx \quad (14.36)$$

has the asymptotic limit 1 for $\lambda \rightarrow \infty$. In fact we had the estimate

$$\begin{aligned} \frac{1}{2}(1 - F(\lambda)) &\leq \Delta_1 + \Delta_2 \\ \Delta_1 &\leq \lambda a^2 \eta(a) \quad \text{with } \eta(a) = \frac{1}{a^2} \max_{0 < x < a} f(x) \\ \Delta_2 &\leq \frac{e^{-\lambda a^2}}{\sqrt{2\pi\lambda a^2}} \end{aligned} \quad (14.37)$$

where the quantity a is a number that enters when the integral from 0 to ∞ is broken into two pieces, $(0, a)$ and (a, ∞) . By judicious choice of a , such as $a = (\log \lambda)/\sqrt{\lambda}$, it follows that $F \rightarrow 1$ (assuming η goes to zero fast enough).

To show that $G^{(N)}$ of (14.34) becomes, to leading order in μ , just $G_2^{(N)}$ of (14.30) we have N integrals of the sort (14.36), and not only must we control their individual errors, but we must show that the sum of all errors goes to zero with μ too.

Our goal is therefore to show that the quantity

$$F(u) = \frac{\int_{-\infty}^{\infty} dc_1 \cdots \int_{-\infty}^{\infty} dc_N \exp\left(-\frac{1}{\mu} \sum c_i^2 l_i + \frac{1}{6\mu} \sum c_i c_j c_k U_{ijk} + \dots\right)}{\int_{-\infty}^{\infty} dc_1 \cdots \int_{-\infty}^{\infty} dc_N \exp\left(-\frac{1}{\mu} \sum c_i^2 l_i\right)} \quad (14.38)$$

approaches unity as $\mu \rightarrow 0$. The dots in the numerator stand for terms $\delta^m S$, $m \geq 4$. As in Section 11, we shall split the problem into two pieces. First we shall see that breaking off the integrations at $|c_i| \leq a_i$ for sufficiently large a_i does not change the asymptotic limit. Then we shall show that if the a_i are not taken *too* large the cubic and higher terms are dominated by the quadratic term.

Suppose then that each integral is cut off at a_i . The error is essentially (assuming as usual that for large $|c_i|$ the integral converges nicely)

$$\Delta_2 = \sum_i \frac{\exp(-l_i a_i^2 / \mu)}{\sqrt{l_i a_i^2 / \mu}} \quad (14.39)$$

Just as in Section 11, it is clear that a_i going to zero as $\sqrt{\mu/l_i}$ is *almost* good enough, but that the a_i must go to zero just a bit more slowly to guarantee that Δ_2 goes to zero.

The other side of the problem is to guarantee that within the set $|c_i| \leq a_i$, $i = 1, \dots, N$ the quadratic term dominates higher-order terms and that the higher terms are in fact small compared to unity. We shall look only at the worst term, the cubic term. Again the error is essentially

$$\Delta_1 \leq \max_{|c_i| \leq a_i} \frac{1}{\mu} \sum c_i c_j c_k U_{ijk} \leq \frac{1}{\mu} \sum a_i a_j a_k |U_{ijk}| \quad (14.40)$$

The U 's are Fourier components of the function $\partial^3 V / \partial x^3$ evaluated along the classical path. As such their drop-off with increasing i, j , or k is determined by the smoothness of $\partial^3 V / \partial x^3$ through the Riemann-Lebesgue lemma. Having no ambitions in the direction of great generality, we assume that for some number κ

$$|U_{jkl}| \leq \frac{\kappa}{jkl} \quad (14.41)$$

This assumption can be justified by giving V enough derivatives, but we

shall not go into details as the actual conditions for the validity of what we are doing probably require very little of V . (One remark: the argument of V is the classical path, which is generally smooth and not a Brownian motion type path; hence this does not slow the decline of $|U_{ijk}|$ with increasing magnitude of its subscripts.) Then

$$\Delta_1 \leq \frac{1}{\mu} \sum_{jkl} \frac{\kappa a_j a_k a_l}{jkl} = \frac{\kappa}{\mu} \left(\sum_j \frac{a_j}{j} \right)^3 \quad (14.42)$$

To emphasize how close a_j can be to $\sqrt{\mu/l_j}$ we define

$$a_j = b_j \sqrt{\mu/l_j} \quad (14.43)$$

and show that by letting $b_j \rightarrow \infty$ very slowly, convergence is achieved.

The errors are then

$$\Delta_1 + \Delta_2 \leq \sum_j \frac{e^{-b_j^2}}{b_j} + \kappa \sqrt{\mu} \left(\sum_j \frac{b_j}{j \sqrt{l_j}} \right)^3 \quad (14.44)$$

By taking

$$b_j = \left[\alpha \log \left(\frac{j}{\mu} \right) \right]^{1/2} \quad \alpha > 1 \quad (14.45)$$

the errors clearly go to zero with μ [recalling that $l_j = O(j^2)$ for large j]. The wastefully rapid convergence of Δ_1 is undoubtedly due to the overly restrictive assumptions on V leading to (14.41).

From this demonstration we can also see in what way a proof of the stationary phase approximation would present more difficulty. With pure imaginary μ , the term $\exp(-l_i a_i^2/\mu)$ would be of no use in reducing Δ_2 . The b_j 's would have to go to infinity faster than j for Δ_2 to converge; but then Δ_1 diverges. Hence this proof would not go through, and one presumably needs to make use of cancellations due to the phases of various terms, a much finer question. Moreover, I would expect that the proof in the imaginary case goes through most easily not for G itself but for $G\psi$ for some ample collection of well-behaved ψ 's. The ψ would make sure that the "wildly oscillating phases" to which one appeals in stationary phase arguments in fact cancel. In Hagedorn's method for obtaining the semiclassical limit (see reference below) I believe the coherent states that he defines serve in that capacity.

NOTES

There are mathematical works dealing with asymptotic approximations of the sort considered in this section. For the Wiener integral there is the work of

M. Schilder, *Trans. Amer. Math. Soc.* **125**, 63 (1966)

Related to this are studies of solutions of the heat equation by

S. R. S. Varadhan, *Commun. Pure and Appl. Math.* **19**, 261 (1966)

For the path integral, the asymptotic approximation is an extension of the more sensitive method of stationary phase. This problem has been studied by

S. Albeverio and R. Hoegh-Krohn, *Inventiones Math.* **40**, 59 (1977)

Their development is based on a formulation of the path integral by the same authors

S. Albeverio and R. Hoegh-Krohn, *Mathematical Theory of Feynman Path Integrals*, Lecture notes in Math 523, Springer, Berlin, 1976

A derivation of the semiclassical approximation based on the Trotter product formula is

G. A. Hagedorn, *Commun. Math. Phys.* **71**, 77 (1980)

FIFTEEN

WKB Near Caustics

From our work on the WKB approximation it is clear that an interesting sort of trouble can arise if the function $f(t)$ of (14.33) vanishes at $t=T$ or equivalently if $\partial^2 S/\partial a \partial b$ is infinite. In either the discrete (finite N) or the continuous case, this vanishing has been found to be due to the vanishing of one or more eigenvalues of $\delta^2 S^{(N)}$ (or $\delta^2 S$). For the quadratic Lagrangian we were stuck with this blowup. It meant that there was perfect focusing and the Green's function became a δ -function. But now with $\delta^3 S$ terms available, the blowup occurs only in an approximation and in fact the Green's function remains finite.

Consider, then, values a , b , and T such that one of the eigenvalues of $\delta^2 S$ along a particular classical path vanishes. Thus b at time T is a conjugate point for the trajectory leaving a at time 0. Holding $(a, 0)$ fixed, we consider the Green's function in a neighborhood of b , $G(b+\Delta b, T; a)$. Again we work with $G^{(N)}$, but we do not worry about distinguishing $\bar{x}(t)$, the exact trajectory connecting the conjugate points, and \bar{x}_k , the solution of the discrete problem—for which the vanishing eigenvalue may differ from zero by terms of $O(\varepsilon)$.^{*} Our expansion will differ from those we considered previously. Paths from $(a, 0)$ to $(b+\Delta b, T)$ will not be expanded about the classical path connecting those endpoints but rather about \bar{x} . This means that the variational path does not vanish at the upper limit, and we have some extra terms to carry along. Next define the path $\rho(t)$ to be some fixed arbitrary curve with $\rho(0)=0$, $\rho(T)=\Delta b$; ρ can be chosen in some convenient way, but for our purposes it is left arbitrary. The path $x(\tau)$, the “integration variable” of the path integral, is written

$$x = \bar{x} + \rho + y \tag{15.1}$$

so that $y(0)=0$, $y(T)=0$. As usual by a translation the new integration

*In fact eigenvalues of σ of (14.12) are only shifted by $O(\varepsilon^2)$.

variable is $y(\tau)$, and again we expand

$$S(x) = S(\bar{x}) + \delta S[\rho + y] + \frac{1}{2} \delta^2 S[\rho + y] + \sum_{m=3}^{\infty} \frac{1}{m!} \delta^m S[\rho + y] \quad (15.2)$$

Now, however, δS does not vanish; in particular

$$\delta S[\rho + y] = \delta S[\rho] = \left[\frac{\partial L}{\partial \dot{x}} \Big|_{\bar{x}} \rho(t) \right]_0^T = \bar{p}(T) \Delta b \quad (15.3)$$

\bar{p} is the momentum of \bar{x} at T and for n dimensions it is the vector dot product that appears in (15.3). The second variation is

$$\begin{aligned} \delta^2 S[\rho + y] &= \int_0^T dt [m(\dot{\rho} + \dot{y})^2 - W(\rho + y)^2] \\ &= \delta^2 S[\rho] + 2 \int_0^T (m\dot{\rho}\dot{y} - W\rho y) dt + \delta^2 S[y] \end{aligned} \quad (15.4)$$

(see (14.9) for the definition of W).

The quantity $\delta^2 S[\rho]$ is some fixed number which we ignore for now. In the cross term of (15.4) we do an integration by parts to obtain

$$\text{cross term} = \int_0^T (m\dot{\rho}\dot{y} - W\rho y) dt = m\rho\dot{y}|_0^T - \int_0^T (m\ddot{y} + Wy)\rho dt \quad (15.5)$$

The function y has the same boundary conditions as the Sturm-Liouville eigenfunctions (see discussion at (12.9)–(12.11)) and can be expanded in terms of them:

$$y(t) = \sum_j a_j \varphi_j(t) \quad (15.6)$$

Call the eigenvalues $\lambda_1, \lambda_2, \dots$, with $\lambda_1 < \lambda_2 < \dots$. Again we have

$$\delta^2 S[y] = \sum_j \lambda_j a_j^2 \quad (15.7)$$

but now $\lambda_1 = 0$, because we are at a conjugate point. Thus a_1^2 does not appear in $\delta^2 S$. Equation 15.5 can also be rewritten, and becomes

$$\text{cross term} = m \sum_j a_j \Delta b \dot{\varphi}_j(T) + \sum_j \lambda_j a_j \int_0^T \varphi_j(t) \rho(t) dt \quad (15.8)$$

We next change variables to $\{a_i\}$ instead of $\{y_i\}$ but ignore details of the Jacobian, which would have been calculable had we been strict in remembering that we are dealing with $G^{(N)}$. We have obtained therefore the

following form for the Green's function:

$$\begin{aligned}
 G(b+\Delta b, T; a) = & \int da_1 \cdots da_N \exp \left\{ \frac{i}{\hbar} \left[S(\bar{x}) + \bar{p}(T) \Delta b + \frac{1}{2} \delta^2 S[\rho] \right. \right. \\
 & + m \sum_j a_j \Delta b \dot{\varphi}_j(T) + \sum_j \lambda_j a_j \int_0^T \varphi_j(t) \rho(t) dt \\
 & \left. \left. + \frac{1}{2} \sum_j \lambda_j a_j^2 + \sum_{m=3}^{\infty} \frac{1}{m!} \delta^m S[\rho+y] \right] \right\} \\
 \end{aligned} \tag{15.9}$$

What is new and interesting about this integral is the absence of an a_1^2 term in the exponent so that for $\Delta b=0$ the leading term in a_1 is cubic. However, to see the consequences of this feature a good deal of underbrush must be cleared away, namely the other $N-1$ integrals over a_2, \dots, a_N . We do this in two ways: one, by using previous experience with asymptotic limits to assign powers of \hbar to various contributions and then noting that everything works out self consistently. We apply this method only for Δb small. The second way appeals to some theorems from the general body of knowledge known as catastrophe theory to obtain a normal form for the integrand for which the limitation to small Δb can be lifted so as to obtain a "uniform asymptotic approximation." That is done in Section 16.

In our first approach to the integral (15.9) we look upon the argument of the exponent as a power series in the a 's. In previous sections we discarded terms of $O(a_i a_j a_k)/\hbar$ as being $O(\hbar^{3/2}/\hbar)=O(\hbar^{1/2})$ relative to the leading quadratic term. Now, with a_1^2 missing, the procedure must be reexamined.

First let us estimate the size of a_1 . Consider an integral of the form

$$F_k(\mu) = \int \exp(i\mu x^3) x^k dx$$

By rescaling, this becomes (assuming that the limits of integration are unimportant)

$$F_k(\mu) = \mu^{-k/3} \mu^{-1/3} F_k(1)$$

so that each power of x contributes $\mu^{-1/3}$. This suggests that for $\hbar \rightarrow 0$, a_1 is $O(\hbar^{1/3})$. Next consider Δb and ρ . By its definition, ρ can be selected so that $O(\rho)=O(\Delta b)$. Δb is an external parameter and is at our disposal. Since we are only seeking a description of the Green's function in the

critical region, it is enough to consider Δb such that $a_1 \Delta b = O(\hbar)$. This is because the principal a_1 contributions to the integral are from a_1^3 and $a_1 \Delta b$. When these are of the same order the character of the integral is no longer dominated by the cubic (in fact, it can be rewritten as a quadratic with coefficient $\sim (\Delta b)^{1/2}$). Hence the critical region—where G is most singular—is characterized by $\Delta b \lesssim \hbar^{2/3}$, and we restrict our attention to this domain.

Terms a_j^2 for $j \neq 1$ are the usual quadratic terms which will ultimately dominate the integrals $\int da_j$, $j \neq 1$. We therefore assume that we can proceed in the ordinary way and take $a_j = O(\hbar^{1/2})$, $j \neq 1$.

Next we must estimate the linear terms in a_j , $j \neq 1$ and higher-order terms involving all a_j 's that appear in $\delta^k S$, $k \geq 3$. Linear terms $a_j \Delta b$ and $a_j \int \varphi \rho dt$, $j \neq 1$, are unimportant to leading order in \hbar since $a_j = O(\hbar^{1/2})$ and $\Delta b = O(\hbar^{2/3})$. Hence this is smaller by $\hbar^{1/6}$ than the quadratic term a_j^2 which gives the dominant contribution. Of the higher-order terms we consider only the third variation, which is

$$\delta^3 S[\rho + y] = - \int_0^T dt \frac{\partial^3 V}{\partial x^3} \left[\rho(t) + \sum_j a_j \varphi_j(t) \right]^3 \quad (15.10)$$

and which includes terms in the following 10 categories:

- I. a_1^3
- II. $a_1^2 a_j \quad j \neq 1$
- III. $a_1 a_j a_k \quad j, k \neq 1$
- IV. $a_j a_k a_n \quad j, k, n \neq 1$
- V. $a_1^2 \int \rho$
- VI. $a_1 a_j \int \rho \quad j \neq 1$
- VII. $a_j a_k \int \rho \quad j, k \neq 1$
- VIII. $a_1 (\int \rho)^2$
- IX. $a_j (\int \rho)^2 \quad j \neq 1$
- X. $(\int \rho)^3$

The term a_1^3 is of course the important contribution for the a_1 integral. By our previous estimate it is $O(\hbar)$. By going through the rest of the list it is immediately evident that all other terms are $O(\hbar^{7/6})$ or smaller. It is also clear that the worst terms in $\delta^k S$, $k \geq 4$, are $O(\hbar^{4/3})$ and so need not be considered.

To leading order in \hbar the Green's function is therefore

$$G(b+\Delta b, T; a) = \exp \left\{ \frac{i}{\hbar} [S(\bar{x}) + \bar{p}(T) \Delta b] \right\}$$

$$\times \int da_1 \dots da_N \exp \left\{ \frac{i}{\hbar} \left[\frac{u}{3} a_1^3 + v a_1 + \frac{1}{2} \sum_{j=2}^N \lambda_j a_j^2 \right] \right\} \quad (15.11)$$

with

$$u = -\frac{1}{2} \int_0^T \frac{\partial^3 V}{\partial x^3} [\varphi_1(t)]^3 dt \quad (15.12)$$

$$v = m \Delta b \dot{\varphi}_1(T) \quad (15.13)$$

The integrals over a_j , $j \neq 1$ will as usual give $(\Pi' \lambda_j)^{-1/2}$, the prime over Π denoting that the term $j=1$ is excluded. Recalling that the product over all λ_j gives a quantity $f(T)$ related to a certain solution of the Jacobi equation [cf. Section 6 or (14.32) and 14.33)], we write $\Pi' \lambda_j = f(T)/\lambda_1$. Although for the conjugate point both $f(T)$ and λ_1 vanish, the ratio is finite. In fact it is just by looking at that ratio that one can sometimes compute $\Pi' \lambda_j$. (Such a calculation occurs in Section 29 in connection with "instantons.")

There remains only the integral over the cubic polynomial. Recall the definition of the Airy function

$$Ai(z) = \frac{1}{2\pi i} \int_{\Gamma} \exp(tz - \frac{1}{3}t^3) dt \quad (15.14)$$

with the contour Γ shown in Fig. 15.1. This can also be written as

$$Ai(z) = \frac{1}{\pi} \int_0^{\infty} \cos \left(\frac{s^3}{3} + zs \right) ds \quad (15.14a)$$

Some properties of the Airy function are

$$\frac{d^2 Ai(z)}{dz^2} - z Ai(z) = 0$$

$$Ai(-z) = \frac{1}{3} \sqrt{z} \left[J_{1/3} \left(\frac{2}{3} z^{3/2} \right) + J_{-1/3} \left(\frac{2}{3} z^{3/2} \right) \right]$$

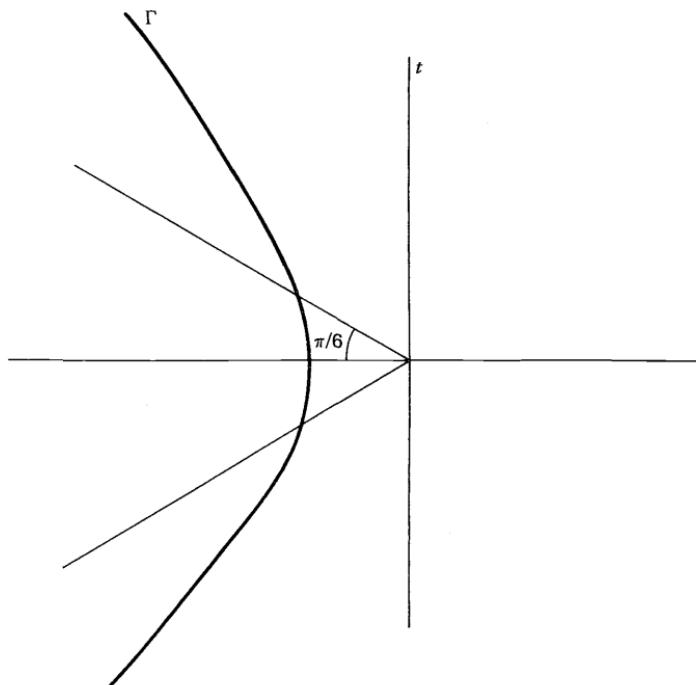


Fig. 15.1 Contour Γ in the complex t plane for evaluating the Airy function (15.14). Γ goes to infinity asymptotically in the left half plane and at angles exceeding $\pi/6$ to the negative real t axis.

(J is the Bessel function). For real z

$$Ai(z) = \begin{cases} \frac{1}{2\sqrt{\pi} z^{1/4}} e^{-(2/3)z^{3/2}} (1 + O(z^{-3/2})), & z \rightarrow \infty \\ \frac{1}{\sqrt{\pi} (-z)^{1/4}} \sin\left[\frac{2}{3}(-z)^{3/2} + \frac{\pi}{4}\right] (1 + O(|z|^{-3/2})), & z \rightarrow -\infty \end{cases} \quad (15.15)$$

It follows that the cubic integral is just an Airy function, and by comparing (15.11) and (15.14) we note that G contains the factor $Ai(vu^{-1/3}\hbar^{-2/3})$. One cosmetic step remains before giving our final form of G . The quantity $\bar{p}(T)$ is the momentum along \bar{x} at the time T and is therefore equal to $\partial S(\bar{x})/\partial x_{\text{final}}$. Therefore

$$S(\bar{x}) + \bar{p}(T)\Delta b = S(\text{class. path } a \rightarrow b + \Delta b) \equiv S_c(b + \Delta b, T; a)$$

The Green's function to leading order in \hbar is

$$G = \sqrt{\frac{\lambda_1}{f(T)}} \exp\left[\frac{iS_c(b+\Delta b, T; a)}{\hbar}\right] Ai\left(\frac{v}{u^{1/3}} \hbar^{-2/3}\right) \quad (15.16)$$

(leaving out factors π , etc.).

The expression (15.16) does not differ by much from that which we derived in Section 16 without the restriction $\Delta b \gtrsim \hbar^{2/3}$ (and therefore $v \lesssim \hbar^{2/3}$). Hence we examine (15.16) for qualitative information on G . The limit $\hbar \rightarrow 0$ corresponds to $|z|$ [of (15.14)] going to infinity. Consequently (15.15) suggests that the behavior of G (in the limit $\hbar \rightarrow 0$) depends strongly on the sign of $v/u^{1/3}$. For $v/u^{1/3} > 0$ the Green's function has an exponential decay—going to zero as $\exp[-(\Delta b)^{3/2} \cdot \text{const.}]$ —while for v , having the opposite sign, G has a sinusoidal dependence. The changing sign of $v/u^{1/3}$ is just that of Δb [cf. (15.13)] as one goes through the conjugate point. Thus one side of the conjugate point is in “shadow.” The Green's function, hence the probability of the system's being there is small. On the other side is “illumination,” that is, high probability of finding the system. The “brightest” region is just near the conjugate point. There the asymptotic expansions for the Airy functions cannot be used. For example, at $x=b$ ($\Delta b=0$)

$$G \sim \int e^{iua_i^3/3\hbar} da_i \sim \left(\frac{\hbar}{u}\right)^{1/3} \quad (15.17)$$

This is large, since the typical contribution of each degree of freedom to G is of order $\hbar^{1/2}$ rather than $\hbar^{1/3}$. Thus within the critical region of the conjugate point there is a greater propensity for nonclassical effects reflected in the appearance of lower powers of \hbar .

We have given arguments in one space dimension. In higher dimension “ v ” depends on

$$v = m\Delta \mathbf{b} \cdot \dot{\varphi}_1(T)$$

where $\varphi_1(t)$ is the solution of the appropriate vector Jacobi equation. In general, there will be a surface of conjugate points; on one side of the surface will be “illumination” and on the other side shadow. Such a surface in optics is known as caustic (see Fig. 15.2). We shall have more to say below about caustics, as well as their relation to catastrophes, but for now let us consider (15.16) a bit more closely.

The phenomenon revealed by G of (15.16), the passage from a high probability region to a low probability region as one crosses the conjugate

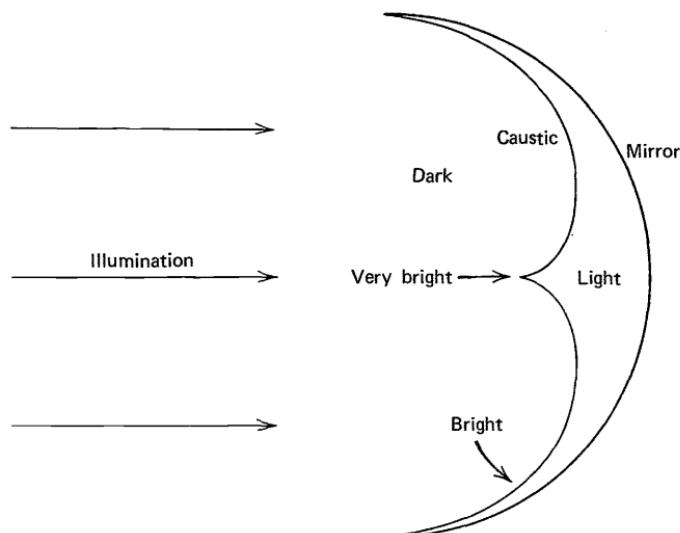


Fig. 15.2 A commonly observed caustic—one that arises from a cylindrical mirror with circular cross section. This is often seen on the surface of a cup of coffee under appropriate conditions of illumination.

point, does not present a new mystery, but rather solves an old problem that we have swept under the rug again and again in the course of this work. In using the semiclassical approximation we have always required the classical paths (solutions of the Euler-Lagrange equations) to be well separated from one another. The Green's function we have just derived is precisely that appropriate for the case of two coalescing classical paths.

First a heuristic justification of the statement above: we digress a bit, to say what problem the Airy function solves. By a change of variable we can write

$$Ai(\alpha x^{2/3}) = \frac{x^{1/3}}{2\pi i} \int_{\Gamma} \exp\left[x\left(\alpha y - \frac{1}{3}y^3\right)\right] dy \quad (15.18)$$

with $x > 0$. For fixed α and large x one of the two forms of (15.15) is obtained. These two forms arise because when one uses steepest descent methods for (15.18) he encounters the equation

$$0 = \frac{\partial}{\partial y} \left(\alpha y - \frac{1}{3}y^3 \right) = \alpha - y^2$$

For $\alpha > 0$ this has no solution on Γ (cf. Fig. 15.1) while for $\alpha < 0$ there are two solutions. For $\alpha < 0$ one therefore obtains the larger values of Ai while

for $\alpha < 0$ the integral has only an exponentially small term arising from a critical point off the range of integration. If α is allowed to vary, and in particular if it can be small on the order of $x^{-2/3}$, then all bets are off. The asymptotic approximations do not give useful results, and the best that one can say is "it's an Airy function." The trouble of course arises because of the coalescing of the two critical points $y = \pm \sqrt{-\alpha}$. For α small their contributions interfere (so the asymptotic approximation which writes A_i as a sum of two distinct terms is inaccurate) while for the wrong sign of α , although the appropriate value of y does not appear on Γ , it is not far away.

For the path integral it is zeros of δS that are the critical points and for the given initial point and the given T the classical mechanics (boundary value problem) has two solutions on one side of the conjugate point and none on the other. To see this rigorously, observe that the argument of the exponent in (15.9) is just the classical action, expressed in rather peculiar coordinates, the variables a_j , $j = 1, \dots, N$. The quantities \bar{x} , \bar{p} , φ_j , and so on, are fixed, and the problem of finding a solution to $\delta S = 0$ is the same as requiring that $\partial/\partial a_j$ vanish for each j . For $j \neq 1$ there is always a solution because the derivative of the leading quadratic term $\frac{1}{2} \lambda_j a_j^2$ gives a linear equation for a_j which has a solution near zero (for small Δb). For a_1 , referring to (15.11), the equation is $\partial(ua_1^3/3 + va_1)/\partial a_1 = 0$ which may have zero or two solutions depending on the sign of v/u ($v \neq 0$). This justifies the description of the phenomenon in (15.16) as the coalescing of classical paths.

There is an obvious generalization of the results obtained so far. Suppose there is a region in coordinate space where u of (15.12) vanishes along with λ_1 so that the leading term in a_1 is a_1^4 . Then the propagator can be brought to the form

$$G(\mathbf{b} + \Delta\mathbf{b}, T; \mathbf{a}) = \exp(iS_c(\mathbf{b} + \Delta\mathbf{b}, T; \mathbf{a})/\hbar) \int da_1 \exp\left[\frac{i}{\hbar}\left(\frac{a_1^4}{4} + \frac{v_2 a_1^2}{2} + v_1 a_1\right)\right] \quad (15.19)$$

with v_1 and v_2 dependent on $\Delta\mathbf{b}$ and vanishing when $\Delta\mathbf{b} = 0$. The point \mathbf{b} is even more brightly illuminated than in previous cases. Moreover, in some directions out of the point \mathbf{b} the polynomial

$$P(a_1) = \frac{1}{4}a_1^4 + \frac{1}{2}v_2 a_1^2 + v_1 a_1 \quad (15.20)$$

has points where it is stationary and also its second derivative vanishes. This means that by a shift in variable we are again on a caustic but one of

lower order, one in which a cubic polynomial determines the behavior. The equation of this caustic is obtained by finding those (v_1, v_2) for which the equations $dP/da_1 = d^2P/da_1^2 = 0$ can be solved simultaneously. The required relation between v_1 and v_2 is

$$\frac{1}{27}v_2^3 + \frac{1}{4}v_1^2 = 0 \quad (15.21)$$

which is a cusp. The quantities v_1 and v_2 are definite calculated functions of $\Delta\mathbf{b}$, v_1 being a coordinate in the direction $\dot{\phi}_1(T)$, and v_2 a coordinate in a perpendicular direction. Thus this particular caustic assumes a specific form, that of the cusp. See the "very bright" point and its neighborhood in Fig. 15.2.

Suppose that at some \mathbf{b} , both λ_1 and λ_2 are zero. G will then involve a double integral, and it is obvious that terms in a_1 and a_2 will appear with coefficients related to $\Delta\mathbf{b} \cdot \dot{\phi}_1(T)$ and $\Delta\mathbf{b} \cdot \dot{\phi}_2(T)$. But this is not all. The component of $\Delta\mathbf{b}$ along the direction $\dot{\phi}_1(T) \times \dot{\phi}_2(T)$ will also control the development of the caustic due to terms from $\delta^3 S$. The propagator can therefore be written

$$G \sim e^{iS_c/\hbar} \int da_1 da_2 e^{iP/\hbar} \quad (15.22)$$

where P can be brought to the form

$$P(a_1, a_2) = a_1^3 + a_2^3 - xa_1 - ya_2 - za_1 a_2 \quad (15.23)$$

We can in this way obtain an entire class of functions giving a description of the wave function near a caustic. Those of the kind presented in (15.20) have been given the name "generalized Airy functions," and it would seem that the propagator of (15.22) with a polynomial in several variables would warrant such a designation also.

The picture that emerges is that a certain polynomial [e.g., (15.23)] by virtue of its appearing in an integral with a large parameter governs the form taken by the caustic. Specifically, several of the coefficients of the lower powers in the polynomial are functions of position. The location of the caustic corresponds to the values of the coefficients for which the polynomial has simultaneously vanishing first and second derivatives.

To anyone exposed to catastrophe theory these results declare in no uncertain terms that caustics are catastrophes. To go into an exposition of catastrophe theory at this point would be too great a digression, so our next few remarks will have to be obscure to those unfamiliar with the subject. (In this we shall be loyal to the traditions of catastrophe theory: it always sounds like a snow job.)

The foregoing work has in effect been an explicit example of the successful operation of the principles of catastrophe theory, namely morphology has been determined by a breakdown in structural stability. The caustics are located by finding those values of the coefficients of the polynomials P for which P , considered as a potential, is not structurally stable. Moreover the theory provides us with a map from the space of coefficients to physical coordinate space. The form of the caustic is not sensitive to the details of this map, and so for example the cusp of (15.21) may be expected to be a common sort of caustic: not in its detailed global behavior but in its form near the singularity.

The present development is also an object lesson in the limitations of catastrophe theory. Although the loci of breakdown in structural stability (which is all catastrophe theory can give you) are correctly obtained, the details of the propagator (e.g., how much brighter is the caustic region than other regions) can only be obtained by looking at the underlying theory (in this case quantum mechanics) governing the phenomenon.

In the next section we turn to the problem of obtaining the propagator at larger distances from the caustic, that is, relaxing the condition $\Delta b \gtrsim \hbar^{2/3}$ and finding "uniform" approximations. Oddly enough I believe that here an important mathematical role is played by one of the theorems used in catastrophe theory, but not the much publicized classification of the "seven elementary catastrophes."

NOTES

Section 15 is an expansion of parts of the following paper

L. S. Schulman, Caustics and Multivaluedness: Two Results of Adding Path Amplitudes, in *Functional Integration and Its Applications*, Proceedings of a 1974 conference, A. M. Arthurs, Ed., Oxford University Press, London, 1975

As mentioned there, parts of that work were in turn based on the dissertations of Steve Coyne and David McLaughlin:

S. Coyne, Semi-Classical Asymptotic Evaluation of Feynman Path Integrals, thesis, Indiana University, Bloomington, Ind., 1972

D. W. McLaughlin, The Path Integral: Its Approximation in Flat Space and its Representation in Curved Space, thesis, Indiana University, Bloomington, Ind., 1970

In obtaining the canonical form of a fourth-order caustic, (15.19), we ignored a subtlety that did not occur for the third-order caustic. For third-order caustics we went through a list of terms that appear in the action [just following (15.10)] and determined that only the term involving a_1^3 need be considered. A similar list for a fourth-order caustic would include terms $a_1^2 a_k$ ($k = 2, 3, \dots$) in the action. These are of order \hbar , the

same as a_1^4 , and cannot be dropped. However, these terms do not affect the validity of (15.19), and if no convergence problems erupt they simply serve to change the coefficient of the quartic term. To see this, consider what happens to such a term when the integral over $a_k (k \neq 1)$ is performed

$$\int da_k \exp \left[\frac{i}{\hbar} \left(\frac{1}{2} \lambda_k a_k^2 + \frac{1}{2} a_1^2 a_k \int dt \frac{\partial^3 V}{\partial x^3} \Big|_{\bar{x}} \varphi_1^2 \varphi_k \right) \right] \\ = \sqrt{\frac{2\pi\hbar i}{\lambda_k}} \exp \left[\frac{ia_1^4}{8\hbar\lambda_k} \left(\int dt \frac{\partial^3 V}{\partial x^3} \Big|_{\bar{x}} \varphi_1^2 \varphi_k \right)^2 \right]$$

Consequently the coefficient of ia_1^4/\hbar is

$$\left[\frac{1}{4!} \int \frac{\partial^4 V}{\partial x^4} \Big|_{\bar{x}} \varphi_1^4 dt \right] - \frac{1}{8} \sum_k \frac{\left[\int \frac{\partial^3 V}{\partial x^3} \Big|_{\bar{x}} \varphi_1^2 \varphi_k dt \right]^2}{\lambda_k}$$

rather than merely the first term alone. To get the form of (15.19) one must in any case perform a rescaling of a_1 and provided the above mentioned sum exists this can be done.

For caustics of yet higher order the same sort of thing is feasible. What is important is that one should never get terms of the form $a_1^n a_j a_k (1 < j < k)$ to be significant, for then one could not perform the integrals $da_m, m > 1$ one at a time to eliminate such terms. But both a_j and a_k are $O(\hbar^{1/2})$ and so this problem cannot develop.

The argument just given for the $a_1^2 a_k$ terms going over to a_1^4 terms makes no pretense of rigor and as often happens in asymptotic analysis the desired result is best obtained by the use of the implicit function theorem. The coordinate system that we use for integration was gotten from the diagonalization of $\delta^2 S$, whereas for the elimination of the cross terms a slightly different coordinate system would be best. The existence of such a coordinate system is shown in

R. S. Ellis and J. S. Rosen, Asymptotic Expansions of Gaussian Integrals, *Bull. Amer. Math. Soc.*, 3, 705, 1980

Ellis and Rosen were looking at the Wiener integral and were able to extend rigorous results on the asymptotic expansion of the Wiener integral (see the Schilder reference in Section 14) to the case of degenerate and coalescing minima. Because they looked at the real Wiener integral they never had the privilege of treating the “easy” cubic case and used the implicit function theorem to show that at degenerate critical points the measure picks up a non Gaussian part, like our a_1^4 term.

The derivation of the caustic-catastrophe connection through path integration has proved to be an attractive method, and further work on this theme has been done by

Cecile DeWitt-Morette and P. Tschumi, Catastrophes in Lagrangian Systems, in *Long Time Predictions in Dynamics*, ed. V. Szebehely and B. D. Tapley, Reidel, Dordrecht, Holland, 1976, pp. 57–69

Cecile DeWitt-Morette, The Semiclassical Expansion, *Ann. Phys.* **97**, 367 (1976)

S. Levit, U. Smilansky, and D. Pelte, A New Semiclassical Theory for Multiple Coulomb Excitation, *Phys. Lett.* **53B**, 39 (1974)

A good expository article on catastrophe theory, containing many references, is

M. Golubitsky, An Introduction to Catastrophe Theory and its Applications, *SIAM Review* **20**, 352 (1978)

Caustics and Uniform Asymptotic Approximations

In Section 15 we had the following integral to evaluate

$$\begin{aligned}
 G(b+\Delta b, T; a) = & \int da_1 \dots da_N \exp \left\{ \frac{i}{\hbar} \left[S(\bar{x}) + \bar{p}(T) \Delta b + \frac{1}{2} \delta^2 S(\rho) + m \sum_j a_j \Delta b \dot{\varphi}_j(T) \right. \right. \\
 & + \sum_j \lambda_j a_j \int_0^T \dot{\varphi}_j(t) \rho(t) dt + \frac{1}{2} \sum_j \lambda_j a_j^2 \\
 & \left. \left. + \sum_{m=3}^{\infty} \frac{\delta^m S(\rho+y)}{m!} \right] \right\} \quad (16.1)
 \end{aligned}$$

This is essentially (15.9); the notation is defined there. For small Δb , specifically for $\Delta b = O(\hbar^{2/3})$ we found the leading asymptotic ($\hbar \rightarrow 0$) contribution to G . In this section we consider what must go into lifting the small Δb restriction.

The overall structure of the argument of the exponent in (16.1) is as follows. It is a function of the a_j 's; in $a_j, j > 1$ it is basically quadratic, while for a_1 it is basically cubic. However, this "basic" character is marred by all sorts of linear and higher power terms. We therefore try to justify a change of variables such that the argument becomes a polynomial of the form

$$z_1^3 + \beta z_1 + \sum_{j=2}^N z_j^2 + \text{const.} \quad (16.2)$$

With these new variables we obtain an asymptotic approximation to the integral which, because we no longer require $\Delta b < \hbar^{2/3}$, is slightly different from that obtained in Section 15.

However, we do not arrive at the normal form (16.2) in one great leap but rather work our way up from simpler integrals in order to see some of the subtleties inherent in (16.1). First we consider problems that arise even for one-dimensional integrals when critical points coalesce; then we go on to two-dimensional integrals and finally discuss multidimensional integrals.

Consider the integral

$$J = \int_{\Gamma} g(w) \exp(-xf(w, \alpha)) dw \quad (16.3)$$

over some contour Γ , such as the real w axis. We wish to find the asymptotic approximation for $x \rightarrow \infty$. The dependence of f on the parameter α is such that as α varies a pair of points at which $\partial f / \partial w = 0$ coalesce. For simplicity we take the value of α at which these critical points coalesce to be $\alpha = 0$. The way (16.3) was handled in our previous work was to expand f in a power series about the location of the coalesced critical points, to observe that the cubic part of the expansion gave the dominant contribution and to discard higher-order terms in w . This can only work for α small on some scale related to x , for example, $\alpha = O(x^{-2/3})$. It is obvious how this restriction enters. Suppose that $f = -\alpha w + w^3/3 + w^4$ and $\Gamma = [0, \infty)$. If $\alpha \ll x^{-2/3}$, then the w^3 term dominates, the effective range of integration for w is $x^{-1/3}$, and the w^4 term is a negligible $O(x^{-4/3})$. On the other hand, for fixed α , with increasing x , the critical point at $w = \sqrt{\alpha}$ will be at values of w greater than $x^{-1/3}$ and the overall size of w will cease going to zero with x . Then the w^4 term cannot be neglected.

The method of Chester, Friedman, and Ursell solves this problem by showing that one can find a new variable z such that

$$f(w, \alpha) = \frac{1}{3}z^3 - \beta z + \gamma \quad (16.4)$$

with β and γ functions of α alone. Equation 16.4 is not a power series expansion. It says that there is some variable z , which stands in one to one correspondence with w , for which f is equal to a cubic polynomial. Call the mapping $z = Z(w)$. Although the proof of the existence of the representation (16.4) involves all sorts of smoothness or analyticity assumptions (we shall not give the proof here) one should bear in mind that the justification for (16.4) is basically topological. That is, the coalescing of critical points with α means that for (say) $\alpha > 0$ f is not monotonic, it has a maximum and a minimum near each other, while for $\alpha < 0$ the merging of the maximum and minimum means that f never backtracks and is monotonic. This

topological essence, the pattern of bumps and their disappearance with changing sign of α , is what the cubic in (16.4) captures.

For the transformation Z to be one to one, both functions in (16.4) must reverse direction at corresponding points; in particular $\partial/\partial z$ of both sides must vanish simultaneously:

$$\frac{\partial f}{\partial w} \frac{dw}{dz} = z^2 - \beta \quad (16.5)$$

(This can also be seen in terms of the implicit function theorem.) Let the points at which f' vanishes be $w_1(\alpha)$ and $w_2(\alpha)$. Then (16.5) demands that $z_1 \equiv Z(w_1)$ and $z_2 \equiv Z(w_2)$ be the roots of the right hand side of (16.5) namely $\pm \sqrt{\beta}$. When inserted in (16.4) this condition becomes

$$f(w_1, \alpha) = -\frac{2}{3}\beta(\alpha)^{3/2} + \gamma(\alpha)$$

$$f(w_2, \alpha) = \frac{2}{3}\beta(\alpha)^{3/2} + \gamma(\alpha) \quad (16.6)$$

Obviously $\beta=0$ when $\alpha=0$. Equation 16.6 fixes β and γ as functions of α . The integral (16.3) is transformed to yield

$$J = \int_{\Gamma'} k(z) \exp \left[-x \left(\frac{1}{3}z^3 - \beta z + \gamma \right) \right] dz \quad (16.7)$$

with Γ' the image of Γ under Z and

$$k(z) = g(w) \frac{dw}{dz} \quad (16.8)$$

At this stage an important point must be made: it will *not* do to expand k as a power series in z and obtain an asymptotic series for J by formally interchanging summation and integration. Such a procedure will not hold uniformly in β since for $\beta > x^{-2/3}$ the effective size of z is $\sqrt{\beta}$, not $x^{-1/3}$, and so successive terms in the formal series will not decrease as $x \rightarrow \infty$. Instead write

$$k(z) = \sum_{m=0}^{\infty} p_m(\alpha)(z^2 - \beta)^m + \sum_{m=0}^{\infty} q_m(\alpha)z(z^2 - \beta)^m \quad (16.9)$$

with the coefficients p_m and q_m determined by standard techniques. In

particular

$$p_0 \pm \sqrt{\beta} q_0 = g(w_i) \sqrt{2\sqrt{\beta} / f''(w_i)} \quad (16.10)$$

with the upper (lower) sign corresponding to $i=1(2)$. Now if integration and summation are interchanged, successive terms in m will decrease for $x \rightarrow \infty$. What is worth noting though is that for the same m , q_m and p_m contributions are of the same order. This is why even to leading order the uniform approximation produces not only the Airy function (multiplying p_0) but its derivative as well (multiplying q_0). To see that indeed the function multiplying q_0 is the derivative of the Airy function recall that from (15.14) $Ai(y) = (1/2\pi i) \int_{\Gamma} \exp(-ty - t^3/3) dt$ (with Γ shown in Fig. 15.1). The derivative with respect to y simply brings down a t and the integrand is $t \exp(ty - t^3/3)$, which is what the integration of the q_0 term in (16.9) yields. Higher terms in m also give the Airy function and its derivative as can be seen by integration by parts. Successive integrations by parts bring down higher powers of $1/x$ providing the asymptotic series. Since we shall only get the leading asymptotic term we shall not go into detail on this point.

The integral J of (16.3) is therefore

$$J = e^{-x\gamma} [p_0 x^{-1/3} Ai(x^{2/3}\beta) + q_0 x^{-2/3} Ai'(x^{2/3}\beta)] [1 + O(x^{-1})] \quad (16.11)$$

with β and γ functions of α defined through (16.6) and p_0 and q_0 given by (16.10). One can check the correctness of the powers of x given in (16.11) (namely, $x^{-1/3}$ and $x^{-2/3}$) by performing a change in variable such as was done in deriving (15.8). The derivative Ai' is taken with respect to its argument. Note that unless $\beta = O(x^{-2/3})$ the p_0 and q_0 terms in (16.11) are of the same asymptotic size in x .

The next step toward our ultimate goal of doing the integral (16.1) is to consider

$$J_2 = \int g(w_1, w_2) \exp[-xf(w_1, w_2; \alpha)] dw_1 dw_2 \quad (16.12)$$

Critical points are defined by $f_1 \equiv \partial f / \partial w_1 = 0$ and $f_2 \equiv \partial f / \partial w_2 = 0$. Again we assume that for $\alpha > 0$ there are two critical points that coalesce for $\alpha \rightarrow 0$, and that for $\alpha < 0$ solutions of $f_1 = 0$ and $f_2 = 0$ do not lie in the range of integration. For $\alpha = 0$ we take f to have one cubic critical point, as in the integral J above, and one uneventful quadratic critical point. Our goal in examining J_2 is to see to what extent having the extra “uneventful” quadratic critical point tagging along complicates the expansion obtained in (16.11).

Before doing anything elegant, let us make a quick examination of J_2 to see that there is nothing really different about it. Consider the equation $f_2(w_1, w_2; \alpha) \equiv 0$. This defines a function $W(w_1, \alpha)$ (through $f_2(w_1, W(w_1, \alpha); \alpha) \equiv 0$) which, unless we have made a poor choice of coordinates, will be well defined for $\alpha > 0$. Now if $f(w_1, w_2; \alpha)$ is expanded about the point $(w_1, W(w_1, \alpha))$, a Gaussian integral in w_2 results, which yields a factor $[\pi/f_{22}(w_1, W(w_1, \alpha); \alpha)]^{1/2}$ in J_2 . What is left in the exponent is an explicit function of w_1 , namely $xf(w_1, W(w_1, \alpha); \alpha)$. As $\alpha \rightarrow 0$ the critical points of this function coalesce—in fact, the critical points of this function are at the same values of w_1 as in the original function. This follows from

$$\frac{\partial}{\partial w_1} f(w_1, W(w_1, \alpha); \alpha) = \frac{\partial f}{\partial w_1} + \frac{\partial f}{\partial w_2} \Big|_{w_2=W} = \frac{\partial f}{\partial w_1}$$

since by the definition of W , $f_2(w_1, W; \alpha) = 0$. Consequently, the integral over w_1 is of the form (16.3), and the Airy function and its derivative emerge. It is clear that for larger numbers of variables w_1, \dots, w_N this method for evaluating (16.12) will get messy.

A more straightforward approach to (16.12) would be to expand f in a power series about the unique (\bar{w}_1, \bar{w}_2) , the single root of $f_1 = f_2 = 0$ for $\alpha = 0$. There the matrix $\partial^2 f / \partial w_i \partial w_j$ would have a zero eigenvalue, and by appropriate rotation of coordinates we would be back to the situation of Section 15, namely

$$\begin{aligned} f(w_1, w_2; \alpha) = & f + \frac{1}{2} \frac{\partial^2 f}{\partial w_2^2} (w_2 - \bar{w}_2)^2 + \frac{1}{6} \frac{\partial^3 f}{\partial w_1^3} (w_1 - \bar{w}_1)^3 \\ & + \alpha \frac{\partial f}{\partial \alpha} + \alpha(w_1 - \bar{w}_1) \frac{\partial^2 f}{\partial \alpha \partial w_1} + \alpha(w_2 - \bar{w}_2) \frac{\partial^2 f}{\partial \alpha \partial w_2} \\ & + O(\alpha^2) + O((w_2 - \bar{w}_2)^3) + O((w_1 - \bar{w}_1)^4) + \dots \end{aligned} \quad (16.13)$$

[all f 's on the right hand side are evaluated at $(\bar{w}_1, \bar{w}_2; \alpha)$; we have assumed that the coordinates have already been chosen so that the eigenvector of $\partial^2 f / \partial w_i \partial w_j$ with zero eigenvalue is along the $(w_1 - \bar{w}_1)$ direction]. But we know that this procedure gives a nonuniform approximation, and a different strategy is needed.

From the form of (16.3) and the strategy of (16.4) it is clear that what we seek are variables z_1 and z_2 such that

$$f(w_1, w_2; \alpha) = \frac{1}{3} z_1^3 - \beta(\alpha) z_1 + \frac{1}{2} z_2^2 + \gamma(\alpha) \quad (16.14)$$

exactly, and not as the first terms in a power series expansion. Topologically, we expect the method to work since with changing α (and β , which is a function of α) both left and right hand sides of (16.14) have coalescing critical points. Again the size of x [the asymptotic variable in (16.12)] has nothing to do with the validity of (16.14), and the range of α for which the transformation exists is set only by the global behavior of f .

Once one has the canonical form on the right side of (16.14) the integration reduces to the previous case. The product of the function $g(w_1, w_2)$ with the Jacobian $\partial(w_1, w_2)/\partial(z_1, z_2)$ introduces some function of z_1 and z_2 into the nondangerous part of the integral (i.e., the part unaffected by $x \rightarrow \infty$). The leading term in the z_2 integral is obtained just from the Gaussian while the z_1 integration yields two terms, as in (16.11).

We postpone justifying the existence of the transformation (16.14) until we discuss the general N variable case.

One further integral is examined to elucidate another possible complication. Consider

$$K = \int g(w_1, w_2) \exp[-xf(w_1, w_2; \alpha)] dw_1 dw_2 \quad (16.15)$$

The quantity α now has several components (three will turn out to be enough for present purposes), and as $|\alpha| \rightarrow 0$ as many as four critical points may coalesce. Obviously the representation (16.14) will not do, being a bit short on critical points. Saying that not more than four critical points coalesce does not quite characterize the singularity, and we further assume that $\partial^2 f / \partial w_i \partial w_j$ has in fact two zero eigenvalues (and is therefore zero) for $\alpha=0$. (The other way of getting four critical points is typified by $f=w_2^2 + w_1^5 + \alpha_3 w_1^3 + \alpha_2 w_1^2 + \alpha_1 w_1$). The canonical form for f for the case of both directions being involved is

$$f(w_1, w_2; \alpha) = \frac{1}{3}z_1^3 + \frac{1}{3}z_2^3 - \beta_1(\alpha)z_1 - \beta_2(\alpha)z_2 - \beta_{12}(\alpha)z_1 z_2 + \gamma(\alpha) \quad (16.16)$$

Again this is an equality that holds for some functions $z_1 = Z_1(w_1, w_2)$ and $z_2 = Z_2(w_1, w_2)$ and with β and γ definite functions of α . And again to justify the existence of the transformations we appeal to topological arguments. The pattern of maxima and minima on both sides of (16.16) must be the same for all values of α in some neighborhood of $\alpha=0$. Hence as the β 's are varied all patterns available through the variation of α must be obtained (and vice versa). The need for β_1 and β_2 is obvious from our previous treatment of the single variable case. That β_{12} is also needed can

be seen from Fig. 16.1 where contour maps of

$$P(z_1, z_2; \beta) = \frac{1}{3}z_1^3 + \frac{1}{3}z_2^3 - \beta_1 z_1 - \beta_2 z_2 - \beta_{12} z_1 z_2 \quad (16.17)$$

have been plotted for the cases (a) $\beta_1 > 0, \beta_2 > 0, \beta_{12} = 0$, (b) $\beta_2 > 0, \beta_1 = \beta_{12} = 0$, and (c) $\beta_1 = \beta_2 = 0, \beta_{12} \neq 0$. Notice that with β_1 and β_2 alone there is no way of generating a contour map with the topology of that shown in Figure 16.1c [Case (c) can be obtained from (a) by allowing the minimum and two saddles to merge into a single saddle, but without β_{12} there is no way to do this.] Of course the display of Fig. 16.1 hardly constitutes a proof that the representation (16.16) exists and is unique, but here too we do not take up the issue until it arises for the N variable case.

The integral K therefore has become

$$K = \int k(z_1, z_2) \exp[-xP(z_1, z_2; \beta)] dz_1 dz_2 \quad (16.18)$$

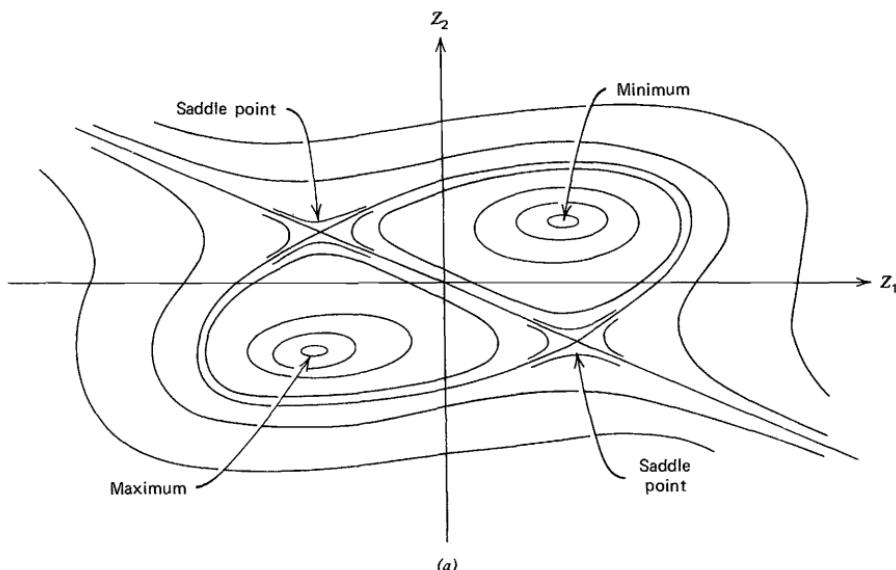
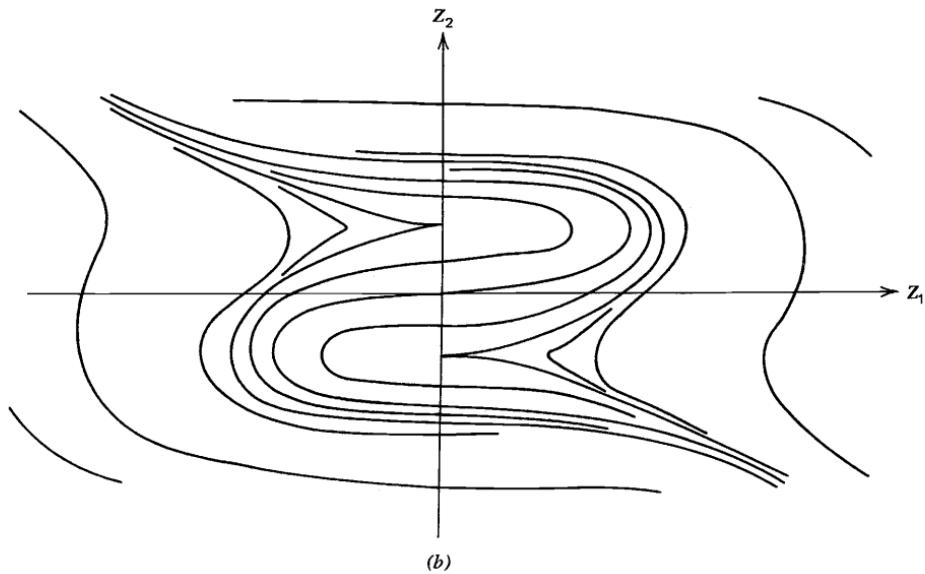
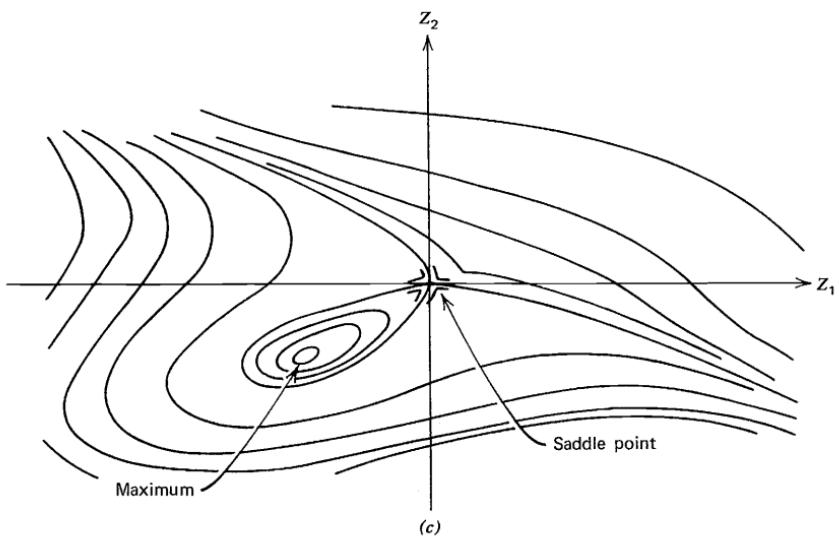


Fig. 16.1 Contour maps of the function $P(z_1, z_2; \beta) = \frac{1}{3}z_1^3 + \frac{1}{3}z_2^3 - \beta_1 z_1 - \beta_2 z_2 - \beta_{12} z_1 z_2$ appearing in (16.17). z_1 and z_2 are not drawn to the same scale. (a) $\beta_1 = \beta_2 > 0, \beta_{12} = 0$. The saddle points occur on the contour $P=0$. (b) $\beta_1 = \beta_{12} = 0, \beta_2 > 0$. This map has no maxima or minima and its contour lines neither close or cross. (c) $\beta_1 = \beta_2 = 0, \beta_{12} > 0$. The saddle point occurs on the curve $P=0$.



(b)



(c)

with $k = g\partial(w_1, w_2)/\partial(z_1, z_2)$. The function k is to be expanded, but as found above for J , in order to get an asymptotic expansion it will not do simply to take a power series in z_1 and z_2 . Furthermore, just as for the Airy function, where both the Airy function and its derivative were needed for the leading terms in the asymptotic expansion, more than one function will be needed for the leading contribution to K . Specifically all the following functions will appear:

$$\begin{aligned} U(\beta) &= \int \exp[-xP(z_1, z_2; \beta)] dz_1 dz_2 \\ \frac{1}{x} \frac{\partial U(\beta)}{\partial \beta_i} &= \int z_i \exp[-xP(z_1, z_2; \beta)] dz_1 dz_2 \quad i = 1, 2 \\ \frac{1}{x} \frac{\partial U(\beta)}{\partial \beta_{12}} &= \int z_1 z_2 \exp[-xP(z_1, z_2; \beta)] dz_1 dz_2 \end{aligned} \quad (16.19)$$

The coefficients of these functions in the expansion of K can be found by essentially the same method as used for (16.10) to (16.11).

Now we go to the general case of N variables. Referring to (16.1), we have a function $W(a_1, \dots, a_N; \Delta b)$ (the argument of the exponent) whose power series for $\Delta b = 0$ starts as $a_1^3 + \sum_{j=2}^N a_j^2$. For $\Delta b \neq 0$ it picks up additional terms in these variables. Our handling of the integrals J , J_2 , and K suggests that we seek variables z_1, \dots, z_N such that W can be equated exactly to a polynomial that looks like the first terms in its power series expansion. Thus we want a representation of the form

$$W(a_1, \dots, a_n; \Delta b) = z_1^3 - \beta(\Delta b)z_1 + \sum_{j=2}^N z_j^2 + \gamma(\Delta b) \quad (16.20)$$

with β and γ definite functions of Δb .

Here is where catastrophe theory has done the work. A significant part of that theory concerns itself with taking a function having some sort of singularity and bringing it to a canonical form. The tools are various preparation theorems. We state one of these, the easiest one in that it covers analytic functions, just to give some of the flavor:

WEIERSTRASS PREPARATION THEOREM: Let $F: U \rightarrow \mathbb{C}$ be analytic where U is a neighborhood of the origin in $\mathbb{C} \times \mathbb{C}^n$. Suppose that F satisfies $F(w, 0) = w^k g(w)$ where $(w, 0) \in \mathbb{C} \times \mathbb{C}^n$, g is an analytic function of a single complex variable w in some neighborhood of the origin, $g(0) \neq 0$, and k is a positive integer. Then there exists an analytic function $h: U' \rightarrow \mathbb{C}$ with U' a neighborhood of 0 in $\mathbb{C} \times \mathbb{C}^n$ and analytic functions $\gamma_0, \dots, \gamma_{k-1}: V \rightarrow \mathbb{C}$ with

V a neighborhood of 0 in \mathbb{C}^n such that

$$(hF)(w, z) = w^k + \sum_{j=0}^{k-1} \gamma_j(z) w^j$$

for all (w, z) in some neighborhood of 0 in $\mathbb{C} \times \mathbb{C}^n$ and $h(0) \neq 0$.

Of course at the bottom of the proof of this theorem lies the implicit function theorem, but quite a bit of effort must go into lining things up so that that powerful theorem can be called into play.

Corresponding to the Weierstrass theorem for analytic functions there is the Malgrange preparation theorem for smooth (C^∞) functions. Then there are refinements, generalizations, extensions, and whatnot associated with the names of various workers—Mather, Nirenberg, Thom, and many others—until one gets the normal form given in (16.20). To be precise one does not usually phrase the results the way I have done in (16.20); rather what one reads about are germs, jets, bundles, and the unfolding of singularities. But I do not think I will be misleading the reader if I state that the existence of the coordinate change implicit in (16.20) ultimately follows from the work I have cited.

The form (16.20) is just one of many possibilities, and here the famous seven elementary catastrophes come to tell us all possible forms that can be taken by a caustic. The shape of the caustic is governed by the singular part of W (i.e. the nonquadratic part) where the singular part can be identified by the simultaneous vanishing of first and second derivatives. Besides the cubic in (16.20) we give examples of other sorts of singularity:

$$z^n, \quad n \geq 3; \quad z_1^n + z_2^m, \quad n, m \geq 3; \quad \exp\left(-\frac{1}{z^2}\right) \quad (16.21)$$

For each singularity there is an “unfolding.” In plain (but undoubtedly obscure) English the unfolding consists of all lower-order terms needed to obtain all topologically distinct functions, just as the coefficient β_1 , β_2 , and β_{12} were needed for (16.16). (By “topological equivalence” we mean: let $f, g: M \rightarrow N$ be smooth mappings from a manifold M to a manifold N . Then f is equivalent to g if there exist homeomorphisms $h: M \rightarrow M$ and $k: N \rightarrow N$ such that $f \circ h = k \circ g$ where \circ is composition). The occurrence of such a singularity is called a catastrophe, and the dimension of the catastrophe is the number of parameters needed in the universal unfolding. As we saw above, these parameters could be mapped onto spatial coordinates and the coordinate space form of the caustic reflected a canonical morphology in an abstract parameter space of the unfolding. Consequently any phenomenon in space time which arises from a catastrophe in its underlying

dynamics will involve a catastrophe whose unfolding has four or fewer parameters (exceptions to this may arise from symmetries; we consider only the generic case). Thom has found all possible catastrophes of four or fewer dimensions. They are seven in number: x^3 (fold), x^4 (cusp), x^5 (swallow tail), x^6 (butterfly), x^3+y^3 (hyperbolic umbilic), x^3+3xy^2 (elliptic umbilic), x^2y+y^4 (parabolic umbilic). Catastrophes involving four parameters can be seen in three-dimensional moving pictures of caustics.

From the standpoint of the method of stationary phase, each catastrophe has its characteristic special function (as the Airy function is associated with the fold), and in the asymptotic expansion each parameter in the unfolding requires a term corresponding to the derivative of the characteristic function with respect to that parameter.

The main physical applications of the foregoing theory are to scattering theory and of course to optics. In a later section we give a path integral formulation of scattering theory and show how the semiclassical approximation (i.e., replacing the path integral by the classical path term alone with appropriate path density) is particularly useful for scattering of some heavy complex objects. The approximation is most useful when the dominant motion is along the classical trajectory of the heavy objects but quantum effects of the internal degrees of freedom still play an important role. Such situations arise in atomic, molecular, and nuclear physics. It also turns out that at some energies and scattering angles focusing effects, that is, caustics, come into play (by the generic property). Then one needs the machinery developed in this section. We refer the reader to the papers given below for specific applications.

NOTES

The uniform asymptotic expansion for a pair of coalescing critical points was given by

C. Chester, B. Friedman, and F. Ursell, An Extension of the Method of Steepest Descents, *Proc. Camb. Phil. Soc. (Ser. Physica)* **53**, 599 (1957)

Further developments with an eye toward atomic and molecular scattering theory applications are to be found in

J. R. Stine and R. A. Marcus, Semiclassical Transition Probabilities by an Asymptotic Evaluation of the S Matrix for Elastic and Inelastic Collisions. Bessel Uniform Approximation, *J. Chem. Phys.* **59**, 5145 (1973)

J. N. L. Connor, Asymptotic Evaluation of Multidimensional Integrals for the S Matrix in the Semiclassical Theory of Inelastic and Reactive Molecular Collisions, *Mol. Phys.* **25**, 181 (1973)

J. N. L. Connor, Semiclassical Theory of Molecular Collisions: Three Nearly Coincident Classical Trajectories, *Mol. Phys.* **26**, 1217 (1973)

J. N. L. Connor, Evaluation of Multidimensional Canonical Integrals in Semiclassical Collision Theory, *Mol. Phys.* **26**, 1371 (1973)

J. N. L. Connor, Multidimensional Canonical Integrals for the Asymptotic Evaluation of the S-Matrix in Semiclassical Collision Theory, *Faraday Discussions of the Chemical Society* **55**, 51 (1973)

J. N. L. Connor, Catastrophes and Molecular Collisions, *Mol. Phys.* **31**, 33 (1976)

W. H. Miller, Semiclassical Theory of Atom-Diatom Collisions: Path Integrals and the Classical S-Matrix, *J. Chem. Phys.* **53**, 1949 (1970)

W. H. Miller, Classical S-Matrix: Numerical Application to Inelastic Collisions, *J. Chem. Phys.* **53**, 3578 (1970)

W. H. Miller and T. F. George, Analytic Continuation of Classical Mechanics for Classically Forbidden Collision Processes, *J. Chem. Phys.* **56**, 5668 (1972)

P. Pechukas and M. S. Child, *Mol. Phys.* **31**, 973 (1976)

P. Pechukas, *Phys. Rev.* **181**, 166 (1969)

P. Pechukas, *Phys. Rev.* **181**, 174 (1969)

A. P. Penner and R. Wallace, *Phys. Rev. A* **7**, 1007 (1973)

A. P. Penner and R. Wallace, *Phys. Rev. A* **9**, 1136 (1974)

A. P. Penner and R. Wallace, *Phys. Rev. A* **11**, 149 (1975)

For nuclear physics the uniform expansion, starting from path integration, is given by

S. Levit, U. Smilansky, and D. Pelte, *Phys. Lett.* **53B**, 39 (1974)

S. Levit, Semiclassical Approximation to Path Integrals—Phases and Catastrophes, in G. J. Papadopolis and J. T. Devreese, eds., *Path Integrals and Their Applications in Quantum, Statistical and Solid State Physics* (Proc. NATO Adv. Study Ins., Antwerp 1977), Plenum, New York, 1978 (and earlier references therein)

A general study of oscillatory integrals using all the machinery of catastrophe theory (machinery employed by us in this section with clashing gears) is that of

J. J. Duistermaat, *Commun. Pure and Appl. Math.* **27**, 207 (1974)

The need for uniform expansions is mentioned in

M. V. Berry and K. E. Mount, Semiclassical Approximations in Wave Mechanics, *Rep. Prog. Phys.* **35**, 315 (1972), p. 384

and catastrophe theory's relation to the semiclassical approximation is discussed in

M. V. Berry, *Adv. Phys.* **25**, 1 (1976)

SEVENTEEN

The Phase of the Semiclassical Amplitude

The semiclassical amplitude is of the form $\sqrt{\partial^2 S_c / \partial a \partial b} \exp(iS_c(b, t; a)/\hbar)$ with a and b initial and final points. Since $(\partial^2 S_c / \partial a \partial b)^{-1}$ vanishes at focal points the sign of the square root can be ambiguous. In this section we show how to obtain that sign (hence the phase of the amplitude) in two ways: by Morse theory and by a straightforward differential equations method. This phase turns out to be important for deriving the correct form of the Bohr-Sommerfeld quantization rule; in particular it supplies the $\frac{1}{2}$ in the relation $\int p dq = (n + \frac{1}{2})\hbar$.

The Morse theory result is implicit in our work of Sections 12 to 14. The term $\partial^2 S_c / \partial a \partial b$ (or its many-dimensional generalization as a determinant) arises as the product of eigenvalues of $\delta^2 S$. If all these eigenvalues are positive then the overall phase of the coefficient in the propagator (i.e., that which multiplies $\exp(iS_c/\hbar)$) is zero; this is because the original normalization of the path integral defined as a limit has just enough $\sqrt{-i}$'s [cf. (1.24)] to cancel all those arising from Gaussian integrals over the normal modes of $\delta^2 S$. Each eigenvalue of $\delta^2 S$ that is negative replaces an integral $\int da \exp(i|\lambda|a^2)$ by an integral $\int da \exp(-i|\lambda|a^2)$ and therefore in the coefficient the factor $\sqrt{1/-i}$ becomes $\sqrt{1/i}$. This gives a phase $\exp(i\pi/2)$ for each negative eigenvalue of $\delta^2 S$. But we know from Morse theory that the number of negative eigenvalues of $\delta^2 S$ (evaluated about some classical trajectory) is just the number of focal points along that trajectory, each focal point counted with its multiplicity. To summarize, the contribution to the semiclassical propagator from each classical trajec-

tory is

$$(-2\pi i\hbar)^{-d/2} \exp(-i\pi n/2) \left| \sqrt{\det(\partial^2 S_c / \partial a \partial b)} \right| \exp[iS_c(b, t; a)/\hbar]$$

with n the sum of the multiplicities of the focal points along the trajectory and d the dimension of the coordinate space.

Next we supply a simple differential equations proof of this result. Since our discussion in this section is within the context of the semiclassical approximation, we can restrict attention to quadratic Lagrangians of the form $L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}W(t)x^2$. The Green's function for such a Lagrangian satisfies

$$\begin{aligned} G(x, t; 0, 0) &= \exp(iS_c/\hbar) \int_{(0, 0)}^{(0, t)} dy(\tau) \exp\left\{ \frac{i}{2\hbar} \int_0^t (m\dot{y}^2 - Wy^2) d\tau \right\} \\ &= \exp\left(\frac{iS_c}{\hbar} \right) G(0, t; 0, 0) \end{aligned} \quad (17.1)$$

with S_c the action along the classical path from $(0, 0)$ to (x, t) . Let f satisfy the differential equation

$$m \frac{d^2 f}{dt^2} + W(t)f = 0 \quad (17.2)$$

with boundary conditions

$$f(0) = 0, \quad \dot{f}(0) = 1$$

By linearity the classical path from $(0, 0)$ to (x, t) is

$$x(\tau) = x \frac{f(\tau)}{f(t)} \quad (17.3)$$

and the action is easily calculated to be

$$S_c = \frac{1}{2}mx^2 \frac{\dot{f}(t)}{f(t)} \quad (17.4)$$

Define

$$g(t) = G(0, t; 0, 0)$$

We know from previous work that g is essentially $1/\sqrt{f}$, but we derive

this again, with special attention given to the phase. First we use the fact that for $t > 0$ $G(x, t; 0, 0)$ satisfies Schrödinger's equation

$$\begin{aligned} 0 &= \left(H - i\hbar \frac{\partial}{\partial t} \right) G(x, t; 0, 0) \\ &= \left\{ -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} W(t)x^2 - i\hbar \frac{\partial}{\partial t} \right\} \left[g(t) \exp\left(\frac{imx^2 \dot{f}}{2\hbar f}\right) \right] \\ &= \left[-\frac{i}{2} \frac{\dot{f}(t)}{f(t)} - i \frac{\dot{g}(t)}{g(t)} \right] \hbar G(x, t; 0, 0) \end{aligned} \quad (17.5)$$

Thus $d(\log g)/dt = -(1/2)d(\log f)/dt$, and we have

$$g(t) = \frac{c(t)}{|f(t)|^{1/2}} \quad (17.6)$$

To get (17.6) we integrated the differential equation for t such that $f(t) \neq 0$. The quantity $c(t)$ is constant for $f(t) \neq 0$ but could change as t passes through zeros of $f(t)$. We next show how this change occurs.

Let t_0 be a (simple) zero of $f(t)$. Then from

$$G(0, t_0 + \varepsilon; 0, 0) = \int dx G(0, t_0 + \varepsilon; x, t_0 - \delta) G(x, t_0 - \delta; 0, 0) \quad (17.7)$$

we have for small ε and δ

$$\frac{c(t_0^+)}{|f(t_0 + \varepsilon)|^{1/2}} = \int dx G(0, t_0 + \varepsilon; x, t_0 - \delta) \frac{c(t_0^-)}{|f(t_0 - \delta)|^{1/2}} \exp\left[\frac{imx^2 \dot{f}(t_0 - \delta)}{2\hbar f(t_0 - \delta)}\right] \quad (17.8)$$

As usual the short time Green's function is given by

$$G(0, t_0 + \varepsilon; x, t_0 - \delta) = \sqrt{\frac{m}{2\pi i\hbar(\varepsilon + \delta)}} \exp\left[\frac{imx^2}{2\hbar(\varepsilon + \delta)}\right] \quad (17.9)$$

The Gaussian integration yields

$$\frac{c(t_0^+)}{|f(t_0 + \varepsilon)|^{1/2}} = \frac{c(t_0^-)}{|f(t_0 - \delta)|^{1/2}} \left[1 + (\varepsilon + \delta) \frac{\dot{f}(t_0 - \delta)}{f(t_0 - \delta)} \right]^{-1/2} \quad (17.10)$$

which, with the assumption that f has a simple root, yields

$$c(\epsilon_0^+) = c(\epsilon_0^-) e^{-i\pi/2} \quad (17.11)$$

where the validity of the limit does not depend on the relative sizes of ϵ and δ .

Exercise: If f has a double root our demonstration appears to give nonsense. Why? How could you provide an analogous demonstration to support the assertion made above that for higher-order focal points the phase $\pi/2$ is added according to the multiplicity of the focal point?

What I think is most entertaining about this demonstration is that it is in some sense a differential equations proof of Morse Theory.

NOTES

The idea of following the propagator through the focal point by means of (17.7) appears in

P. Pechukas, *Phys. Rev.* **181**, 166 (1969)

Phase changes arising from passage through focal points have long been known in optics. See

M. Born and E. Wolf, *Principles of Optics*, Pergamon, New York, 1959.

EIGHTEEN

The Semiclassical Propagator as a Function of Energy

The approximation that we have been calling WKB has borne little resemblance to the form of that approximation most commonly encountered. To recover the familiar form we look at the propagator as a function of energy, as described in Section 7:

$$\tilde{G}(x, y; E) = \int_0^\infty dt \exp\left(\frac{iEt}{\hbar}\right) G(x, t; y, 0) \quad (18.1)$$

where $E \sim E + i\varepsilon$, $\varepsilon \downarrow 0$, as is usual. The WKB approximation for $G(x, t; y, 0)$ is

$$\begin{aligned} G_{\text{WKB}}(x, t; y, 0) &= \sum_\alpha \sqrt{\det\left(\frac{i}{2\pi\hbar} \frac{\partial^2 S_\alpha}{\partial x \partial y}\right)} e^{iS_\alpha/\hbar} \\ &= \sum_\alpha \left(\frac{i}{2\pi\hbar}\right)^{d/2} \sqrt{\left|\det \frac{\partial^2 S_\alpha}{\partial x \partial y}\right|} e^{i[(S_\alpha/\hbar) - n\pi/2]} \end{aligned}$$

where α labels the classical paths from $(y, 0)$ to (x, t) , d is the dimension of the space, and n is the number of negative eigenvalues of $\delta^2 S$ along the path α (or alternatively the sum of the orders of the focal points along α). We insert G_{WKB} in (18.1) and as usual do a stationary phase approximation to get the leading term as $\hbar \rightarrow 0$:

$$\frac{\partial}{\partial t} (Et + S_\alpha(x, y; t)) = 0 \quad (18.2)$$

For a given E and given α this picks (in general, many) values of t (labeled by β) satisfying

$$E = - \frac{\partial S_\alpha(x, y; t_{\alpha\beta})}{\partial t} \quad (18.3)$$

For each $t_{\alpha\beta}$ a Gaussian approximation to (18.1) gives the leading term:

$$\begin{aligned} \tilde{G}_{\text{WKB}}(x, y; E) &= \left(\frac{i}{2\pi\hbar} \right)^{d/2} \sum_{\alpha, \beta} \sqrt{|\det \partial^2 S / \partial x \partial y|} \\ &\times \exp \left[- \frac{in\pi}{2} + \frac{iS_{\alpha\beta}}{\hbar} + \frac{iEt_{\alpha\beta}}{\hbar} \right] \int_0^\infty dt e^{i(\partial^2 S / \partial t^2)(t - t_{\alpha\beta})^2 / 2\hbar} \end{aligned} \quad (18.4)$$

(suppressing some $\alpha\beta$ subscripts). The integral leaves us with an expression

$$\frac{\det(\partial^2 S / \partial x \partial y)}{\partial^2 S / \partial t^2} \quad (18.5)$$

whose square root is a factor in the propagator (phases will be handled separately). Going from t as independent variable to E is, by (18.3), a Legendre transform, and we define

$$W(x, y, E) = S + Et \quad (18.6)$$

The expression (18.5) will next be rewritten in terms of W . Using the relations

$$\frac{\partial W(x, y, E)}{\partial E} = t, \quad \frac{\partial S(x, y; t)}{\partial u} = \frac{\partial W(x, y, E)}{\partial u} \quad (18.7)$$

(u =any component of x or y ; other arguments of each function fixed) we find

$$\frac{\partial^2 S}{\partial u \partial v} \Big|_t = \frac{\partial^2 W}{\partial u \partial v} \Big|_E + \frac{\partial^2 W}{\partial u \partial E} \frac{\partial^2 W}{\partial v \partial E} \frac{\partial^2 S}{\partial t^2} \quad (18.8)$$

where u and v are any components of x and y respectively. Using (18.7) and (18.3) it follows that

$$\frac{\partial^2 S}{\partial t^2} = - \left(\frac{\partial^2 W}{\partial E^2} \right)^{-1} \quad (18.9)$$

Combining the last several equations, it is convenient to consider an enlarged matrix and the expression (18.5) becomes

$$-\frac{\det(\partial^2 S / \partial x \partial y)}{\partial^2 S / \partial t^2} = \det \begin{bmatrix} \frac{\partial^2 W}{\partial x \partial y} & \frac{\partial^2 W}{\partial x \partial E} \\ \frac{\partial^2 W}{\partial y \partial E} & \frac{\partial^2 W}{\partial E^2} \end{bmatrix} \equiv \tilde{D} \quad (18.10)$$

The problem of phases, solved earlier for $G(x, t; y, 0)$, must now be reconsidered since $\partial^2 S / \partial t^2$ has its own zeros, infinities, and sign changes. As we shall next show, $\partial^2 S / \partial t^2$ is infinite when $\det \partial^2 S / \partial x \partial y$ is, so that the focal points that gave phase factors in the t representation are no longer important. The infinity arises because $\partial^2 S / \partial t^2$ is $-\partial E(x, y, t) / \partial t$. Considering instead $\partial t / \partial E$ we shall see how a small change in energy leads to a higher order ($O[(\Delta E)^2]$) change in time at a focal point. Consider therefore the boundary value problem with the same initial and final points, y and x , but with slightly different energy. The variation is about a specific path x_α . That y and x are conjugate means that there is a one-parameter family of paths, with various initial momenta, all of which end at (x, t) (to lowest order). Thus varying E will not change the time to lowest order. Hence $\partial t / \partial E$ is zero.

The foregoing argument is heuristic, but we can appeal to (18.10) for ultimate justification. The infinities of that determinant in general are different from those of $\det \partial^2 S / \partial x \partial y$ and we now seek to determine them. From the derivation of (18.10) we know that those infinities will be at the zeros of $\partial^2 S / \partial t^2$ which are the same as the infinities of $\partial t / \partial E$. For fixed y , we claim that those infinities occur when x is at the turning points of the classical motion. At the turning points the system spends relatively longer periods of time, and for fixed x (and of course y) changing E will mean that x is no longer a turning point. More precisely, for motion in a potential V

$$t = \int_y^x \frac{\sqrt{m} du}{\sqrt{2(E - V(u))}} \quad (18.11)$$

Even for x at a turning point ($V(x) = E$) this integral is finite. But the derivative with respect to E is not. (For trajectories that have gone through a turning point the derivative is again finite, but some care is needed in its evaluation.)

The phases for the WKB Green's function as a function of energy are therefore determined by the number of times the determinant on the right

hand side of (18.10) passes through zero, and these passages represent turning points of the classical motion. Strictly we should check the first assertion either by counting negative eigenvalues of the appropriate infinite dimensional quadratic form (for the variational principle with fixed energy) or by checking the Schrödinger equation near the turning point. Having done both these checks for $G_{\text{WKB}}(x, t; y, 0)$ we shall do neither for its Fourier transform. The final expression is

$$\tilde{G}_{\text{WKB}}(x, y; E) = \left(\frac{i}{2\pi\hbar}\right)^{d/2 - 1/2} \sum_{\alpha, \beta} \sqrt{|\tilde{D}|} e^{-in\pi/2 + iW/\hbar} \quad (18.12)$$

with \tilde{D} given in (18.10) above and with each W and associated quantities understood to have indices α and β . Here n , also dependent on α and β , is the number of turning points.

Having struggled to get (18.12) let us recover some familiar formulas. In one space dimension, let the shape of the potential and the value of the energy be such that the particle is constrained to a finite region and bounces back and forth between its turning points $x_L(E)$ and $x_R(E)$. See Fig. 18.1. Take the arguments x and y of $\tilde{G}_{\text{WKB}}(x, y; E)$ to be within the classically accessible region ($x_L \leq x, y \leq x_R$), and for convenience take $y \leq x$. Before summing in (18.12), we enumerate all classical paths from y to x having energy E . These fall into four classes depending on whether

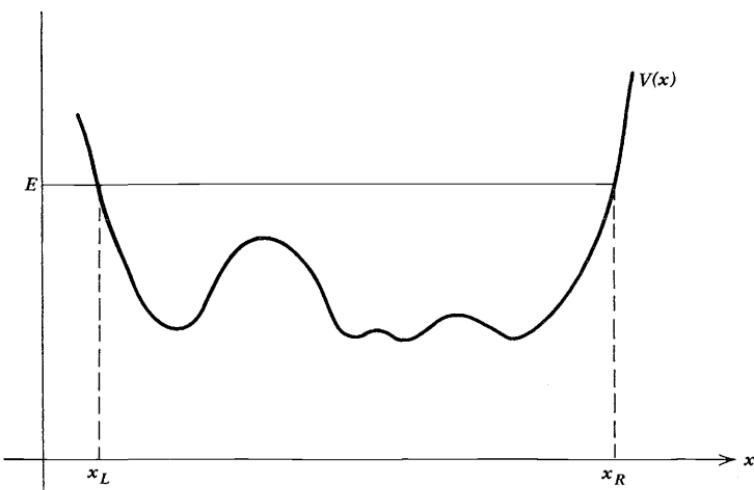


Fig. 18.1 Bounded periodic motion in a potential. For energy E the turning points are x_L and x_R .

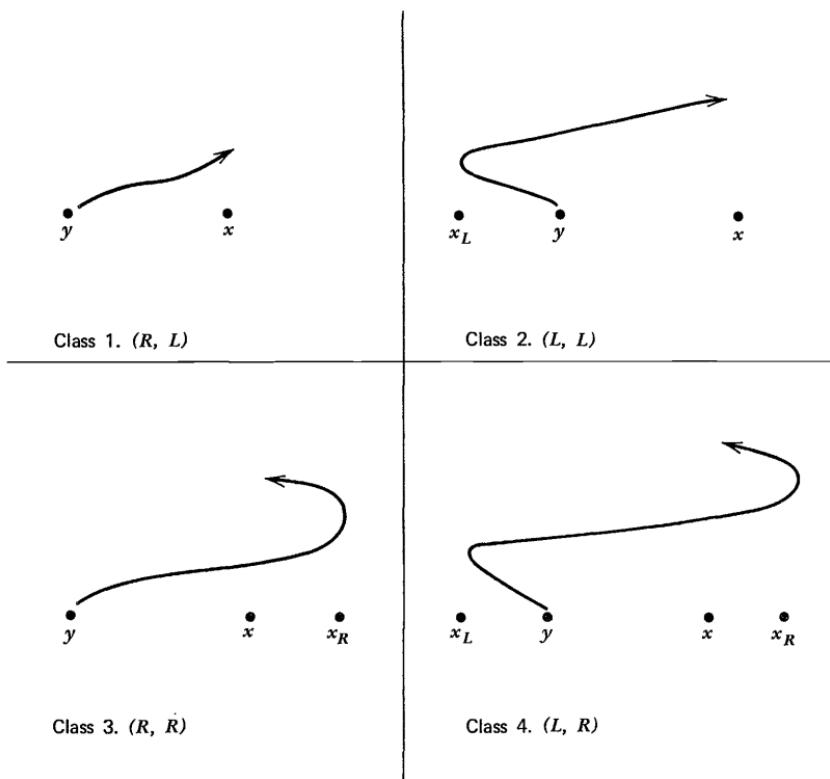


Fig. 18.2 Paths from y to x . The shortest member of each class is illustrated for classes 1–4. The time coordinate is vertical.

they leave y going left or right and arrive at x from the left or right. The shortest member of each class is illustrated in Fig. 18.2. For each path illustrated there is an infinity of paths of the same energy E in which the particle goes through an integer number of periods, as in Fig. 18.3. To compute the action W along each path no detailed knowledge of the path is needed. Since $S(x, y; t)$ satisfies the Hamilton-Jacobi equation, W satisfies

$$E = H\left(\frac{\partial W}{\partial x}, x\right) = \frac{1}{2m} \left(\frac{\partial W}{\partial x}\right)^2 + V(x) \quad (18.13)$$

Defining

$$w(x, E) = \int_{x_L}^x \sqrt{2m(E - V(u))} du \quad (18.14)$$

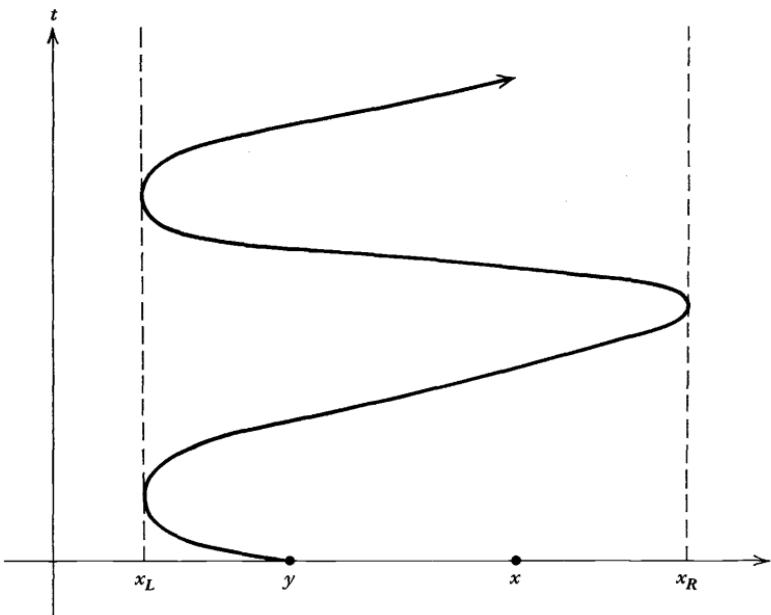


Fig. 18.3 A path in class 2 of Fig. 18.2.

it is clear that $W(x, y, E)$ must have the form

$$W(x, y, E) = \pm w(x, E) \pm w(y, E) + \text{const.} \quad (18.15)$$

with the constant some integer multiple of

$$J = \oint p dx = 2w(x_R, E) = 2 \int_{x_L}^{x_R} \sqrt{2m(E - V)} dx \quad (18.16)$$

Sorting out the signs in (18.15) involves looking in detail at the various classes of paths, and it is easy to see that

$$\begin{aligned} W_1(n) &= w(x, E) - w(y, E) + nJ \\ W_2(n) &= w(x, E) + w(y, E) + nJ \\ W_3(n) &= -w(x, E) - w(y, E) + (n+1)J \\ W_4(n) &= -w(x, E) + w(y, E) + (n+1)J \end{aligned} \quad (18.17)$$

with $n = 0, 1, \dots$ being the number of additional full cycles, besides those illustrated in Fig. 18.2. The number of turning points along each of these

paths is $2n + n_i$ where $n_1 = 0$, $n_2 = 1$, $n_3 = 1$, and $n_4 = 2$. In doing the sum in (18.12) there is a determinant to evaluate. For this determinant the element $\partial^2 W / \partial E^2$ may always be ignored (even for dimension greater than 1), since its cofactor $\partial^2 W / \partial x \partial y$ has determinant zero. To see that $\partial^2 W / \partial x \partial y$ is a singular matrix recall that W satisfies

$$H\left(\frac{\partial W(x, y, E)}{\partial x}, x\right) = E \quad (18.18)$$

Apply to this $\partial / \partial y$ to get

$$\sum_i \frac{\partial H}{\partial p_i} \frac{\partial^2 W}{\partial x_j \partial y_i} = 0 \quad (18.19)$$

(indices i, j for the components of the vectors x, y). The quantity $\partial H / \partial p_i$, being the velocity, is not in general zero. Returning to our one-dimensional case the determinant is simply the product of $\partial^2 W / \partial x \partial E$ and $\partial^2 W / \partial y \partial E$. It is the absolute value of the determinant that enters, and this is the same for all classes. Putting together all the foregoing results

$$\begin{aligned} \tilde{G}_{\text{WKB}}(x, y; E) &= \sqrt{\frac{m/2}{\sqrt{(E - V(x))(E - V(y))}}} \\ &\times \sum_{n=0}^{\infty} \exp\left(\frac{inJ}{\hbar} - in\pi\right) \sum_{i=1}^4 \exp\left[\frac{iW_i(0)}{\hbar} - \frac{in_i\pi}{2}\right] \end{aligned} \quad (18.20)$$

The sum over n is evaluated to yield

$$\begin{aligned} \tilde{G}_{\text{WKB}}(x, y; E) &= \frac{(m/2)^{1/2} (E - V(x))^{-1/4} (E - V(y))^{-1/4}}{1 - \exp[(i/\hbar)(J - \hbar\pi)]} \\ &\times \sum_{j=1}^4 \exp\left[\frac{i}{\hbar} \left(W_j(0) - \frac{n_j\hbar\pi}{2}\right)\right] \end{aligned} \quad (18.21)$$

The propagator $\tilde{G}(x, y; E)$ can be expressed in terms of its singularities and near a value of E in the point spectrum of H (a bound state) has a pole with residue related to the bound state wave function, specifically

$$\tilde{G}(x, y; E) = \frac{i\hbar}{E - E_n} \psi_n(x) \psi_n^*(y) + \text{nonsingular function} \quad (18.22)$$

The singularities of (18.21) occur when

$$1 = \exp\left[\frac{i}{\hbar}(J - \hbar\pi)\right] \quad (18.23)$$

that is, for

$$J = 2\hbar\pi\left(k + \frac{1}{2}\right) \quad k = 0, 1, \dots \quad (18.24)$$

($k \geq 0$ since we have defined J as intrinsically positive).

Equation 18.24 is a familiar semiclassical result: the Bohr-Sommerfeld quantization rules. But the Green's function contains more information, and we examine the residue at the pole which is

$$\sqrt{\frac{m}{2}} \left(\frac{i\hbar}{\tau}\right) \frac{2 \sin\left[\frac{w(x, E)}{\hbar} + \frac{\pi}{4}\right]}{(E - V(x))^{1/4}} \cdot \frac{2 \sin\left[\frac{w(y, E)}{\hbar} + \frac{\pi}{4}\right]}{(E - V(y))^{1/4}} \quad (18.25)$$

where

$$\tau = 2 \int_{x_L}^{x_R} dx \sqrt{\frac{m}{2(E - V(x))}} = \text{time for complete cycle} \quad (18.26)$$

Equation 18.25 should be compared with the wave function in the semiclassical approximation (see, for example, Landau and Lifshitz, *Quantum Mechanics*, eq. 48.3). In a word, the residue is, as it should be, the product of the wave functions.

A variant of the foregoing theme involves not only the passage from t to E but also the use of momentum coordinates p instead of x . This is closely related to "phase space path integrals," but the corresponding variational principles (and Morse theory) for the classical mechanics are not so well developed. Gutzwiller has found the remarkable fact that the approximate WKB wave functions using p and E are in fact exact for the Coulomb potential. It is not clear whether this result is related to the exactness of the sum over classical paths for Lie groups (to be discussed later) inasmuch as the Coulomb problem (with given energy) can be recast as a trivial dynamics on the group manifold of $O(4)$. This entire subject is one that possesses many remarkable properties generally characterized as the exactness of the correspondence principle of quantum mechanics. This topic has also been explored by Norcliffe et al.

In physical applications it is of course necessary to consider potentials that are not perfectly spherically symmetric, and this opens a whole series of problems in both the description of the classical paths and in their summation. Gutzwiller has discussed the spectrum of the Hamiltonian using path integrals in this more difficult situation, and the reader is referred to his papers for further elaboration.

So far we have calculated the wave function in the “allowed” region—that is, we assumed that for a given E there were solutions to (18.2) and for the given endpoints there were one or more classical paths.

Barrier penetration is a quantum phenomenon in which the wave function does not vanish despite the classical inaccessibility of some spatial region (for given energy). As such it is analogous to optical diffraction phenomena. There is a WKB description of barrier penetration. What can path integration say?

For given energy the path integral takes the form

$$\tilde{G}(b, a; E) = \int_0^\infty dt \int_{\substack{x(t)=b \\ x(0)=a}} dx(\cdot) \exp\left\{\frac{i}{\hbar} (Et + S[x(\cdot)])\right\} \quad (18.27)$$

Barrier penetration or tunneling occurs when there is no simultaneous (real) solution to the equations

$$\delta S = 0 \quad (\text{Euler-Lagrange equations}) \quad (18.28a)$$

$$E + \frac{\partial S}{\partial t} = 0 \quad (18.28b)$$

The most elegant state of affairs would obtain if one could use complex valued functions $x(t)$ satisfying (18.28) as the critical points in a *path integral over complex paths* and thereby get a full analogue for the path integral of the stationary phase approximation. No one seems to know how to do this. What can be done is to approximate the time dependent propagator using the WKB expressions which involves $S(b, a; t)$, the action evaluated along the classical path. Then one is only lacking a solution to (18.28b) and by going to complex time the integral $\int_0^\infty dt$ can be shifted so as to pick up the stationary point. This is what McLaughlin did, and he recovered many WKB results in this way. I believe this approach is bound to succeed in the sense of recovering the WKB results for the following reason: analytic continuation of $S(b, a; t)$ is related to analytic continuation of the Hamilton-Jacobi equation of which it is a solution. But the Hamilton-Jacobi equation is what one obtains by formal manipulation of the Schrödinger equation in seeking the largest contribution for $\hbar \rightarrow 0$.

Consequently, although having started from the path integral, once one has fixed upon its WKB approximation all analytic continuations leave one with solutions of the Hamilton-Jacobi equation, hence with the leading \hbar contribution to the Schrödinger equation. (See also page 224.)

For a completely different view of tunneling, in which bona fide paths are found as solutions to a continued classical mechanics, see Section 29. There the potential V becomes $-V$ as a result of $t \rightarrow -it$, the latter transformation being the result of looking at the Feynman-Kac formula.

Exercise: Use the methods of this section to get the Green's function, eigenvalues, and eigenvectors for a particle in an "infinite square well" of length a . Find the time dependent propagator.

Remarks on the Solution: Because the potential cannot be approximated by a line of finite slope at the turning points [a feature influencing the analytic behavior of the integral (18.11)] the phase change in the propagator on going through a turning point is not $\pi/2$. Rather, the phase change is π , since this is what is needed to make the Green's function, hence all the wave functions, vanish at the walls. See (6.43) and the subsequent discussion for the case of a single hard wall. Equation 18.20 now becomes

$$\tilde{G} \sim \sum_{n=0}^{\infty} \exp\left(\frac{inJ}{\hbar} - 2in\pi\right) \sum_{k=1}^4 \exp\left(\frac{iW_k(0)}{\hbar} - in_k\pi\right)$$

The $2ni\pi$ can be dropped. The numbers n_k (defined above) are just right to make \tilde{G} vanish for x or y approaching the walls. Since V is constant for $0 \leq x \leq a$, $J = 2a\sqrt{2mE}$ and the sum over n yields a pole for $J/\hbar = 2\nu\pi$, $\nu = 1, 2, \dots$. Hence

$$E = \frac{1}{2m} \left(\frac{\hbar\nu\pi}{a} \right)^2, \quad \nu = 1, 2, \dots$$

are the quantized energy levels. $W_k(0)$, $k = 1, \dots, 4$ are computed to give the usual square well wave functions.

Finally it is easy to check that the time dependent propagator is just a Jacobi theta function, and when the fundamental identity for theta functions is implemented [(23.12)] a simple expression for $G(x, t; y)$ emerges in terms of classical paths that bounce various numbers of times off the walls. This can also be stated in terms of the method of images (see also Exercise 3, Section 7, page 46).

NOTES

Many of the calculations presented here can be found in

M. C. Gutzwiller, Phase-Integral Approximation in Momentum Space and the Bound States of an Atom, *J. Math. Phys.* **8**, 1979 (1967)

M. C. Gutzwiller, Phase-Integral Approximation in Momentum Space and the Bound States of an Atom II, *J. Math. Phys.* **10**, 1004 (1969)

M. C. Gutzwiller, Energy Spectrum According to Classical Mechanics, *J. Math. Phys.* **11**, 1791 (1970)

M. C. Gutzwiller, Periodic Orbits and Classical Quantization, *J. Math. Phys.* **12**, 343 (1971)

These papers go well beyond what I have presented in this section. Exactness of the WKB approximation for the Coulomb potential and the consequences of this property has been studied under the designation "Correspondence Identities" by A. Norcliffe, I. C. Percival, and M. J. Roberts, and is reviewed in

A. Norcliffe, Correspondence Identities and the Coulomb Potential, *Case Studies in Atomic Physics* **4**, 1 (1973), North Holland

Special properties of the Coulomb potential are given in

M. Bander and C. Itzykson, Group Theory and the Hydrogen Atom. I, *Rev. Mod. Phys.* **38**, 330 (1966); II, *Rev. Mod. Phys.* **38**, 346 (1966)

Barrier penetration and path integrals are discussed in

D. W. McLaughlin, Complex Time, Contour Independent Path Integrals and Barrier Penetration, *J. Math. Phys.* **13**, 1099 (1972)

McLaughlin works out a number of explicit examples, and his method rests in particular on analytic continuation in time. The idea of a complex path and in general the combining of the seemingly contradictory ideas of diffraction and ray optics is due to

J. B. Keller, A Geometrical Theory of Diffraction, in *Calculus of Variations and its Applications*, L. M. Graves, Ed., McGraw-Hill, New York, 1958

J. B. Keller, Geometrical Theory of Diffraction, *J. Opt. Soc. Amer.* **52**, 116 (1962)

NINETEEN

Scattering Theory

An area in which the semiclassical approximation to the path integral has proved particularly useful is scattering theory. The first formulations of scattering theory in path integral language were in the context of atomic scattering, but the large energies (and quantum numbers) involved in heavy ion scattering have led to applications to nuclear physics too.

The S -matrix is

$$\langle p''|S|p'\rangle = \lim_{\substack{t' \rightarrow -\infty \\ t'' \rightarrow +\infty}} \left\langle p'' \left| \exp\left(\frac{iH_0 t''}{\hbar}\right) \exp\left[\frac{-iH(t'' - t')}{\hbar}\right] \exp\left(\frac{-iH_0 t'}{\hbar}\right) \right| p' \right\rangle \quad (19.1)$$

where H_0 is the free Hamiltonian and H includes interactions. For simplicity we assume the particle to be spinless and interacting with a potential, although this is certainly not the case of major physical interest. The operator H_0 in (19.1) is trivial and gives respectively E'' and E' for the momentum states p'' and p' . What is left to evaluate after this is simply the Fourier transform of the usual path integral, namely

$$\begin{aligned} \langle p''|S|p'\rangle &= \lim \exp\left[\frac{i}{\hbar}(E''t'' - E't')\right] \int dx' dx'' \frac{1}{2\pi} \exp\left[\frac{i}{\hbar}(p'x' - p''x'')\right] \\ &\quad \times \left\langle x'' \left| \exp\left[-\frac{i}{\hbar}H(t'' - t')\right] \right| x' \right\rangle \end{aligned} \quad (19.2)$$

The last matrix element is expressed as a path integral,

$$G \equiv \left\langle x'' \left| \exp\left[-\frac{i}{\hbar}H(t'' - t')\right] \right| x' \right\rangle = \int dx(\tau) \exp\left[\frac{i}{\hbar} \int_{t'}^{t''} \left(\frac{m\dot{x}^2}{2} - V(x)\right) d\tau\right] \quad (19.3)$$

and we make the usual semiclassical approximation

$$G = \sum_{\alpha} \sqrt{\det\left(\frac{i}{2\pi\hbar} \frac{\partial^2 S_{\alpha}}{\partial x_i'' \partial x_j'}\right)} e^{iS[x_{\alpha}]/\hbar} \quad (19.4)$$

where each x_{α} is a *classical path* (solution of the Euler-Lagrange equations) starting at (x', t') and ending at (x'', t'') . $S(x_{\alpha})$ is the action along that classical path. The integrals over x' and x'' are also handled by the stationary phase approximation and, for example, the x' integral takes its major contribution from that x' for which

$$\frac{\partial}{\partial x'} [p' x' + S_{\alpha}] = 0$$

This of course is one of the properties of the classical path:

$$p' = -\frac{\partial S}{\partial x'} \quad (19.5)$$

Equation 19.5 is to be satisfied simultaneously with

$$p'' = \frac{\partial S}{\partial x''} \quad (19.6)$$

yielding a pair \bar{x}' and \bar{x}'' , the classical endpoints for the given boundary condition in p' and p'' . To order \hbar we have merely performed a canonical transformation. The square root in (19.4) is slowly varying and can be removed from the integral yielding

$$\begin{aligned} \langle p'' | S | p' \rangle &= \lim_{t', t''} \exp\left[\frac{i}{\hbar}(E''t'' - E't')\right] \sum_{\alpha} \sqrt{\det\left(\frac{i}{2\pi\hbar} \frac{\partial^2 S_{\alpha}}{\partial x'' \partial x'}\right)} \\ &\times \exp\left[\frac{i}{\hbar}(S_{\alpha} + p'\bar{x}' - p''\bar{x}'')\right] \int d(x'' - \bar{x}'') d(x' - \bar{x}') \\ &\times \exp\left\{\frac{i}{\hbar}\left[\frac{\partial^2 S}{\partial x'^2}(x' - \bar{x}')^2 + \dots\right]\right\} \end{aligned} \quad (19.7)$$

The Gaussian integral brings factors into the square root that change it to a second derivative of the action with respect to momenta. The action S_{α} can be written

$$S_{\alpha} = \int (p dx - H dt)$$

For the classical path H is a constant, E_α , and using the integration by parts formula, $\int p dx = px|_{x'}^{x''} - \int x dp$, S becomes

$$\begin{aligned}\langle p'' | S | p' \rangle &= \lim_{t'', t'} \sum_{\alpha} \exp \left\{ \frac{i}{\hbar} [(E'' - E_\alpha) t'' - (E' - E_\alpha) t'] \right\} \\ &\quad \times \sqrt{\det \left(\frac{\partial^2 F_\alpha(p'', p')}{\partial p'' \partial p'} \right)} e^{(i/\hbar) F_\alpha(p'', p')} \end{aligned} \quad (19.8)$$

where

$$F_\alpha(p'', p') = - \int x dp \quad (19.9)$$

the integral taken along the classical path. It is assumed here—as is customary in scattering theory—that wave packets have been formed, making things converge when necessary. Sandwiching S between narrow wave packets leads to $E_\alpha = E' = E''$. Ignoring the δ -function that also arises from this we write S as

$$\langle p'' | S | p' \rangle = \sum_{\alpha} \sqrt{\det \left(\frac{\partial^2 F_\alpha}{\partial p' \partial p''} \right)} e^{i F_\alpha / \hbar} \quad (19.10)$$

We work the details of the formalism above only in the simplest of cases, but it is worth pointing out the extensive applications available for (19.10).

Consider the scattering of a pair of atoms. The energies involved often put the relative motion of the atoms well into the semiclassical domain. One can use an effective potential derived by looking at the details of the more rapid (and not semiclassical) electronic motion. For nuclear ion-ion scattering one can go even further. There are collective modes of the individual nuclei that are well described by harmonic oscillator coordinates. These coordinates can be included in the classical action so that these classical like degrees of freedom of the nucleus can also be handled on a first principle basis (rather than by the Born-Oppenheimer type approximation implicitly suggested above).

While these may seem matters for the specialists (see references given at the end of this section) their practical implementation calls forth some rather attractive ideas, in particular the way in which analytic continuation plays a role in the tunneling phenomenon.

The example that we now give to illustrate the formalism does not begin to cover the applications; in particular, it misses the refinements

needed for handling inelastic scattering and the complex potentials and paths invoked there. Nevertheless, it may at least shed some light on classical tunneling.

Consider a one-dimensional square well potential

$$V(x) = U\theta\left(\frac{a}{2} - |x|\right) \quad (19.11)$$

A particle of momentum $\hbar k$ is incident from the left. In classical mechanics if $V < \hbar^2 k^2 / 2m$, the particle goes right on through, except that while in the well it has momentum $\hbar K$, with

$$k^2 = K^2 + \frac{2mU}{\hbar^2} \equiv K^2 + \lambda \quad (19.12)$$

which defines λ . To evaluate the integral in (19.9) consider the function

$$p(t) = \hbar k + \hbar(K - k)\theta(t - T_1)\theta(T_2 - t) \quad (19.13)$$

where T_1 is the time at which the particle enters the well and T_2 the time at which it leaves. The function $p(t)$ is the classical motion of the momentum coordinate. Since dp/dt vanishes except for $t = T_1$ and T_2 the only information needed about x is

$$\begin{aligned} x(T_1) &= -\frac{a}{2} \\ x(T_2) &= \frac{a}{2} \end{aligned}$$

The derivative of p is a pair of δ -functions in t and clearly

$$\begin{aligned} \frac{1}{\hbar} F &= \frac{-1}{\hbar} \int x \frac{dp}{dt} dt = (-1) \left[\left(\frac{-a}{2} \right) (K - k) + \frac{a}{2} (K - k)(-1) \right] \\ &= a(K - k) = a \left[\sqrt{k^2 - \lambda} - k \right] \end{aligned} \quad (19.14)$$

By (19.10) this is the scattering phase shift. This can be compared to the exact phase shift δ , which is (see any text in quantum mechanics) the solution of the equation

$$k \tan\left(\frac{Ka}{2}\right) = k \tan\left(\frac{ka}{2} - \delta\right) \quad (19.15)$$

For both ka and λ/k^2 small there is agreement between (19.15) and

(19.14). But for other parameter values the results can disagree—even to the point of qualitative differences. Consider the attractive δ -function potential limit, for which $a \rightarrow 0$, $U \rightarrow -\infty$, $aU \rightarrow \text{const}$. This gives a finite phase shift quantum mechanically, but in the semiclassical approximation this potential becomes completely transparent.

Another consequence of (19.14) is the derivation of tunneling or barrier penetration effects. Let U become large and positive. Then F/\hbar develops an imaginary part, and we choose the branch cuts so that

$$\operatorname{Re}\left(\frac{iF}{\hbar}\right) = -a\sqrt{\lambda - k^2} \quad (19.16)$$

This is precisely the WKB approximation for barrier penetration. The way we got (19.16) was by analytic continuation, and there was a choice of variables suitable for this continuation: λ or k^2 . By developing the semiclassical approximation in other ways one could continue in t or other variables. In effect we have been doing a stationary phase approximation, and for barrier penetration the critical points of the integrand lie outside the domain of integration. The continuation is equivalent to deforming the contour of integration to include the distant critical point. The most intriguing ideas in this direction suggest that it is a complex trajectory $x(\tau)$ that is wanted in (19.3) and that the domain for the path integral has been deformed to include complex paths. As remarked earlier in this book, as far as I know, the implementation of these ideas remains an open problem.

NOTES

Derivation of the S -matrix [given in (19.1)] can be found in

R. G. Newton, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York, 1966, Sec. 6.3

The phase shift for scattering by a square well potential is given in

R. H. Dicke and J. P. Wittke, *Introduction to Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1960

Some references on scattering theory and path integrals are:

P. Pechukas, Time Dependent Semiclassical Scattering Theory, I, Potential Scattering, *Phys. Rev.* **181**, 166 (1969); II, Atomic Collisions, *Phys. Rev.* **181**, 174 (1969)

D. Gelman and L. Spruch, Feynman Path Integrals and Scattering Theory, *J. Math. Phys.* **10**, 2240 (1969)

A. P. Penner and R. Wallace, Semiclassical Collision Theory within the Feynman Path-Integral Formalism, *Phys. Rev. A* **7**, 1007 (1973); Application to Vibrational Excitation, *Phys. Rev. A* **9**, 1136 (1974)

W. H. Miller, Semiclassical Theory of Atom-Diatom Collisions: Path Integrals and the Classical *S*-Matrix, *J. Chem. Phys.* **53**, 1949 (1970); Classical *S*-Matrix: Numerical Application to Inelastic Collisions, *J. Chem. Phys.* **53** 3578 (1970); *Adv. Chem. Phys.* **25**, 69 (1974)

W. H. Miller and T. F. George, Analytic Continuation of Classical Mechanics for Classically Forbidden Processes, *J. Chem. Phys.* **56**, 5668 (1972)

Miller and George make the point that one-dimensional examples of barrier penetration that analytically continue in a variable other than the coordinate (such as we did in λ or k^2) may be unrealistic in that in higher dimension a single complexified parameter may not be enough to find a way around the barrier.

W. B. Campbell, P. Finkler, C. E. Jones, and M. N. Misheloff, Path-Integral Formulation of Scattering Theory, *Phys. Rev. D* **12**, 2363 (1975)

T. Koeling and R. A. Malfliet, Semi-classical Approximations to Heavy Ion Scattering Based on the Feynman Path-Integral Method, *Physics Reports* **22**, 181 (1975)

This is a useful review of applications to nuclear physics and examines in detail the effectiveness of various approximations. The authors are particularly concerned with scattering by complex potentials and with the sorts of "classical paths" needed to provide accurate approximations.

S. Levit, U. Smilansky, and D. Pelte, A New Semi-Classical Theory for Multiple Coulomb Excitation, *Phys. Lett.* **53B**, 39 (1974)

S. Levit and U. Smilansky, A New Approach to Gaussian Path Integrals and the Evaluation of the Semi-Classical Propagator, *Ann. Phys.* **103**, 198 (1977)

These authors also have the nuclear physics applications in mind and study caustics and other focusing phenomena along lines similar to our presentation in earlier sections. Rainbow and glory phenomena are also taken up by Pechukas (see above).

Barrier penetration and complex potentials were studied by D. W. McLaughlin in his 1970 Indiana University thesis and published by him:

D. W. McLaughlin, *J. Math. Phys.* **13**, 1099 (1972)

He does the inverted parabolic barrier (as do Miller and George, see above) and several other examples. Complex paths in a more traditional WKB setting—not using the path integral—are discussed by

J. Knoll and R. Schaeffer, Semiclassical Scattering Theory with Complex Trajectories, I, Elastic Waves, *Ann. Phys.* **97**, 307 (1976)

Another paper along these lines is that of

R. Balian and C. Bloch, Solution of the Schrödinger Equation in Terms of Classical Paths, *Ann. Phys.* **85**, 514 (1974)

The use of ray tracing for diffraction (analogous to the use of classical mechanics for tunneling) can be found in

J. B. Keller, Geometrical Theory of Diffraction, *J. Opt. Soc.* **52**, 116 (1962)

TWENTY

Geometrical Optics

Mathematically speaking, electromagnetism and the nonrelativistic quantum mechanics of a spinless particle differ in two important ways. First, the electromagnetic field is a *vector* field. Second, it satisfies a partial differential equation of second order in time, in fact hyperbolic, while quantum mechanics is first order in time and parabolic. As far as path integrals are concerned the vector character of the electromagnetic field is indeed a drawback and the would be path integrator of the electromagnetic field is faced by the same challenges discussed later in this book in connection with the path integration of quantum mechanical particles with spin. For this reason, in our treatment of the electromagnetic field we treat it as a scalar quantity—a good deal of useful optics can be understood with this simplified treatment. (Note that we are here discussing a path integral for the electromagnetic wave equation—not the second quantization of the electromagnetic field, which is another matter altogether.)

A way around the fact that the electromagnetic field satisfies a wave equation which is second order in time can be found through a formal trick reminiscent of the way Feynman quantized the Klein-Gordon equation (Section 25).

Consider, then, free electromagnetic radiation moving through a medium with a spatially slowly varying index of refraction $n(x)$. Then the “scalar electromagnetic field” $u(x, t)$ satisfies the equation

$$\frac{n(x)^2}{c^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = 0 \quad (20.1)$$

Consider only waves of a definite frequency ω , so that

$$u(x, t) = u(x) \exp(-i\omega t) \quad (20.2)$$

One must be especially careful in defining what parameter will become

large in the forthcoming geometrical optics approximation. In fact characteristic distances are large compared to the wavelength of light. Hence when x is measured in units of light wavelengths it will be large, and we are motivated to define

$$x = By \frac{c}{\omega} \quad (20.3)$$

where y is a new variable with characteristic size unity while $B \rightarrow \infty$ will be the geometrical optics limit. Combining the foregoing equations gives

$$-\frac{\omega^2}{c^2} [n(x)]^2 u - \frac{\omega^2}{c^2} \frac{1}{B^2} \nabla_y^2 u = 0 \quad (20.4)$$

with $\nabla_y^2 = \partial^2/\partial y_1^2 + \partial^2/\partial y_2^2 + \partial^2/\partial y_3^2$. For expositional convenience we assume we are in a situation where $n(x)^2 \rightarrow E$ as $x \rightarrow \infty$, with E some constant ($E=1$ for vacuum). Define

$$V(y) = E - n(x)^2 \quad (20.5)$$

Equation 20.4 becomes

$$-\frac{1}{B^2} \nabla_y^2 u + V(y)u = Eu \quad (20.6)$$

which, one must admit, is beginning to look like Schrödinger's equation. The next transformation has a very artificial look to it, but it's what it takes to get the job done, and, if one bears in mind Feynman's way of dealing with Klein-Gordon particles, is not all that unexpected. Write

$$\psi(y, \tau) = e^{-iE\tau B} u(y) \quad (20.7)$$

The function ψ clearly satisfies

$$-\frac{1}{B^2} \nabla_y^2 \psi + V\psi = \frac{i}{B} \frac{\partial \psi}{\partial \tau} \quad (20.8)$$

so that the parallel to the time dependent Schrödinger equation is manifest. Moreover, the limits $\hbar \rightarrow 0$ and $B \rightarrow \infty$ are formally identical. On the other hand, to recover a solution of (20.6) from a solution of (20.8) one integrates

$$u(y) = \int_{-\infty}^{\infty} d\tau \exp(iE\tau B) \psi(y, \tau) \quad (20.9)$$

At this stage one can write down a path integral solution for ψ , but some conceptual clarity is gained by backtracking to (20.6) and considering the Green's function for that equation, that is, the solution to

$$\left[-\frac{1}{B^2} \nabla_b^2 + V(b) - E \right] \tilde{G}(b, a; E) = -\frac{i}{B} \delta(b-a) \quad (20.10)$$

in precise analogy to (7.8). As in Section 7, $\text{Im } E > 0$ will provide a convergence factor whenever necessary. $\tilde{G}(b, a; E)$ represents a solution to the optics problem in that it gives the illumination at b for a point source at a . It can also be used for the “scattering problem,” that is, to find the scattered wave for a given incoming wave, just as in the quantum mechanical Lippmann-Schwinger equation.

Now the way is clear. The time independent Green's function $\tilde{G}(b, a; E)$ is obtained from the time dependent one exactly as in Section 7, namely by integration, but with range of integration $0 \leq \tau \leq \infty$ rather than $-\infty$ to ∞ as in (20.9). The time dependent Green's function is given by the path integral, and we have finally

$$\tilde{G}(b, a; E) = \int_0^\infty d\tau e^{iE\tau B} \int_{(a, 0)}^{(b, \tau)} dy(\cdot) \exp(iBS[y(\cdot)]) \quad (20.11)$$

with

$$S[y(\cdot)] = \int_0^\tau \left[\frac{1}{4} \left(\frac{dy}{d\tau'} \right)^2 - V(y) \right] d\tau' \quad (20.12)$$

where the notation $dy(\cdot)$ has been used for “volume in path space” to help us slide the more easily through our hand waving arguments. The notation indicates that only paths satisfying $y(0)=a$, $y(\tau)=b$ are to be summed over. Of course $1/\hbar$ has everywhere been replaced by B .

For the multiple integral (20.11) the large value of B indicates the use of the stationary phase approximation. The integrals are over τ and $y(\cdot)$, and the argument of the exponent is

$$iB \left[E\tau + \int_0^\tau \left[\frac{1}{4} \left(\frac{dy}{d\tau'} \right)^2 - V(y) \right] d\tau' \right] \quad (20.13)$$

This quantity is to be stationary with respect to variation of both τ and $y(\cdot)$. Variation with respect to $y(\cdot)$ gives the Euler-Lagrange equations and a “classical” path $\bar{y}(\tau)$ in the unphysical variable τ , while variation of

(20.13) with respect to τ gives

$$E + \frac{\partial S}{\partial \tau} = E - \left[\frac{1}{4} \left(\frac{dy}{d\tau} \right)^2 + V \right] = 0 \quad (20.14)$$

which says that the “energy” constant along the path \bar{y} is just E .*

To formulate a variational principle for $y(\tau)$ without the unphysical variable τ , we make use of (20.14) to rewrite (20.13) as

$$E\tau + \int_0^\tau \left[\frac{1}{4} \left(\frac{dy}{d\tau'} \right)^2 - V \right] d\tau' = \int_0^\tau \frac{1}{2} \left(\frac{dy}{d\tau'} \right)^2 d\tau' \equiv A \quad (20.15)$$

It now follows that among paths satisfying (for all τ')

$$\frac{1}{4} \left(\frac{dy}{d\tau'} \right)^2 + V = E \quad (20.16)$$

which begin at a and end at b , allowing the final time τ to vary also, the classical path $\bar{y}(\cdot)$ is that which makes the quantity

$$A = \int_0^\tau \frac{1}{2} \left(\frac{dy}{d\tau'} \right)^2 d\tau' \quad (20.17)$$

stationary. The freedom to vary τ comes about because with (20.16) given as a constraint, (20.14) becomes the condition that fixes τ .

But the variational principle just formulated is that of Maupertuis. To get rid of the τ in the formulation of the problem we observe that the arc length ds' along $y(s')$ is given by

$$ds' = \sqrt{\left(\frac{dy_1}{d\tau'} \right)^2 + \left(\frac{dy_2}{d\tau'} \right)^2 + \left(\frac{dy_3}{d\tau'} \right)^2} d\tau' \quad (20.18)$$

which is of course independent of τ . Moreover, (20.16) can be rewritten as

$$\left(\frac{dy}{d\tau'} \right)^2 = 4(E - V) \quad (20.19)$$

*Notice that $\partial S / \partial \tau$ is $-H = -[\frac{1}{4}(dy/d\tau)^2 + V]$ and not L as one might have thought. This is because in changing τ one also changes the classical path along which the integral in (20.13) is evaluated.

Substituting,

$$A = \int_0^s \sqrt{E - V} \, ds' \quad (20.20)$$

where s in the integration limit is the length of the path in physical space. We now recall the definition of V , (20.5), so that $(E - V)^{1/2}$ is just $n(x)$, which is proportional to $1/v(x)$ the local light velocity in the medium.

We can now formulate our conclusion as follows: The path $x(s')$ (parametrized by arc length s') which gives the largest contribution to the illumination at a point b from a point source at a is that which makes stationary the integral

$$T = \int_0^s \frac{1}{v(x)} \, ds' \quad (20.21)$$

among paths $x(s')$ beginning at a and ending at b , but with total arc lengths not specified.

This is Fermat's principle! T is just the travel time for the light ray, and our demand is for the path of minimum (stationary, actually) time.

Several comments are in order. First, the stationary path alone does not really give the "largest contribution," but the paths near it do. I allowed myself some license to make the statement of the conclusion more dramatic. Second, the Gaussian integral around the stationary path can be done, just as for path integrals, and gives the ray density—certainly an important quantity in applications. Third, the entire range of interference and diffraction phenomena is easily handled with this formalism. In case there are two or more stationary solutions there will be interference. When solutions are "close" in function space focusing and caustics exist, exactly as was treated in other sections of this book. Finally it may happen that no solutions exist but that by analytic continuation complex times or complex paths may be found that are stationary. Such "complex paths" suggest a relation to geometric diffraction theory, and were it not for the formal nature of our stationary phase approximation for the path integral the way would be open to a good deal of useful mathematics.

NOTES

The derivation given in this section, or some variant thereof, must surely have been known in the early days of path integrals. So although I doubt that what has been presented here is original, I do not know of any source

for it. One paper based on a similar approach is that of

M. Eve, The Use of Path Integrals in Guided Wave Theory, *Proc. Roy. Soc. Lond. A* **347**, 405 (1976)

Geometrical optics of course has its own extensive literature. Ultimately its mathematical justification depends on the stationary phase approximation. I am unable to provide a complete or balanced guide to the literature, but the following papers will at least give the interested reader a handle on further sources:

C. S. Morawetz and D. Ludwig, An Inequality for the Reduced Wave Operator and the Justification of Geometrical Optics, *Comm. Pure and Appl. Math.* **21**, 187 (1968)

D. Ludwig, Uniform Asymptotic Expansions at a Caustic, *Comm. Pure and Appl. Math.* **19**, 215 (1966)

R. A. Handelsman and N. Bleistein, Asymptotic Expansions of Integral Transforms with Oscillatory Kernels: A Generalization of the Method Stationary Phase, *SIAM J. Math. Anal.* **4**, 519 (1973)

“Geometrical diffraction”, a seeming contradiction in terms, is described in

J. B. Keller, Geometrical Theory of Diffraction, *J. Opt. Soc.* **52**, 116 (1962)

Keller has found a way to use ray optics methods in the analysis of diffraction, and it has been suggested that further understanding of his method could come from the sort of path integral approach described in this section.

TWENTY-ONE

The Polaron

An electron moving in a polar crystal distorts the ions around it. This both lowers the energy of the electron and increases its effective mass. A "distortion" of a crystal lattice corresponds to the creation of phonons, and we treat the problem as that of an electron in interaction with the phonon field. The physically important contributions come from the high frequency, "optical" modes, and it is both convenient and apparently not a bad approximation to assume that all these modes have the same frequency.

Historically, the use of path integration in this problem was an early and powerful application by Feynman of his new method. The application is significant in several ways. First, the problem is of interest in its own rights in solid state physics. Second, it is one of the simplest examples of a field interacting with a particle. The particle becomes "dressed" in the finest style of quantum field theory. All the mass renormalization of quantum electrodynamics, which has been known to discourage people with delicate mathematical sensitivities, is here accomplished without mysterious infinities. The electron does indeed travel in a cloud of field excitations (phonons), but both its bare mass and renormalized mass are perfectly defined. The third reason for the special significance of the polaron problem is that it illustrates one of the most effective uses of path integration. Namely, there are two systems, say X and Y , in interaction; using path integrals one of them may be solved exactly and leaves the problem entirely; the propagator then involves only the other with some self interaction. Thus the path integral has the form

$$G = \int dX(\cdot) dY(\cdot) \exp \left\{ - \int L_X - \int L_Y - \int L_I \right\}$$

with L_X and L_Y the free Lagrangians for X and Y respectively and L_I the interaction. Suppose that $\int dY(\cdot) \exp \{- \int L_Y - \int L_I\}$ can be carried out exactly to give some $g(X) \equiv \exp(-V(X))$. Then G is only a path integration over X involving a new self interaction term $V(X)$. This method works

best when Y is a harmonic oscillator, but that includes, besides the example treated in this section, the electron-photon interaction. There the induced self interaction of the electron is just the Lagrangian for action at a distance electrodynamics.

Consider, then, an electron in a continuous polarizable solid. The electron coordinate and momentum are r and p and its mass m , really its "effective mass" in the static lattice potential, is taken to be 1. Our units are also tailored to having the assumed single frequency ω of the optical phonons to be 1 as well as $\hbar = 1$. The important physical parameter is the coupling constant α which measures the extent to which the electron is capable of distorting the lattice. In terms of the usual dimensional quantities α is given by

$$\alpha = \frac{1}{2} \left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon} \right) \frac{e^2}{\hbar \omega} \left(\frac{2m\omega}{\hbar} \right)^{1/2}$$

where ϵ and ϵ_∞ are the static and high frequency dielectric constants. Typical values of α are between 1 and 20. Let the quantity k be the phonon wave number, and phonon coordinates p_k and q_k are defined in the usual way in terms of creation and annihilation operators of the second quantized field of vibrations of the solid. In terms of all these quantities the Hamiltonian is

$$H = \frac{1}{2} p^2 + \frac{1}{2} \sum_k (p_k^2 + q_k^2) + (2\sqrt{2} \pi \alpha)^{1/2} \sum_k \frac{1}{k} q_k e^{ik \cdot r} \quad (21.1)$$

where by \sum_k we mean $\int d^3 k / (2\pi)^3$. A Lagrangian giving rise to (21.1) is

$$L = \frac{1}{2} \dot{r}^2 + \frac{1}{2} \sum_k (\dot{q}_k^2 - q_k^2) - (2\sqrt{2} \pi \alpha)^{1/2} \sum_k \frac{1}{k} q_k e^{ik \cdot r} \quad (21.2)$$

Our goal is to estimate the ground state energy of this system using the Feynman-Kac formula. To use this formula we go to an imaginary time parameter and study

$$G(x, -i\hbar T; y, 0) = \sum_{\alpha} e^{-E_{\alpha} T} \psi_{\alpha}(x) \psi_{\alpha}^*(y) \quad (21.3)$$

so that

$$E_0 = \lim_{T \rightarrow \infty} \left(-\frac{1}{T} \right) \log G(x, -i\hbar T; y, 0) \quad (21.4)$$

The path integral expression for G , in the case of $L = \frac{1}{2} \sum m_i \dot{x}_i^2 - V(x)$ and

with $\hbar = 1$, is

$$G(-iT) = \sum_{\text{paths}} \exp \left\{ - \int_0^T d\tau \left[\frac{1}{2} \sum m_i \dot{x}_i^2 + V(x) \right] \right\} \quad (21.5)$$

Note that the argument in the integral in the exponent is just the Hamiltonian. We are thus led to evaluate

$$G = \int dr(\cdot) \int \prod_k dq_k(\cdot) \exp \left\{ - \int_0^T dt \left[\frac{1}{2} \dot{r}^2 + \frac{1}{2} \sum (\dot{q}_k^2 + q_k^2) \right. \right. \\ \left. \left. + (2\sqrt{2} \pi \alpha)^{1/2} \sum_k \frac{1}{k} q_k e^{i\mathbf{k} \cdot \mathbf{r}} \right] \right\} \quad (21.6)$$

for some set of end points (designated r' , q'_k for the initial points and r'' , q''_k for the final positions).

Now we use the trick mentioned above. The interacting field and particle have been pulled apart from each other enough so that one of them becomes a completely solvable problem. We can do the path integral on the q 's, and the q 's will disappear.

Consider the path integral over one of the q_k and for convenience drop the subscript k :

$$\int_{(q', 0)}^{(q'', T)} dq(\tau) \exp \left\{ - \int_0^T \left[\frac{1}{2} (\dot{q}^2 + q^2) + q \gamma(\tau) \right] d\tau \right\} \quad (21.7)$$

where

$$\gamma(\tau) = (2\sqrt{2} \pi \alpha)^{1/2} \frac{\exp(i\mathbf{k} \cdot \mathbf{r}(\tau))}{k} \quad (21.8)$$

The path integral for a quadratic Lagrangian with arbitrary forcing term was done in Section 6 [(6.36), (6.41), and (6.42)]. The result (for imaginary times) is

$$[2\pi \sinh T]^{-1/2} \exp(-S_c)$$

where

$$S_c = (2 \sinh T)^{-1} \\ \times \left[(q'^2 + q''^2) \cosh T - 2q'q'' + 2q'' \int_0^T \gamma(t) \sinh t dt \right. \\ \left. + 2q' \int_0^T \gamma(t) \sinh(T-t) dt - 2 \int_0^T dt \int_0^t ds \gamma(t) \gamma(s) \sinh(T-t) \sinh s \right] \quad (21.9)$$

For the Feynman-Kac formula all we really need is the propagator at points where the wave function does not vanish. However, it is easy to take the matrix elements of G in the harmonic oscillator (q_k) ground state (and so obtain the vacuum-vacuum transition element), since all one has is the tedium of two more Gaussian integrals. The final result is

$$\sqrt{\frac{1}{2\pi \sinh T}} \exp \left[\frac{1}{2} \int_0^T \int_0^T dt ds \gamma^*(t) \gamma(s) \exp(-|t-s|) \right]$$

A number of remarks must be made about this formula. First, the expression $\exp(-|t-s|)$ appears only for T large and with the assumption that the important part of the path integral arises from X such that correlations between $X(s)$ and $X(t)$ are negligible unless $|t-s| \ll T$. (The actual expression should be

$$e^{-|t-s|} \frac{1}{1 - e^{-T}} [1 + e^{2|t-s|-T}] \quad)$$

Second, the appearance of $\gamma^*(t)$ rather than $\gamma(t)$ arises because the q_k coordinates, being defined as Fourier transforms of the real ionic coordinates, are inherently complex [so that H of (21.1) is in fact Hermitian], and when this is taken into account the complex conjugate $\gamma^*(t)$ results. (Formula 21.9 holds for real γ and real q .) Finally, the prefactor $[2\pi \sinh T]^{-1/2}$ should be dropped. In the Feynman-Kac formula it will give a term $\frac{1}{2}$ (or $\hbar\omega/2$ in other units) for each degree of freedom k . Dropping the prefactor thus means dropping the infinite zero point energy of the field. We conclude therefore that

$$G = \int dr(\cdot) \exp \left\{ -\frac{1}{2} \int_0^T \dot{r}^2 dt + 2\sqrt{2} \pi \alpha \times \sum_k \frac{1}{2k^2} \int_0^T \int_0^T dt ds \exp(i\mathbf{k} \cdot (\mathbf{r}(t) - \mathbf{r}(s))) e^{-|t-s|} \right\} \quad (21.10)$$

The quantity G is a propagator for r to go from some r' to r'' with the phonon field starting and ending in its vacuum state. The integral over k in (21.10) can be done (it is the Fourier transform of $1/x$) to yield

$$G = \int dr(\cdot) \exp \left\{ -\frac{1}{2} \int_0^T \dot{r}^2 dt + \frac{\alpha}{\sqrt{8}} \int_0^T \int_0^T ds dt \frac{e^{-|t-s|}}{|r(t) - r(s)|} \right\} \quad (21.11)$$

Looking at either (21.10) or (21.11) we see that getting rid of the field (of phonons) has effectively given the electron a self interaction that is nonlocal in time. When the same sort of elimination process is done on the photon field for a pair of electrons, the electrons are left with an interaction that is nonlocal in space and time—this is the action at a distance formulation of electrodynamics. The equations of motion there are difference differential equations with both advance and retardation. The singular denominator in the integral in (21.11) makes the associated difference differential (or renewal) equations a bit hard to interpret. There is also the problem of appropriate boundary conditions, since difference differential equations generally require boundary conditions on one or more intervals rather than the specification of a function and a finite number of its derivatives at a point.

The singular and apparently nonintegrable denominator in (21.11) is not a problem, since the paths entering the Wiener integral effectively have large (infinite) velocity and $1/|r(t)-r(s)| \sim 1/(v_{\text{effective}})|t-s|$. In fact, we shall never have to confront this problem directly, since to do the functional integral we shall not use (21.11) but rather go back to (21.10), do the integral for arbitrary k , and afterward integrate over k .

Actually, the integral (21.11) [or (21.10)] is too hard to do, and we shall use a variational principle to get a bound on the ground state energy, E_0 . The variational principle is based on Jensen's inequality.

Let φ_i , $i=1, \dots, N$ be a set of positive numbers and define the average of a function f_i as

$$\langle f \rangle_\varphi \equiv \frac{\sum_{i=1}^N \varphi_i f_i}{\sum_{i=1}^N \varphi_i} \quad (21.12)$$

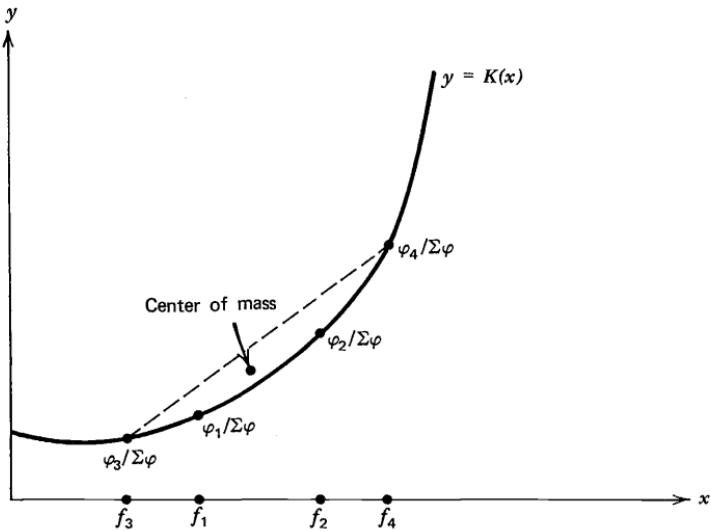
Then for any function $f(i)$ it is true that

$$\langle e^{-f} \rangle_\varphi \geq e^{-\langle f \rangle_\varphi} \quad (21.13)$$

It is just as easy to prove something more general. Let $K(x)$ be any convex* function of x , that is,

$$\frac{1}{2} [K(x) + K(y)] \geq K\left(\frac{x+y}{2}\right) \quad (21.14)$$

*Beware: in some terminology this is called concave.

Fig. 21.1 Graph of $K(x)$.

Then

$$\langle K(f) \rangle_{\varphi} \geq K(\langle f \rangle_{\varphi}) \quad (21.15)$$

We prove this pictorially. The graph of $K(x)$ curves upward—see Fig. 21.1. One can think of

$$\langle K(f) \rangle_{\varphi} = \sum_i \left(\frac{\varphi_i}{\sum_j \varphi_j} \right) K(f_i)$$

as the y coordinate of the center of mass of a system of mass points with masses $\varphi_i/\sum_j \varphi_j$, each located at $(f_i, K(f_i))$. This is illustrated in the figure. Similarly the quantity $\sum_i f_i (\varphi_i/\sum_j \varphi_j)$ is the x coordinate of the center of mass. It is “obvious” that the center of mass must lie above the curve. But this is exactly what was to be proved.

It only remains to show that the exponential is convex. This follows from

$$0 \leq (e^{a/2} - e^{b/2})^2 = e^a + e^b - 2e^{(a+b)/2} \quad (21.16)$$

We use (21.13) to formulate a variational principle for the energy. We reintroduce the notation

$$\langle F(r) \rangle_P = \frac{\int dr(\tau) e^{-P(r(\tau))} F(r(\tau))}{\int dr(\tau) e^{-P}} \quad (21.17)$$

for some action P . We now seek an S' such that $\int dr(\tau) \exp(-S')$ can be evaluated—that is, a simpler action for which $S - S'$ will be a perturbation. This is to be used coupled with (21.13) as follows:

$$\begin{aligned} G &= \int dr(\tau) e^{-S} = \int dr(\tau) e^{-S' - (S - S')} \\ &= \left(\int dr e^{-S'} \right) \langle e^{-(S - S')} \rangle_{S'} \geq \left(\int dr e^{-S'} \right) e^{-\langle (S - S') \rangle_{S'}} \end{aligned} \quad (21.18)$$

Asymptotically ($T \rightarrow \infty$) G will be

$$G \sim e^{-E_0 T} \quad E_0 = \text{ground state electron energy} \quad (21.19)$$

Similarly

$$\int dr e^{-S'} \sim e^{-E'_0 T} \quad E'_0 = \text{ground state energy with action } S' \quad (21.20)$$

Suppose that $\langle S - S' \rangle_{S'}$ grows linearly with time (asymptotically) as

$$\langle S - S' \rangle_{S'} \sim -\delta T \quad (21.21)$$

Then

$$\frac{e^{-E_0 T}}{e^{-E'_0 T}} \sim \langle e^{-(S - S')} \rangle_{S'} \geq e^{-\langle S - S' \rangle_{S'}} \sim e^{\delta T} \quad (21.22)$$

Taking logarithms,

$$E_0 \leq E'_0 - \delta \quad (21.23)$$

which is the variational principle we are after.

A good choice for S' will be an action that resembles S as much as possible but is at the same time solvable. This suggests the action

$$S' = \frac{1}{2} \int_0^T \dot{r}^2 dt + \frac{1}{2} C \int_0^T \int_0^T dt ds |r(t) - r(s)|^2 e^{-w|t-s|} \quad (21.24)$$

For this action we must evaluate E'_0 and δ . First

$$\begin{aligned} \delta \sim & \frac{1}{T} \langle S' - S \rangle_{S'} = \frac{\alpha}{\sqrt{8}} \int_0^T \int_0^T ds dt \left\langle \frac{1}{|r(t) - r(s)|} \right\rangle_{S'} e^{-|t-s|} \\ & + \frac{1}{2} C \int_0^T \int_0^T dt ds \langle |r(t) - r(s)|^2 \rangle_{S'} e^{-w|t-s|} \\ & \equiv A + B \end{aligned} \quad (21.25)$$

(which defines A and B). Both expectation values in (21.25) can be evaluated by Fourier transformation if one knows

$$\langle \exp[i\mathbf{k} \cdot (\mathbf{r}(t) - \mathbf{r}(s))] \rangle_{S'} = \frac{\int d\mathbf{r}(\cdot) e^{-S'} e^{i\mathbf{k} \cdot (\mathbf{r}(t) - \mathbf{r}(s))}}{\int d\mathbf{r} e^{-S'}} \quad (21.26)$$

It is possible to calculate the more general quantity

$$J(f(\cdot)) = \int d\mathbf{r}(\cdot) \exp(-S') \exp\left(\int_0^T f(t) r(t) dt\right) \quad (21.27)$$

and then to let $f(\tau)$ be

$$ik[\delta(\tau-t) - \delta(\tau-s)] \quad (21.28)$$

This is still just a harmonic oscillator, but now with a forcing term. J is therefore $\exp(-\text{classical action})$ times a square root term; the square root, however, does not depend on f and will cancel in (21.26), so we ignore it. The equation for the classical path (designated $x(t)$) is gotten by varying the action $S' - \int f r dt$:

$$\frac{d^2x}{dt^2} = -f(t) + 2C \int_0^T ds e^{-w|t-s|} [x(t) - x(s)] \quad (21.29)$$

We are able to confine attention here to a one-dimensional x , since the action is simply a sum of the Cartesian coordinates of \mathbf{r} . Since our

intention is to invoke the Feynman-Kac formula, the end points of x or \mathbf{r} can be treated cavalierly, and we take $x(0)=x(T)=0$. Then for x the solution of (21.29)

$$\text{classical action} = S'(x) - \int f x \, dt = -\frac{1}{2} \int_0^T f(t)x(t) \, dt \quad (21.30)$$

(this will be true for variation of any quadratic form $\sum x_i A_{ij} x_j + \sum f_i x_i$). Now to tackle (21.29). The trick is to define

$$z(t) = \frac{1}{2} w \int_0^T \exp(-w|t-s|) x(s) \, ds \quad (21.31)$$

Then

$$\frac{d^2 z}{dt^2} = w^2(z-x) \quad (21.32a)$$

and

$$\frac{d^2 x}{dt^2} = \frac{4C}{w}(x-z) - f \quad (21.32b)$$

Note that (21.32b) holds only with the neglect of end point contributions. With the function f given by (21.28), (21.32) can be solved by standard methods (again taking everything to vanish at the end points) and inserting the result in (21.30):

$$\begin{aligned} \langle \exp[i\mathbf{k} \cdot (\mathbf{r}(t) - \mathbf{r}(s))] \rangle_{S'} &= \frac{J(ik\delta(t-\tau) - ik\delta(\tau-s))}{J(0)} \\ &= \exp\left[\frac{-2Ck^2}{v^3 w}(1 - e^{-v|t-s|}) - \frac{w^2}{2v^2} k^2 |t-s|\right] \end{aligned} \quad (21.33)$$

where

$$v^2 = w^2 + \frac{4C}{w}$$

The term “ B ” in (21.25) is now evaluated from

$$\begin{aligned} & \left\langle \int_0^T \int_0^T ds dt (x(t) - x(s))^2 e^{-w|t-s|} \right\rangle_{S'} \\ &= - \int_0^T \int_0^T ds dt e^{-w|t-s|} \frac{d^2}{dk^2} \langle e^{ik(x(t)-x(s))} \rangle_{S'} \Big|_{k=0} \\ &= \frac{2T}{vw} \quad (\text{times 3 for 3 dimensions}) \end{aligned} \quad (21.34)$$

for large T . The term “ A ” is evaluated using the fact that $1/|r(t) - r(s)|$ is the Fourier transform of $1/k^2$. The resulting integral cannot be done explicitly and will appear in our final formulas below. The ground state energy of the approximate action S' is obtained by considering its rate of change with respect to the parameter C , thus

$$\begin{aligned} \frac{dE'_0}{dC} &= \frac{d}{dC} \left(-\frac{1}{T} \log \int dx e^{-S'} \right) \\ &= -\frac{1}{T} \left\langle -\frac{1}{2} \int_0^T \int_0^T dt ds |r(t) - r(s)|^2 e^{-w|t-s|} \right\rangle_{S'} \end{aligned}$$

But this is just the quantity evaluated in (21.34). Hence

$$\frac{dE'_0}{dC} = 3w^{-1} \left(w^2 + \frac{4C}{w} \right)^{-1/2}$$

which is integrated to

$$E'_0 = \frac{3}{2}(v-w)$$

where we have used the fact that $E'_0 = 0$ for $C=0$ because the particle is then free.

All these results are collected (including the integral for A which we have not yet written) to give

$$E_0 \leq \frac{3}{4v}(v-w)^2 - \frac{\alpha v}{\sqrt{\pi}} \int_0^\infty \frac{du e^{-u}}{\left[w^2 u + ((v^2 - w^2)/v)(1 - e^{-\nu v}) \right]^{1/2}} \quad (21.35)$$

In this form v and w are taken to be the independent variables, and one can now seek v and w so as to get the best bound on E_0 .

One immediate result is to let S' simply be the free particle action, that is, $v=w$. Then

$$E_0 \leq -\alpha$$

This is actually the same as the second-order perturbation theory result (which purports to give equality rather than a bound), but by varying v and w much better results can be obtained. In particular

$$E_0 \leq -\alpha - \frac{\alpha^2}{81} \quad \text{small } \alpha$$

$$E_0 \leq -\frac{\alpha^2}{2\pi} - \frac{3}{2}(2\log 2 + \gamma) - \frac{3}{4} + O\left(\frac{1}{\alpha^2}\right)$$

large α , $\gamma = \text{Euler's constant} \sim 0.577$

$E_0 \leq$ various numbers obtained from numerical integration of (21.35)

(21.36)

This is the final answer. Having seen it, it behooves one to ask in what way does knowing this result improve his knowledge of the physical world. I think the most important feature of (21.36) is simply the fact that it can be done and that within a relatively simple expression like (21.35) one can obtain the ground state energy over the entire range of coupling constant strengths. Further properties of the polaron can be deduced starting from this approach (see notes), but in the end the polaron itself is less important than the hopes it raises for the solution of more difficult problems, especially those involving strong coupling.

NOTES

The basic paper is

R. P. Feynman, Slow Electrons in a Polar Crystal, *Phys. Rev.* **97**, 660 (1955)

For background on the polaron see

H. Frohlich, *Adv. Phys.* **3**, 325 (1954)

Later work appears in

T. D. Schultz, *Phys. Rev.* **116**, 526 (1959)

R. P. Feynman, R. W. Hellwarth, C. K. Iddings, and P. M. Platzman, *Phys. Rev.* **127**, 1004 (1962)

Extensive discussions are given in

R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965

R. P. Feynman, *Statistical Mechanics*, Benjamin, New York, 1972

Feynman's various presentations are a bit hard to reconcile, as it is not always clear where he has changed his notation or where typographical errors have crept in. I have to confess that I have not reproduced all the derivations and in particular am not entirely certain of what should or need be done about the end points in (21.32).

Further sources for Jensen's inequality are

R. E. Edwards, *Functional Analysis*, Holt-Rinehart & Winston, New York, 1965, p. 291

M. Loeve, *Probability Theory*, Van Nostrand, New York, 1960, p. 159

If the density matrix is written as a product and Jensen's inequality applied to each factor, better bounds may be obtained. This technique is employed by

E. P. Gross, Path Integrals and Lower Bounds for Density Matrices, *J. Stat. Phys.* **21**, 215 (1979)

Recently, improved estimates on the polaron propagator with bounds in both directions have been given by Donsker and Varadhan. Their work is not yet (as of summer 1980) written, but an early paper which presents some of their methods is

M. Donsker and S. R. S. Varadhan, in *Functional Integration and its Applications*, A. M. Arthurs, ed., Oxford University Press, London, 1975

Spin and Related Matters

It is pleasing to think of path integration as an alternative approach to quantum mechanics. There are few who can go through Feynman's introductory exposition to the subject (two-slit experiment, etc.) without feeling that this approach eases the burden of accepting those parts of quantum mechanics that conflict with their naive intuitions. Instead of adding probabilities one adds probability amplitudes—that is really not so traumatic a departure from classical physics. However, before accepting this lightening of the philosophical burden one had better be sure that all the difficult parts of quantum mechanics have been accounted for. It would be—in fact, is—disconcerting to have to tack on little pieces of quantum mechanics in its traditional Schrödinger or Heisenberg formulation to the path integral structure. Feynman was quite sensitive to this difficulty, and the problem that forced on him a retreat to traditional formulations was that of representing spin as a path integral. Many particles have an internal spin degree of freedom, and because this variable takes discrete values it has been difficult to suggest for it a continuous path subsequently to be summed over so as to obtain the propagator.

In this section we discuss two approaches to the problem. One of them is quite straightforward. The other takes us a bit far afield to matters of homotopy theory, but I think it is closer in spirit to the original path integral and in addition, by introducing these ideas, allows some other problems to be considered.

22.1 THE DIRECT METHOD—PRODUCT INTEGRALS OR TIME ORDERED PRODUCTS

In the straightforward approach one ignores any philosophical or poetical considerations and asks quite simply what the path integral is supposed to

be. The answer is, a matrix element of the Green's function. Specification of the matrix element means (1) specification of the state and (2) determination of the Hamiltonian. For a single nonrelativistic particle of spin j state vectors can be labeled by

$$|\mathbf{x}, m\rangle \quad (22.1)$$

where m is the eigenvalue of s_z , the j quantum number is suppressed, and \mathbf{x} is the position. The Hamiltonian for such a particle is generally of the form

$$H = \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A} \right)^2 + V + 2\gamma \mathbf{s} \cdot \mathbf{B} \quad (22.2)$$

\mathbf{A} is the vector potential, $\mathbf{B} = \nabla \times \mathbf{A}$, \mathbf{s} = spin, γ = some constant. For $j > \frac{1}{2}$ there may be additional tensor operators combined with \mathbf{s} to form contributions to H . Nothing could be more straightforward than declaring the Green's function to be

$$G(x'', m'', t''; x', m', t') = \langle x'' m'' t'' | x' m' t' \rangle \quad (22.3)$$

For this G , the procedure of Section 1 gives non commuting pieces $\exp(i\epsilon p^2/2m\hbar)$ and $\exp(i\epsilon 2\gamma \mathbf{s} \cdot \mathbf{B}(\mathbf{x}, t)/\hbar)$ that cannot easily be recombined in a single exponential. It is nevertheless true that (letting $\psi(t)$ be the state vector at time t)

$$\psi(t) = \lim_{\epsilon = t/N \rightarrow 0} \exp(-iH(t_N)\epsilon) \cdots \exp(-iH(t_j)\epsilon) \cdots \exp(-iH(t_1)\epsilon) \psi(0) \quad (22.4)$$

$$t_j = \epsilon \left(j - \frac{1}{2} \right)$$

The object inside the limit, unwieldy as it may seem, is quite common in both mathematics and physics. In the mathematical literature it is known as a product integral, while in physics it is called the time ordered product. When we cut H down to include the spin energy only, we can consider it to be a finite dimensional matrix, and one then has comprehensive theorems that allow extensive manipulations to be performed. However, for the sake of familiarity we stick to physics notation and introduce the time ordered product

$$\begin{aligned} U(b, a) &\equiv T \exp \int_a^b (-iH(t)) dt \\ &\equiv \lim_{\epsilon \rightarrow 0} \exp[-iH(t_N)\epsilon] \cdots \exp[-iH(t_j)\epsilon] \cdots \exp[-iH(t_1)\epsilon] \\ \epsilon &= (b-a)/N \quad t_j = \left(j - \frac{1}{2} \right) \epsilon + a \end{aligned} \quad (22.5)$$

The principal properties of this object are

$$U(b, a) = 1 - i \int_a^b H(x) U(x, a) dx \quad (22.6)$$

$$i \frac{\partial U(t, a)}{\partial t} = H(t) U(t, a) \quad (22.7)$$

$$U(t, a) = 1 - i \int_a^t H(t_1) dt_1 - \int_a^t H(t_1) \int_a^{t_1} H(t_2) dt_2 dt_1 \quad (22.8)$$

$$+ \dots$$

$$\begin{aligned} T \exp \int_a^b (-i) [H_1(t) + H_2(t)] dt &= \left\{ T \exp \int_a^b (-i) H_1(t) dt \right\} \\ &\times \left\{ T \exp \int_a^b (-i) \left[T \exp \int_a^t (-i) H_1(x) dx \right]^{-1} \right. \\ &\quad \left. \times H_2(t) \left[T \exp \int_a^t (-i) H_1(x) dx \right] dt \right\} \quad (22.9) \end{aligned}$$

Of course the domain of validity of these equations depends on the properties of H , H_1 , and H_2 . The formulas are all well known in quantum mechanics; (22.9), for example, is the expression for the propagator in the interaction representation.

To go from physics notation to that of mathematics simply replace the symbol $T \exp \int_a^b$ by $\int_a^b \exp$.

In the absence of spin, if we were to take the coordinate space matrix element of each factor in (22.4) we would have the usual path integral formula. To see what happens for spin let us simplify to spin $\frac{1}{2}$ and assume that the space and spin coordinates can be dynamically separated in the sense that the spatial dynamics are uninfluenced by what is happening to the spin degrees of freedom (but not vice versa!). Thus H can be taken to be

$$H = \gamma \boldsymbol{\sigma} \cdot \mathbf{B}(t) \quad (22.10)$$

The time dependence in \mathbf{B} may arise from time variation of the field as well as from the changing position of the particle in a spatially varying field. This dynamical simplification occurs quite naturally in the WKB approximation. Assuming this simplification to hold, in the remainder of this subsection we consider only the spin propagator for the Hamiltonian (22.10).

While it may seem that an infinite product of matrices is not a very tractable object simple expressions may sometimes emerge. An example is

magnetic resonance, in which the magnetic moment of a system is determined through measurement of the frequency at which a maximum of energy is withdrawn by the system from a time varying magnetic field. The magnetic field can be idealized as being spatially homogeneous and of the form

$$\mathbf{B}(t) = \hat{\mathbf{z}} + \lambda(\hat{\mathbf{x}} \cos \omega t + \hat{\mathbf{y}} \sin \omega t) \quad (22.11)$$

Formula 22.9 is now used with

$$\begin{aligned} H_1 &= -\frac{1}{2}i\omega\sigma_z \\ H_2 &= -i\sigma_z\left(\gamma - \frac{1}{2}\omega\right) - i\gamma\lambda(\sigma_x \cos \omega t + \sigma_y \sin \omega t) \end{aligned} \quad (22.12)$$

It can then be easily checked that

$$U(t, 0) = T \exp \int_0^t -iH(t) dt = \exp\left(\frac{i}{2}\omega\sigma_x t\right) \exp\left\{-i\left[\sigma_z\left(\gamma - \frac{1}{2}\omega\right) + \sigma_x\gamma\lambda\right]t\right\} \quad (22.13)$$

The salient feature—experimentally speaking—of this formula is that if $\omega = 2\gamma$ the right hand factor of U periodically becomes a rotation through π about the x axis. Thus the original spin state, assumed to be in the $+z$ direction, is flipping with this period, and this is the observed resonance phenomenon.

The principle of using a time ordered product when a continuous space path integral fails can as well be applied to the Dirac equation. In an early paper Feynman appears to have given thought to this while C. DeWitt-Morette actually worked out some explicit formulas for a Dirac particle in a constant, external electromagnetic field. It is also clear that one could formulate a “path integral” for the dynamics of any discrete variable in this way, not just spin. And in fact this has been done in studying localized moments due to magnetic ion impurities in a metal.

22.2 CONTINUOUS MODELS FOR SPIN

We turn now to a more ambitious attempt to incorporate spin into a path integral formulation of quantum mechanics. We search for an appropriate space on which to define a *continuous* classical mechanics whose path integral quantization gives spin theory. There are three steps: (1) find the space, (2) define the dynamics, and (3) path integrate.

The main hint in the search for the space is that whatever the coordinates we finally settle on we already know that their conjugate momenta have well-defined properties under rotations. These momenta, after all, are supposed to become spin angular momenta when quantized. But in the context of a classical mechanics the momenta are obtained from the Lagrangian which is a well defined function of the coordinates and coordinate velocities. Hence we expect the coordinates to have definite transformation properties under the rotation group, $SO(3)$.

Let us call the coordinate space we are after M . This will play the role of an internal coordinate whose conjugate momenta are the spin observables. From the foregoing discussion it follows that $SO(3)$ sends M into itself. There are then two possibilities. Either starting from a single point in M any other point can be reached by an $SO(3)$ transformation. That is, $SO(3)$ is *transitive* on M . Or else $SO(3)$ is not transitive on M and in fact breaks M into a discrete or continuous family of orbits (an orbit is an image of a point under the transformation group). In the second case, M has, in addition to its spin degrees of freedom, some other internal coordinates (the orbit labels). Therefore in order to get spin and spin only, we assume $SO(3)$ to be transitive on M . Should the need arise, additional internal degrees of freedom can later be added.

Given reasonable smoothness assumptions, there are only three spaces on which $SO(3)$ is transitive, and these are the quotient spaces of $SO(3)$ relative to its subgroups. Specifically

- 1 $SO(3)$, the group itself.
- 2 $SO(3)/SO(2)$, essentially a two-sphere S^2 .
- 3 $SO(3)/SO(3)$, just a point.

The subgroup relative to which the quotient is taken is the isotropy group of the system, that is, those rotations that leave unchanged the mechanical object to which internal coordinates are ascribed. For example, case 2 might describe a dipole. In case 3, the object is spinless. An object having $M = SO(3)$ itself is the top. Both $SO(3)$ and the position of a top can be parametrized by the Euler angles.

Rejection of case 3 is clearly indicated, since it can give no spin momenta. In subsequent work it will become evident why the dipole is inappropriate for a model encompassing both integral and half integral spin. Thus, subject to our later responsibility to demonstrate this inappropriateness, we take M to be the group manifold of $SO(3)$ and call our dynamical object a top.

The next step in the program is to define the dynamics on M —that is, specify the Lagrangian. In our case, since M is identified with a compact Lie group (or one of its coset spaces) it automatically inherits a mechanics; the paths of the system are geodesics of the space. There is a natural metric associated with this Lie group, and what we are saying is that if M is the group then the metric coincides with the Lagrangian. This, however, does not exhaust the possibilities. Geodesic motion in $SO(3)$ corresponds to a spherical top. Symmetrical and asymmetrical tops are more general and have Lagrangians differing from the $SO(3)$ metric. Again, since the goal is to find the simplest model yielding spin there does not seem to be any reason to go beyond spherical tops. (It is conceivable that one might want to do this in order to fit a mass spectrum for related particles of different spin.) In any case, for present purposes the spinning object is a spherical top. ("Is" here means "mechanically indistinguishable from"—no spatial extension is implied for the spinning particle.) We remark that if one were looking at $SO(3)/SO(2)$ the analogous induced natural metric is that which gives as geodesics great circles on the sphere.

That we come in the end to calling our spinning object a top should not be surprising in view of the parallel requirements of spin models and path integrals. In 1950 Bopp and Haag and Rosen studied the quantum mechanical top with just this goal in mind, namely, to establish it as a model for spin. They brought out the fact that wave functions (and differential operators) could be defined on the internal coordinate space (these wave functions are the components of the rotation matrices) and that both integral and half integral spins appeared. When this top was placed in an electromagnetic field, agreement was found with the usual Pauli spinor theory.

We turn finally to the path integral for this classical mechanical system. The trajectory of the particle is in the six-dimensional space $\mathbb{R}^3 \times SO(3)$ where we indicate also the spatial degrees of freedom. The propagator has as arguments the initial and final times and positions in the product space.

The path integral we now wish to do is beset by two difficulties not previously encountered. First, the space $SO(3)$ is not simply connected. There are paths not continuously deformable into one another. Second, the space has an intrinsic curvature. In effect the "mass" is a function of the coordinates. Because the mass multiplies the most singular term in the Lagrangian (the velocity squared) great care must be exercised in evaluating expressions of this sort. For these reasons we hold off the completion of the path integral for spin until we have presented separate sections on the handling of the two foregoing problems.

NOTES

There have been many attempts at incorporating spin into path integral theory. Feynman has not written much describing his efforts in this direction although in Problem 2-6 of his book with Hibbs a relativistic path integral for the Dirac equation in one space dimension is described.

Some of the earliest formulations of field theory for fermions which included spin degrees of freedom too involved "anticommuting C-numbers." See

P. T. Matthews and A. Salam, Propagators of a Quantized Field, *Nuovo Cim.* **2**, 120 (1955)

Subsequently Klauder took a different approach in which the distinction between fermions and bosons revolved about an ambiguity in the functional integral due to the singular nature of the integrands. His work is

J. R. Klauder, The Action Option and a Feynman Quantization of Spinor Fields in Terms of Ordinary C-Numbers, *Ann. Phys.* **11**, 123 (1960)

Recent work in which pairs of fields are used is that of

J. Tarski, Feynman-Type Integrals for Spin and the Functional Approach to Quantum Field Theory, *Acta Phys. Austriaca* **44**, 89 (1976)

The approaches described in sections 22, 23, and 24 are those I have taken and are less ambitious than the foregoing in that there is neither field theory nor relativity. This work appears in the following sources

L. S. Schulman, A Path Integral for Spin, thesis, Princeton University, Princeton, N.J., 1967

L. S. Schulman, A Path Integral for Spin, *Phys. Rev.* **176**, 1558 (1968)

J. F. Hamilton, Jr., and L. S. Schulman, Path Integrals and Product Integrals, *J. Math. Phys.* **12**, 160 (1971)

The use of the top as a spin model quantized by methods other than path integration appears in

N. Rosen, Particle Spin and Rotation, *Phys. Rev.* **82**, 621 (1951)

F. Bopp and R. Haag, Über die Möglichkeit von Spinmodellen, *Z. Naturforsch.* **5a**, 644 (1950)

The work of C. DeWitt-Morette mentioned at the end of Section 22.1 is

C. Morette (DeWitt), On the Definition and Approximation of Feynman's Path Integrals, *Phys. Rev.* **81**, 848 (1951)

Recent work on the Dirac equation is

G. J. Papadopoulos and J. T. Devreese, Path Integral Solutions of the Dirac equation, *Phys. Rev. D* **13**, 2227 (1976)

T. Miura, Relativistic Path Integrals, *Prog. Theor. Phys.* **61**, 1521 (1979)

The application of path integrals to localized magnetic moments has been quite an active and useful pursuit. An incomplete list of references is:

B. Mühlshlegel, Functional Integrals and Local Many-Body Problems: Localized Mo-

ments and Small Particles, in *Functional Integration and Its Applications*, A. M. Arthurs, Ed., Clarendon, Oxford, 1975

J. R. Schrieffer, lecture notes, Canadian Association of Physicists, Summer School, Banff, 1969 (unpublished)

R. F. Hassing and D. M. Esterling, Functional-Integral Approaches to the Anderson Model: A Comparison of Results in Limiting Cases, *Phys. Rev. B* 7, 432 (1973)

D. J. Amit and C. M. Bender, Path-Integral Approach to the Magnetic Impurity Problem, *Phys. Rev. B* 4, 3115 (1971)

R. Kishore, Functional Integral Approach to the Static Susceptibility of the Anderson Model, *Nuovo Cim.* 55B, 143 (1980)

The works by Mühllegel and by Hassing and Esterling give many earlier references.

TWENTY-THREE

Path Integrals for Multiply Connected Spaces

This section consists of four parts. We first examine a particular example of a path integral on a multiply connected space. Then we go into a certain amount of mathematics (algebraic topology) to facilitate the handling of general cases. Next we return to the definition of the path integral using the mathematical formalism. Finally we use another branch of mathematics—functional analysis—for a deeper analysis of what was done in the preceding subsection.

23.1 PARTICLE CONSTRAINED TO A CIRCLE

Consider a system whose one degree of freedom is constrained to a circle. It can be thought of as a rotator whose axis is fixed, or as a bead on a ring, or—to cite a case somewhat closer to the real world—as an electron in a periodic infinite lattice with periodic boundary conditions. In any case, the coordinate is φ , $0 \leq \varphi \leq 2\pi$ with $\varphi=0$ and $\varphi=2\pi$ identified. As Lagrangian take

$$L = \frac{1}{2} I \dot{\varphi}^2 \quad (23.1)$$

which is also the natural metric if one views this space as the group manifold of $SO(2)$. A continuous path on this system is a continuous function $\varphi(t)$, again with the identification of 0 and 2π . The set of such paths can be broken into subsets labeled by their “winding number,” the

number of times the path goes past some particular point moving counter-clockwise minus the number of times it passes that point going clockwise. Two paths having the same end points and the same winding number can be deformed into one another; that is, holding the end points fixed, one of the paths can be modified in a continuous way until it is the same as the other path. Such deformations cannot be found for paths with different winding numbers.

This allows us to contemplate a new sort of freedom in doing the path integral. Write the propagator for our system as (take $\hbar=1$ for this section)

$$\begin{aligned} K(\varphi'', t''; \varphi', t') &= \sum_{\substack{\varphi(t') = \varphi' \\ \varphi(t'') = \varphi''}} e^{iS[\varphi(t)]} \\ &= \sum_{n=-\infty}^{\infty} \sum_{\varphi(t) \in g_n} e^{iS[\varphi(t)]} \end{aligned} \quad (23.2)$$

where

$$\begin{aligned} g_n = \{ \varphi(t) | t' \leq t \leq t'', \varphi(t') = \varphi', \varphi(t'') = \varphi'', \\ \varphi(t) \text{ is continuous, } \varphi(t) \text{ has winding number } n \} \end{aligned}$$

The restriction to continuous paths is justified by the knowledge that for the Weiner integral, as one expects for the path integral, almost all paths are continuous but nowhere differentiable. Now the proof that the path integral satisfies the appropriate Schrödinger equation is a local one: in Section 4 the wave function was propagated infinitesimally. It is thus at least plausible and we shall now make the assumption that each term in the sum over n in (23.2) *individually* satisfies Schrödinger's equation. (In Section 23.3 this assumption will be justified.) It follows that

$$\sum_n A_n \sum_{\varphi \in g_n} e^{iS(\varphi)} \quad (23.3)$$

with A_n not all unity is fully as good as a candidate for the propagator as the expression offered in (23.2) (i.e., taking all the A_n to be unity). The foregoing statement has to be weakened a bit: there are nontrivial $\{A_n\}$'s that can be used in (23.3) but they are not arbitrary; in fact they must be of the form

$$A_n = e^{in\delta} \quad \delta \text{ real} \quad (23.4)$$

To see how this comes about, let us imagine that we have calculated K and

it has the form

$$K(\varphi'', t''; \varphi', t') = \sum_{n=-\infty}^{\infty} A_n K_n(\varphi'', t''; \varphi', t') \quad (23.5)$$

where K_n is the sum over paths in the set g_n . Let φ'' vary (t'' , etc., fixed), and follow the changing values of K_n and K —the numbers A_n of course do not depend on φ . If φ'' goes through a complete circle, the K_n 's become the K_{n-1} 's while K itself must show no physical change, which is to say it can only be multiplied by a phase factor. Calling this factor $e^{i\delta}$ and assuming the linear independence of the K_n 's we immediately have

$$A_{n+1} = e^{i\delta} A_n \quad (23.6)$$

The magnitude of A_0 is fixed by unitarity to be 1, and taking its (arbitrary) phase to be zero we get (23.4).

Before going to further abstract properties—covering spaces and essential self adjointness—we perform the path integral for (23.1), since the actual result has some interest in itself.

Consider the mapping from the real line \mathbf{R} to the circle

$$p : \mathbf{R} \rightarrow S^1 \quad p(x) = x - \left[\frac{x}{2\pi} \right] 2\pi \quad (23.7)$$

where the bracket means “integer part of,” so that $0 \leq p(x) < 2\pi$. Locally, p is invertible, so that once one particular preimage of some φ is agreed upon, neighborhoods of φ map uniquely up to \mathbf{R} . One need only be careful that neighborhoods containing 0 map into successive multiples of 2π . In Fig. 23.1 the mapping p is illustrated. Thus, given a continuous path $\varphi(t)$ with $\varphi(0) = \varphi'$ and $\varphi(1) = \varphi''$, if a definite $p^{-1}(\varphi')$ is selected, then $\varphi(t)$ can be retracted to \mathbf{R} , that is, $p^{-1}(\varphi(t))$ is well defined. It is clear then that paths $\varphi(t)$ having the same end points and with the same choice of $p^{-1}(\varphi')$ can end at different preimages of φ'' and that the winding number is essentially the number $[p^{-1}(\varphi'')/2\pi]$ that must be subtracted from $p^{-1}(\varphi'')$ to get back to S^1 (“essentially” because the reference point defining the winding number has not been specified). It follows that to calculate K_n , the sum over paths for paths with winding number n , one can do the path integral on \mathbf{R} for the path that begins at the definite $p^{-1}(\varphi')$ that had been selected and ends at $p^{-1}(\varphi'') + 2\pi n$ where $p^{-1}(\varphi'')$ is the preimage of φ'' fixed by the selection of the preimage of φ' . Because the map p is smooth there is no problem in using it to carry the classical Lagrangian (or for that matter the Schrödinger equation) from S^1 to \mathbf{R} .

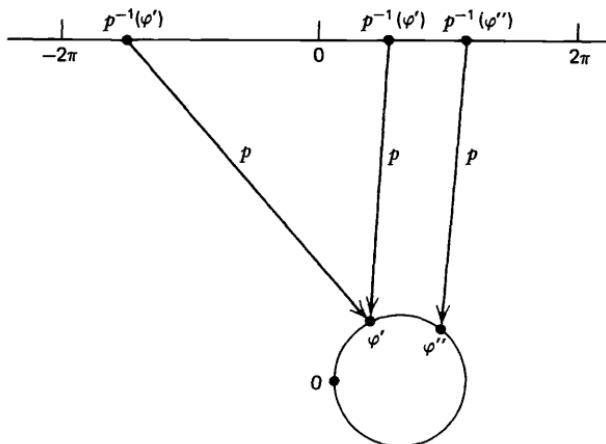


Fig. 23.1 The map $p: \mathbb{R} \rightarrow S^1$ for points on the line to some φ' and to some φ'' . Two points in the inverse image of φ' are shown.

On \mathbb{R} the Lagrangian (23.1) is that of a free particle, and we immediately have

$$K_n(\varphi'', t''; \varphi', t') = \left(\frac{I}{2\pi iT} \right)^{1/2} \exp\left(\frac{iI}{2T} (\varphi - 2n\pi)^2 \right) \quad (23.8)$$

with

$$\varphi = \varphi'' - \varphi', \quad T = t'' - t' \quad (23.9)$$

and the point for reckoning winding number has been taken to be 0. Putting (23.8) and (23.6) into (23.5) yields the propagator

$$K(\varphi'', t''; \varphi', t') = \sum_{n=-\infty}^{\infty} \left(\frac{I}{2\pi iT} \right)^{1/2} \exp\left[in\delta + \frac{iI(\varphi - 2n\pi)^2}{2T} \right] \quad (23.10)$$

K is therefore one of the classical functions, the Jacobi theta function, whose definition we now recall:

$$\theta_3(z, t) = \sum_{n=-\infty}^{\infty} \exp(i\pi tn^2 + i2nz) \quad (23.11)$$

In the half plane $\text{Im}(t) > 0$, θ_3 is analytic in z . It is quasiperiodic and

satisfies

$$\begin{aligned}\theta_3(z + \pi, t) &= \theta_3(z, t) \\ \theta_3(z + \pi t, t) &= e^{-i\pi t - 2iz} \theta_3(z, t)\end{aligned}\quad (23.12)$$

A remarkable property of θ_3 , following from the Poisson summation formula, is

$$\theta_3(z, t) = (-it)^{-1/2} e^{z^2/i\pi t} \theta_3\left(\frac{z}{t}, -\frac{1}{t}\right) \quad (23.13)$$

Comparing (23.10) and (23.11) and using the fact that $\theta_3(z, t) = \theta_3(-z, t)$ we have

$$K(\varphi'', t''; \varphi', t') = \left(\frac{I}{2\pi i T}\right)^{1/2} e^{iI\varphi^2/2T} \theta_3\left(\frac{\pi\varphi I}{T} - \frac{\delta}{2}, \frac{2\pi I}{T}\right) \quad (23.14)$$

Because I/T is real this function is not analytic and is only the boundary value of the analytic function defined by giving I/T a positive imaginary part.

It is worth comparing the propagator (23.14) to the conventional quantum mechanical expression so as to gain some interpretation of δ and incidentally to check the result. Schrödinger's equation on the circle is

$$-\frac{1}{2I} \frac{\partial^2 \psi}{\partial \varphi^2} = i \frac{\partial \psi}{\partial t} \quad (23.15)$$

Stationary state solutions to this equation have the form

$$\begin{aligned}\psi_m(\varphi, t) &= u_m(\varphi) e^{-iE_m t} \quad u_m(\varphi) = \frac{1}{\sqrt{2\pi}} \exp\left(im\varphi + \frac{i\delta\varphi}{2\pi}\right) \\ E_m &= \frac{1}{2I} \left(m + \frac{\delta}{2\pi}\right)^2 \quad m = 0, \pm 1, \dots\end{aligned}\quad (23.16)$$

where we have *not* insisted that $u(2\pi) = u(0)$. The quantity δ is precisely the phase change each vector shows under a 2π rotation, and we later return to possible physical interpretations of δ . The Green's function as a sum over stationary states is

$$G(\varphi'', t''; \varphi', t') = \sum_m \psi_m(\varphi'') \psi_m^*(\varphi') \exp(-iE_m(t'' - t')) \quad (23.17)$$

This too is obviously the Jacobi theta function and is given by

$$G(\varphi'', t''; \varphi', t') = \frac{1}{2\pi} \exp \left[\frac{i\delta\varphi}{2\pi} - \frac{i\delta^2 T}{8I\pi^2} \right] \theta_3 \left(\frac{\varphi}{2} - \frac{T\delta}{4\pi I}, \frac{-T}{2\pi I} \right) \quad (23.18)$$

Expressions 23.18 and 23.14 are two seemingly different quantities for the propagator, the one derived from a sum over stationary states, the other from a sum over classical paths. The fact that they are equal is guaranteed precisely by (23.13), the fundamental identity for theta functions.

The degree of pathology exhibited by this Green's function is entertaining, especially in view of the elementary nature of the example. This pathology of course has nothing to do with path integrals, since the Green's function is that obtained from ordinary quantum mechanics. For $\text{Im } t > 0$ the zeros of the theta function are doubly periodic, forming an infinite lattice in the z plane. As $\text{Im } t \rightarrow 0$ the lattice collapses and the zeros move towards the real z axis. For t real and rational there will be a finite number of zeros in any interval. For t irrational the zeros are dense on the real z axis by Kronecker's theorem. (If q is irrational the set of points $nq - [nq], n=0, 1, \dots$ is dense in the interval $(0, 1)$.) This does not imply that G is identically zero, since we are now very much without analyticity. In fact, we know that as an integral kernel G is norm preserving and has nonzero integrals.

A number of physical interpretations are possible for the phase δ . Suppose that one attempted to study the properties of a periodic one-dimensional crystal by concentrating on a single cell of the lattice. Clearly, perfectly periodic boundary conditions (i.e., no phase change) for such a cell would be inappropriate. In fact, for Bloch waves the phase should advance $2\pi\mathbf{k}\cdot\boldsymbol{\rho}$ where \mathbf{k} is the wave vector and $\boldsymbol{\rho}$ a lattice vector. This phase is identified with δ . A second physical context within which a phase change is appropriate is the Aharonov-Bohm effect. The "experimental" layout for seeing this effect is shown in Fig. 23.2. Although the electron can move in a plane, the idealized impenetrability of the solenoid (so that there is in effect a hole in the plane) makes useful the idea of winding number for the trajectories entering the path integral, and an arbitrary relative phase enters the wave function for trajectories passing through different slits. This phase has the real physical effect of shifting the diffraction maxima. On the other hand, if the idealization of impenetrability is not made, the magnetic field within the solenoid, and its vector potential which extends everywhere, provide the shift of the diffraction pattern. It turns out that in this situation the phase change is just the magnetic flux through the solenoid.

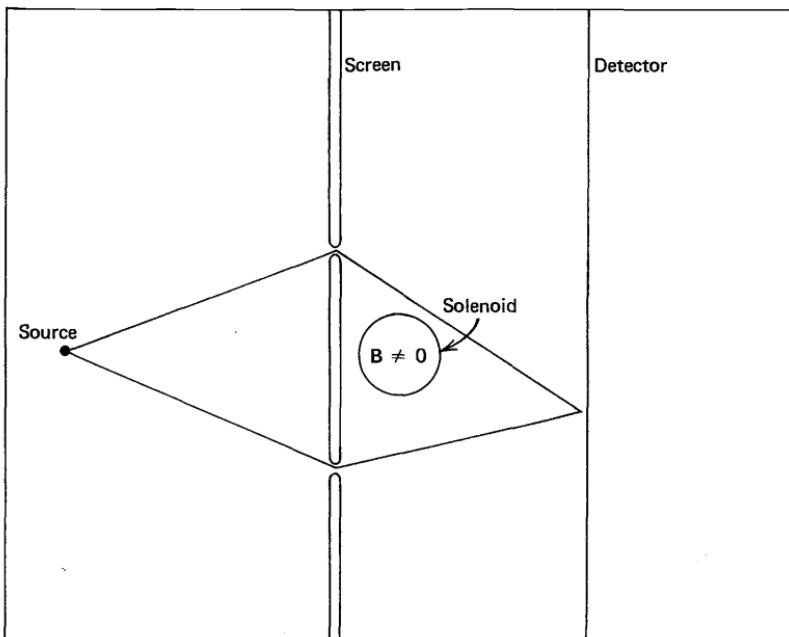


Fig. 23.2 Aharonov-Bohm effect. Electrons leave the source and pass through the slits. A solenoid with $B \neq 0$ within its walls is situated in the shadow of the screen. The walls of the solenoid are impenetrable. It is assumed that conditions are such that the WKB concepts implicit in the foregoing description are appropriate.

Notes for Section 23.1

The path integral on a circle was considered by

W. K. Burton and A. H. DeBorde, *Nuovo Cim.* **2**, 197 (1955)

but exploitation of the additional freedom due to homotopy theory first appeared (to our knowledge) in

L. S. Schulman, *Phys. Rev.* **176**, 1558 (1968)

For Kronecker's theorem see

G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon Press, Oxford, 1962, p. 376

The applications described near the end of Section 23.1 (Aharonov-Bohm effect and periodic lattices) are given in

L. S. Schulman, *Phys. Rev.* **188**, 1139 (1969)

L. S. Schulman, *J. Math. Phys.* **12**, 304 (1971)

23.2 RUDIMENTS OF HOMOTOPY THEORY

The concepts developed to do the path integral on a circle can easily be generalized to other multiply connected spaces. But to state this generalization most elegantly one must get involved in a certain amount of algebraic topology, specifically homotopy theory. We now present an exposition of this subject, providing only the most heuristic of proofs (or none at all) in the hope that some of the concepts can be transmitted even without the full details.

Consider continuous functions from a space X to a space Y

$$f, g: X \rightarrow Y \quad (23.19)$$

It goes without saying that X and Y are topological spaces and have whatever good properties a pathology abhorring nature would bestow on spaces of physical interest. In particular, they are assumed to be arcwise connected: any two points in X (or in Y) can be connected by a continuous curve.

Definition: f is homotopic to g , written $f \sim g$, if there exists a (continuous) function $h: X \times I \rightarrow Y$ (where $I = [0, 1]$, the unit interval) such that for all x in X ,

$$\begin{aligned} h(x, 0) &= f(x) \\ h(x, 1) &= g(x) \end{aligned} \quad (23.20)$$

Thus f has been continuously deformed into g .

THEOREM: Homotopy (\sim) is an equivalence relation.

Proof: As an exercise in the use of the preceding definition we prove one of the properties of an equivalence relation; namely, if for three functions f , g , and k ,

$$f \sim g \quad \text{and} \quad g \sim k, \quad \text{then} \quad f \sim k$$

By definition there are continuous functions h_1 and h_2 such that

$$h_1(x, 0) = f(x), \quad h_1(x, 1) = g(x), \quad h_2(x, 0) = g(x), \quad h_2(x, 1) = k(x)$$

Then the function deforming f to k can be taken to be

$$h(x, t) = \begin{cases} h_1(x, 2t) & 0 \leq t \leq \frac{1}{2} \\ h_2(x, 2t-1) & \frac{1}{2} \leq t \leq 1 \end{cases}$$

The important consequence of the foregoing theorem is that homotopy, like all equivalence relations, can be used to break up a space into equivalence classes. These "homotopy classes" are sets of functions all homotopic to each other. The equivalence class of a function f will be denoted $[f]$.

We shall be interested in classes of paths that are deformable into each other *and* have the same initial point as well as the same final point. For this the concept of *relative* homotopy is appropriate.

Definition: Let $f, g: X \rightarrow Y$ and let $X_0 \subset X$. Then f is homotopic to g relative to X_0 , written $f \sim g$ rel X_0 , if $f \sim g$ and if for the function h deforming f into g

$$h(x_0, t) = f(x_0) = g(x_0) \quad \text{for all } t \in I \text{ and all } x_0 \in X_0$$

Obviously relative homotopy is also an equivalence relation and we denote its equivalence classes in the same way as for homotopy.

The constant map $c: X \rightarrow Y$, $c(x) = y_0$ for some particular y_0 and all x can always be defined and any function homotopic to c is said to be homotopic to a constant. (Use the fact that Y is arcwise connected to prove that all constant maps $X \rightarrow Y$ (for various y_0) are homotopic to each other.) Let i be the identity map, $i: X \rightarrow X$, $i(x) = x$.

Definition: If the identity map $i: X \rightarrow X$ is homotopic to a constant, X is said to be *contractible*.

THEOREM: If Y is a contractible space, then every $f: X \rightarrow Y$ is homotopic to a constant.

Proof: Any $f: X \rightarrow Y$ can be written $f \circ i: X \rightarrow Y$ and then deform i into a constant [$(f \circ g)(x)$ means $f(g(x))$].

This theorem indicates that contractible spaces have no homotopy structure. The n -dimensional Euclidean space (E^n) is contractible, and it is the fact that path integrals are generally done on that space that leaves homotopy out of the picture. On the other hand, S^2 and (the group manifold of) $SO(3)$ are not contractible, this property of $SO(3)$ being the justification for our lengthy digression into homotopy theory.

A comment is here in order on the general philosophy of algebraic topology. *Topology* declares two spaces to be essentially the same if there is a one to one, onto continuous mapping between them. This can be a stringent requirement. *Algebraic* topology tries to classify spaces through

algebraic structures associated with them (we shall soon find that the homotopy equivalence classes can be given a group structure), and while that classification can be a thing of great efficacy and beauty, it is also much coarser than the topological classification. We have just seen an example of this in that homotopy theory lumps together E^n for all n (as well as many other spaces). In fact the role of homotopy theory in classifying spaces can be formalized even before we study homotopy group theory through the following definition:

Definition: Spaces X and Y are said to be of the same homotopy type if there exist functions f and g , $f: X \rightarrow Y$ and $g: Y \rightarrow X$, such that

$$\begin{aligned} g \circ f: X \rightarrow X &\text{ is homotopic to the identity on } X, \text{ and} \\ f \circ g: Y \rightarrow Y &\text{ is homotopic to the identity on } Y \end{aligned}$$

Obviously this defines an equivalence relation.

The preoccupation of homotopy theory with paths makes it appropriate for our intended physical applications. First some definitions:

Definition: A path in a topological space X is a (continuous) map $\alpha: I \rightarrow X$. $\alpha(0)$ is the beginning. $\alpha(1)$ is the end. If $\alpha(0) = \alpha(1)$, α is called closed. The inverse path α^{-1} is the function $\alpha(1-u)$, $u \in I$. If for some $x_0 \in X$, $\alpha(u) = x_0$ for all $u \in I$, α is said to be a null path. For two paths $\alpha, \beta: I \rightarrow X$, such that $\alpha(1) = \beta(0)$, the product path $\alpha\beta$ is that γ for which

$$\gamma(u) = \begin{cases} \alpha(2u) & 0 \leq u \leq \frac{1}{2} \\ \beta(2u-1) & \frac{1}{2} \leq u \leq 1 \end{cases}$$

Confusion will be avoided if it is realized that the product path is not the composition of two functions and that $\alpha\beta$ means—first travel on α , then on β . The concept of homotopy for paths is merely a special case of that for more general functions:

Definition: Two paths α and β are homotopic if $\alpha \sim \beta$ rel $\{0, 1\}$.

The symbols “rel $\{0, 1\}$ ” will be omitted when understood from context. An illustration of homotopy appears in Fig. 23.3. Several properties of the product are important for the later establishment of a group structure:

THEOREM (1) If $\alpha \sim \gamma$, $\beta \sim \delta$ and $\alpha\beta$ exists, then $\gamma\delta$ exists and $\alpha\beta \sim \gamma\delta$.
 (2) If $\alpha \sim \beta$, then $\alpha^{-1} \sim \beta^{-1}$.

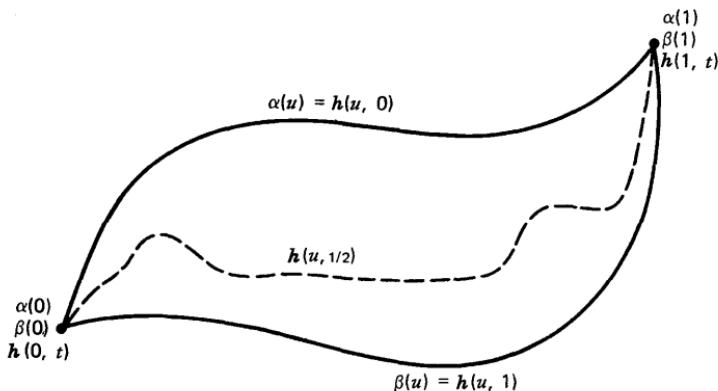


Fig. 23.3 The paths α and β (solid lines) are homotopic and are deformed into one another by the homotopy map $h: I \times I \rightarrow Y$. An intermediate value of h is shown as a dashed line.

(3) If β is null and $\alpha\beta$ exists, then $\alpha\beta \sim \alpha$. If γ is null and $\gamma\alpha$ exists, then $\gamma\alpha \sim \alpha$.

(4) If for three paths α , β , and γ both $\alpha\beta$ and $\beta\gamma$ exist, then $(\alpha\beta)\gamma$ and $\alpha(\beta\gamma)$ exist and $(\alpha\beta)\gamma \sim \alpha(\beta\gamma)$.

(5) For all α , $\alpha\alpha^{-1}$ and $\alpha^{-1}\alpha$ exist and are homotopic to null paths.

(6) If $\alpha\beta^{-1}$ exists and is closed then $\alpha\beta^{-1}$ is homotopic to a null path if and only if $\alpha \sim \beta$.

Proof: As an example we prove statement (3). That $\alpha\beta$ exists simply means that $\alpha(1) = \beta(0)$. Thus $\beta(u) = \alpha(1)$ for all u . We must produce a function $h(u, t)$ continuous in both variables for which $h(u, 0) = \alpha\beta$, $h(u, 1) = \alpha$, $h(0, t) = \alpha(0)$ and $h(1, t) = \alpha(1)$. One such function is

$$h(u, t) = \begin{cases} \alpha\left(\frac{2u}{1+t}\right) & 0 \leq u \leq \frac{1}{2}(1+t) \\ \alpha(1) & \frac{1}{2}(1+t) \leq u \leq 1 \end{cases}$$

The other statements in the theorem are similarly easy applications of the definitions.

Each path belongs to some equivalence class, and the product of equivalence classes is defined in the obvious way:

$$[\alpha][\beta] = [\alpha\beta]$$

The product is well defined in that $[\alpha\beta]$ is independent of which representatives α and β are chosen from the classes $[\alpha]$ and $[\beta]$. Associativity follows from (4) of the theorem. The identity element for the product is the class of a null path, and the inverse of a class is the class of the inverse.

We are now ready to introduce one of the central concepts of homotopy theory: the fundamental group. Consider closed paths all having the same beginning (and end). This common point is called the base point. Then the equivalence classes of these paths (under relative homotopy) form a group with the above mentioned product. On a space X and with base point x_0 this group is designated $\pi_1(X, x_0)$, the fundamental group.

THEOREM: If $x, y \in X$ and X is arcwise connected, then $\pi_1(X, x)$ is isomorphic to $\pi_1(X, y)$.

Proof: Let γ be a path such that $\gamma(0)=x$ and $\gamma(1)=y$. Then if $[\alpha] \in \pi_1(X, y)$ the path $\gamma\alpha\gamma^{-1}$ starts and ends at x . The mapping providing the isomorphism is obviously $[\alpha] \rightarrow [\gamma][\alpha][\gamma^{-1}]$.

For arcwise connected spaces the base point is often deleted in the designation: $\pi_1(X)$. When two spaces are related, so are their fundamental groups. The following is easy to prove.

THEOREM: Suppose that there exists a continuous map $f: X_1 \rightarrow X_2$. Then there exists a homomorphism $f^*: \pi_1(X_1, x) \rightarrow \pi_1(X_2, f(x))$. If f is a homeomorphism, then f^* is an isomorphism.

Actually isomorphism is easier to obtain than might be supposed from the foregoing theorem. It is enough that X_1 and X_2 be homotopic:

THEOREM: If X_1 and X_2 are arcwise connected and of the same homotopy type, then $\pi_1(X_1)$ is isomorphic to $\pi_1(X_2)$.

Proof: Use the maps appearing in the definition of homotopy equivalence for spaces to produce the isomorphism.

Definition: An arcwise connected space X for which $\pi_1(X) = \{e\}$ is said to be simply connected.

Some examples may be useful at this point. The sphere $S^2 = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$ is simply connected. Although S^2 is not contractible its fundamental group is nevertheless trivial. Any curve on S^2 can be shrunk to a point. There are higher homotopy groups that distinguish S^2 from a point, but we shall not study them. Essentially they involve maps from

$I \times I$ or higher dimensional cubes to the space in question. Hence also the index 1 on π_1 .

The circle S^1 has paths that cannot be deformed into each other. In fact a path with a given winding number (defined earlier) cannot be deformed into a path with a different winding number. Moreover, the product of two paths has as its winding number the algebraic sum of the two winding numbers. Thus $\pi_1(S^1)$ can be taken to be the additive group of the integers, designated \mathbb{Z} .

The two-dimensional torus is just $S^1 \times S^1$, and its fundamental group is $\mathbb{Z} \times \mathbb{Z}$.

The figure 8 (e.g., $\text{EIGHT} = \{(x, y) \in \mathbb{R}^2 | x^2 + (y - 1)^2 = 1 \text{ or } x^2 + (y + 1)^2 = 1\}$) has subgroups of its homotopy group generated by single circuits around each of its constituent circles. However, these generators do not commute so that $\pi_1(\text{EIGHT})$ is the noncommutative group consisting of strings of products of these generators.

The group manifold of the rotation group $SO(3)$ is not simply connected. One can take as this space the solid ball of radius π with antipodal points identified. A counterclockwise rotation of φ about an axis \hat{n} corresponds to a point a distance φ from the center of the ball along the radius pointing in the direction \hat{n} . A path in this space (i.e., a continuous family of rotations) that passes through the outer surface of the ball once cannot be deformed into a path that does not pass through at all. However, a path that passes through any even number of times can be so deformed. It follows that the fundamental group is the group of two elements, the identity corresponding to the equivalence class of null paths (as always) and the other element including rotations about a single axis through angles $(2j+1)2\pi$, j integral. (Of course other paths are also included.) For elaboration and pictures see Hamermesh's book.

For purposes of path integration an extremely useful construct is the universal covering space. Consider a space X with a nontrivial fundamental group. The universal covering space is a space that locally looks just like X , but is nevertheless simply connected. In some sense the universal covering space has many copies of X pasted together in such a way that the result is simply connected. The advantage that will accrue to us in path integration is that we shall be able to path integrate on the covering space and then use our results there to produce a propagator on X .

The definition of covering space (and in particular of universal covering space) involves the concept of covering projection. Although the following definition may obscure the point, the covering projection is basically a homeomorphism between small neighborhoods of the covering space and small neighborhoods of the base space—reflecting the notion that locally the two spaces look the same.

We shall always assume the space X to be arcwise connected, locally arcwise connected, and locally simply connected. The last two properties mean that sufficiently small neighborhoods of X , considered as spaces in their own rights, are arcwise connected and simply connected.

Definition: A mapping $p: B \rightarrow X$ is a *covering projection* if p is onto and for all $x \in X$ there exists an open set U containing x such that $p^{-1}(U) \subset B$ is open and p restricted to $p^{-1}(U)$ (i.e., $p|_{p^{-1}(U)}: p^{-1}(U) \rightarrow U$) is a homeomorphism. B is said to be a *covering space* of X .

THEOREM: Suppose that α and α' are paths in X , $\alpha(0)=\alpha'(0)=x_0$, $\alpha(1)=\alpha'(1)=x_1$ and $\alpha \sim \alpha'$ rel $\{x_0, x_1\}$. Then given a specific preimage of x_0 , $p^{-1}(x_0)$, both $p^{-1}(\alpha)$ and $p^{-1}(\alpha')$ are uniquely defined paths in B and $p^{-1}(\alpha) \sim p^{-1}(\alpha')$ rel $\{p^{-1}(x_0), p^{-1}(x_1)\}$ (the preimage $p^{-1}(x_1)$ is uniquely picked by $p^{-1}(\alpha)$).

Proof: To get the unique preimage of a path (say α), start from $p^{-1}(x_0)$ and move along in steps by surrounding $p^{-1}(x_0)$ and successive points by the sufficiently small neighborhoods for which p is a homeomorphism. To show that $p^{-1}(\alpha)$ and $p^{-1}(\alpha')$ are homotopic just use p^{-1} to lift to B the mapping $h = I \times I \rightarrow X$ that deforms α into α' .

The significance of the theorem is that it allows us to map homotopy classes in X uniquely to the homotopy classes of B with preservation of the product. The theorem is a simple but important example of a “lifting theorem” concerned with the circumstances under which a mapping $f: Y \rightarrow X$ can be lifted to a mapping $f^*: Y \rightarrow B$ (where $p: B \rightarrow X$) in such a way that $p \circ f^* = f$ (the diagram is commutative). In our case $Y = I$.

While it is true that every closed path in B projects downward to a closed path in X , the closed paths in X may very well be lifted so that $p^{-1}(x_0)$ takes different values at the beginning and end of the lifted path (x_0 is the base point). Thus p induces a map $\bar{p}: \pi(B) \rightarrow \pi(X)$ which is in fact an isomorphism between $\pi(B)$ and its image (which is in general smaller than $\pi(X)$). The property of isomorphism follows because homotopic paths in B are homotopic in X , using p to send the homotopy map down to X . On the other hand, paths not homotopic in B could not be homotopic in X for otherwise, by the theorem, the homotopy mapping could be lifted to B . Thus $\pi(B)$ can only be the same or smaller than $\pi(X)$.

In general X can have many covering spaces with any of the subgroups of $\pi(X)$ being $\pi(B)$. There is one special covering space known as the universal covering space, for which $\pi(B) = \{e\}$. There is a canonical way of

constructing this space, and it is universal in the sense that a covering projection can be produced so that it covers any other space that covers X . We shall not go into any further details and will study instead examples to develop a feeling for the concepts.

The universal covering space for the circle is the line. Let $S^1 = \{z \in \mathbb{C} \mid |z|^2 = 1\}$. As covering projection we take $p: \mathbb{R} \rightarrow S^1$, $p(t) = \exp(it)$, $t \in \mathbb{R}$. For $z \in S^1$, any neighborhood of z less than the entire circle will provide a neighborhood on which p is a homeomorphism (as demanded by the definition). Then $p^{-1}(z) = \{\arg z + 2n\pi, n=0, \pm 1, \dots\}$, and we take $0 < \arg z < 2\pi$ so that there is a preimage for each element of $\pi_1(S^1)$. \mathbb{R} is of course simply connected, and the only paths on S^1 that lift to closed paths on \mathbb{R} are those homotopic to the null path on S^1 . We could also take the circle as a covering space of itself letting $p(z) = z^2$ so that p^{-1} is the two-sheeted root function (rather than the infinitely sheeted log). The mapping of fundamental groups would be the mapping of the integers into the subgroup of even integers.

The universal covering space for the torus is the plane.

For $SO(3)$ the universal covering space is the group manifold of $SU(2)$. The mapping p is part of every physicist's education. For example if $x \in \mathbb{R}^3$, let x correspond (uniquely) to the Hermitian matrix

$$\underline{x} = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3 \quad (23.21)$$

with σ_1 , σ_2 , and σ_3 the Pauli spin matrices. Then for $U \in SU(2)$, the rule for rotating x is

$$\underline{x}' = U\underline{x}U^\dagger$$

If R_{ij} is the 3×3 matrix that accomplishes this same rotation

$$R_{ij} = \frac{1}{2}\text{Tr}(\sigma_i U \sigma_j U^\dagger) \equiv (p(U))_{ij} \quad (23.22)$$

which defines p . Obviously U and $-U$ give the same rotation. $SU(2)$ is simply connected although the way in which preimages of $SO(3)$ are sewn together here is not so simple as for S^1 and \mathbb{R} . A clear picture of the map $p: SU(2) \rightarrow SO(3)$ can be obtained by realizing that $U \in SU(2)$ can be specified by a pair of complex numbers α and β satisfying $|\alpha|^2 + |\beta|^2 = 1$, that is, a point on S^3 . To go to $SO(3)$ identify the points (α, β) and $(-\alpha, -\beta)$ (corresponding to $-U$), that is, antipodal points on the 3-sphere. Thus the two copies of $SO(3)$ in $SU(2)$ can be pictured as the "northern" and "southern" hemispheres of S^3 , but along the equator antipodal points must be identified even within $SO(3)$. This is just projective 3 space and of

course is homeomorphic to the solid ball described earlier as a model of the $SO(3)$ group manifold.

Notes for Section 23.2

References on homotopy theory abound. At an introductory level the following are recommended:

P. J. Hilton, *An Introduction to Homotopy Theory*, Cambridge University Press, London, 1964

J. G. Hocking and G. S. Young, *Topology*, Addison-Wesley, Reading, Mass., 1961

E. M. Patterson, *Topology*, Oliver and Boyd, Edinburgh, 1959

A more advanced text is

E. H. Spanier, *Algebraic Topology*, McGraw-Hill, New York, 1966

Much information on the rotation group can be found in

M. Hamermesh, *Group Theory*, Addison-Wesley, Reading, Mass., 1962

The $SU(2)$ description of rotations is found in Hamermesh's book, and also of course in

H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, Mass., 1950.

It may be worth mentioning that we have pitched our homotopy theory at a very simple and hopefully intuitive level. Unfortunately the reader wishing to penetrate further (say, in Spanier's book) is likely to come up against categories and functors, which are good things but may not be too accessible to the physicist. Many of the theorems stated in the text are true under less stringent conditions than indicated, but generality is at the bottom of our list of priorities, since it is only physically interesting spaces that we have in mind.

23.3 HOMOTOPY APPLIED TO THE PATH INTEGRAL

Schrödinger's equation on a space X takes the form

$$H\psi = i \frac{\partial \psi}{\partial t} \quad H = -\frac{1}{2}\Delta + V(x) \quad (23.23)$$

where Δ is the Laplacian (or Laplace-Beltrami operator), the generalization of $\text{div} \cdot \text{grad}$ for curved spaces (more on this in Section 24). The propagator can be written

$$K(x, t; y, 0) = \sum_{\substack{x(\cdot) \\ x(0)=y \\ x(t)=x}} \exp(iS(x(\cdot))) \quad (23.24)$$

a sum over paths of the “classical action,” where for a Riemannian space the kinetic energy is just the metric (again, we shall be more explicit in the appropriate place).

Now suppose that X is not simply connected. Then as in Section 23.1 it is worth exploring the possibility of grouping the paths entering the sum into those that can be deformed into one another, a collection of sets we now recognize to be the homotopy equivalence classes. Moreover, we again propose that sums over different homotopy classes enter the path integral with coefficients not all unity, allowing additional freedom for the propagator. The proposed form of the propagator is

$$K(x, t; y, 0) = \sum_{\alpha} c_{\alpha} \sum_{\substack{x(\cdot) \in \alpha \\ x(0) = y \\ x(t) = x}} \exp(iS(x(\cdot))) \quad (23.25)$$

where each α represents a homotopy class of paths from y to x . In Section 23.1 the argument for the introduction of $\{c_{\alpha}\}$ was that our proof that the path integral satisfied Schrödinger’s equation was a local one (cf. Section 4). Since (as was claimed) *each* sum of the form

$$K_{\alpha}(x, t; y, 0) \equiv \sum_{\substack{x(\cdot) \in \alpha \\ x(0) = y \\ x(t) = x}} \exp(iS(x(\cdot))) \quad (23.26)$$

satisfies Schrödinger’s equation the local proof does not preclude the appearance of $\{c_{\alpha}\}$. With the perspective of Section 23.2 the justification of our claim concerning K_{α} can be seen as follows.

The space X has a universal covering space B and there is a covering projection p such that $p: B \rightarrow X$ is locally a homeomorphism. Each path in X can be lifted to B once a specific preimage for a single point has been selected. To fix ideas take a specific $y^* \in B$ such that $p(y^*) = y$ = (the initial point in the propagators above). Thus paths in each different α end at a different point in B corresponding to a different preimage of x . (If two ended at the same point, by the simple connectedness of B they could be deformed into each other in B and the homotopy map projected down to X .) Call the different preimages of x , x_{α}^* . Thus $p(x_{\alpha}^*) = x$ for all α and each $x(\cdot) \in \alpha$ is lifted to a path from y^* to x_{α}^* . Now we make a smoothness assumption. Our interest in physical spaces suggests the reasonableness of taking p smooth enough to lift the Laplacian (a purely local object) to B where it has the same form as on X . Thus p lifts the Schrödinger equation

to B and the function K_α of (23.26) is just the propagator

$$K^B(x_\alpha^*, t; y^*, 0) = \sum_{\substack{\text{paths in } B \\ x^*(0)=y^* \\ x^*(t)=x_\alpha^*}} \exp(iS(x^*(\cdot))) \quad (23.27)$$

where S and the Lagrangian have also been lifted and (23.27) is the path integral expression—on B —for the propagator. When $K^B(x_\alpha^*, t; y^*, 0)$ is mapped down to x it is exactly the function K_α . Since K^B satisfies Schrödinger's equation locally on B , K_α satisfies the equation locally on X .

Restrictions on the coefficients $\{c_\alpha\}$ arise from the following considerations. Let $\xi(u), u \in I$ be a closed path beginning and ending at x . The quantity u is not to be thought of as a time variable. As u varies so do all the various propagators entering the sum for K and so do the end points x_α^* of the various paths in B . Thus

$$\begin{aligned} f(u) &\equiv K(\xi(u), t; y, 0) = \sum_\alpha c_\alpha K_\alpha(\xi(u), t; y, 0) \\ &= \sum_\alpha c_\alpha K^B(\xi_\alpha^*(u), t; y^*, 0) \end{aligned} \quad (23.28)$$

where the second equality recognizes the numerical equality of the functions K_α and $K^B(\xi_\alpha^* \dots)$. The point $\xi_\alpha^*(u)$ is well defined and for $u=0$ is just x_α^* . As u varies the change in x is lifted to a change in x_α^* and that position is just ξ_α^* . The most significant aspect of the variation is that while $\xi(1)=x$, $\xi_\alpha^*(1)$ can certainly differ from $\xi_\alpha^*(0)$ and in general will be a different preimage of x from the initial x_α^* . In fact if $\xi_\alpha^*(1)=x_\beta^*$ then the homotopy classes have the relation

$$\beta = [\xi] \alpha \quad (23.29)$$

This is the usual product of homotopy classes and $[\xi]$ is moreover an element of the fundamental group $\pi_1(X, x)$.

The transformation we have just described can have no physical effect, and the object we obtain at $u=1$ must equal that at $u=0$ except possibly for a phase factor φ . The phase factor can depend only on $[\xi] \equiv h$, and we write

$$\begin{aligned} f(1) &= \sum_\alpha c_\alpha K^B(x_\beta^*, t; y^*, 0) = \sum_\beta c_{\beta h^{-1}} K_\beta(x, t; y, 0) \\ &= \exp(-i\varphi_h) f(0) = \sum_\beta \exp(-i\varphi_h) c_\beta K_\beta(x, t; y, 0) \end{aligned} \quad (23.30)$$

The third and fifth expressions in (23.30) must be equal on a term by term basis. To see this, allow y to vary for sufficiently short t . Then one summand in each sum tends toward a delta function, and this pair (one from each sum) must be equal as functions for all t , x , and y (they are propagators between the same points in B). Now let y vary along some nonidentity element of $\pi_1(X, y)$, and another pair of terms in the respective sums must be equal. It follows that for all β and for any particular h

$$c_{\beta h^{-1}} = \exp(-i\varphi_h) c_\beta \quad (23.31)$$

Now perform a second transformation using a path $\eta(u) \in k \in \pi_1(X, x)$. It follows that

$$c_{\beta h^{-1} k^{-1}} = \exp(-i(\varphi_h + \varphi_k)) c_\beta \quad (23.32)$$

for all β . On the other hand, a transformation in the class kh in the first place would have yielded

$$c_{\beta(kh)^{-1}} = \exp(-i\varphi_{hk}) c_\beta \quad (23.33)$$

But successive transformation by h and k must yield the same result as the hk transformation. This implies

$$\varphi_{hk} = \varphi_h + \varphi_k \quad (23.34)$$

Equation 23.34 is our main result. The phases form an additive (hence commutative) representation of the fundamental group. Picking some particular c_α to be unity, any β can be written $\beta = \alpha(\beta^{-1}\alpha)^{-1}$ and

$$c_\beta = \exp(-i\varphi_{\beta^{-1}\alpha}) = \exp(i\varphi_{\alpha^{-1}\beta}) \quad (23.35)$$

The final form of the propagator is therefore

$$K(x, t; y, 0) = \sum_\beta \exp(i\varphi_{\alpha^{-1}\beta}) K_\beta(x, t; y, 0) \quad (23.36)$$

where it is understood that α is a fixed homotopy class of paths from y to x and each term K_β is the partial sum of $\exp(iS)$ over paths in the homotopy class of β .

(In terms of the path integral on the circle of Section 23.1 the quantity $\varphi_{\alpha^{-1}\beta}$ is $n\delta$ where the winding number label " n " corresponds to $\alpha^{-1}\beta$, since the fundamental homotopy group for the circle is the additive group of the integers \mathbb{Z} .)

There is still a bit more machinery needed before the path integral on $SO(3)$ can be undertaken. In particular the curvature of $SO(3)$ will require separate consideration. However, there is an ingenious flat space application of (23.34) given by Laidlaw and DeWitt which we now present.

The usual way of dealing with N indistinguishable particles is to define the coordinate space as \mathbf{R}^{3N} and then use the fact that the Hamiltonian commutes with the exchange operator to restrict attention to representations of the symmetric group on N objects (S_N). This approach has both an aesthetic and a practical disadvantage. First, the indistinguishability is handled on an *ad hoc* basis. From the formalism it still seems to be meaningful to ascribe to electron A position \mathbf{r}_A and to B , \mathbf{r}_B . Second, there are many representations of S_N that do not appear to be realized for elementary particles, and some effort is needed to get rid of these *parastatistics* and restrict attention to fermions and bosons alone.

Laidlaw and DeWitt said not to look at \mathbf{R}^{3N} , but rather at the quotient space gotten from \mathbf{R}^{3N} by identifying N -tuples which can be reached from one another through the exchange of single vectors. To make a formal statement of their result one (seemingly) technical point intrudes itself. Write points in \mathbf{R}^{3N} as

$$y = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \quad \text{each } \mathbf{x}_i \text{ a 3 vector.}$$

From \mathbf{R}^{3N} we exclude y 's in which any pair of vectors \mathbf{x}_i and \mathbf{x}_j coincide. The correct coordinate space for N indistinguishable particles is therefore

$$X = [\mathbf{R}^{3N} - \{\text{coincidence points}\}] / S_N$$

where “ $/S_N$ ” means modulo the symmetric group, that is, N -tuples y which are permutations of each other are identified. It follows from various theorems of homotopy theory (which we do not go into in detail here) that the fundamental group of X is just S_N . But now the requirement of commutativeness for the representations comes into play when we construct a propagator for systems on X . In particular S_N has only two commutative representations:

$$D^B(g) = +1 \quad \text{all } g \in S_N$$

$$D^F(g) = [\text{sign of the permutation } g]$$

These correspond respectively to bosons and fermions, and the commutativity requirement excludes parastatistics.

Notes for Section 23.3

Implicit use of the main result of this section (equation 23.34) was made in

L. S. Schulman, A Path Integral for Spin, *Phys. Rev.* **176**, 1558 (1968)

but a statement in terms of representations of the fundamental homotopy group did not appear until the following two papers:

L. S. Schulman, Approximate Topologies, *J. Math. Phys.* **12**, 304 (1971)

M. G. G. Laidlaw and C. M. DeWitt, Feynman Functional Integrals for Systems of Indistinguishable Particles, *Phys. Rev. D* **3**, 1275 (1971)

Laidlaw and DeWitt gave a detailed proof of the result, as well as the application to parastatistics described in text. Another proof is

J. S. Dowker, Quantum Mechanics and Field Theory on Multiply Connected and on Homogeneous Spaces, *J. Phys. A* **5**, 936 (1972)

Regarding the requirement for Abelian representations, note that for n -component spinors K is an $n \times n$ matrix and (23.30) generalizes to $f(1) = Uf(0)$ with U a unitary $n \times n$ matrix. As far as I can tell the U s yield a unitary but not necessarily Abelian representation of π_1 . For indistinguishable particles with spin there would still be the possibility of parastatistics.

Quantization with the aid of homotopy theory, as described in this section, has been related to the nonintegrable phase factors of Wu and Yang and to the "prequantising" of Kostant and Souriau. This was done by

P. A. Horvathy, Quantisation in Multiply Connected Spaces, *Phys. Lett.* **76A**, 11 (1980)

Nonintegrable phase factors and gauge fields are studied in

T. T. Wu and C. N. Yang, *Phys. Rev. D* **12**, 3845 (1975)

These authors are concerned with characterizing gauge fields such as the electromagnetic field. They observe that in view of the Aharonov-Bohm effect (Section 23.1) the field strengths \mathbf{E} and \mathbf{B} alone are insufficient. However, the message to be gotten from Aharonov and Bohm is *not* that $\oint \mathbf{A} \cdot d\mathbf{x}$ (around all possible curves) is the desired characterization of magnetic fields, but rather that $\exp[i\oint e\mathbf{A} \cdot d\mathbf{x}/\hbar c]$, that is, the phase factor alone, is that characterization. They bring their point home in terms of gauge transformations, but we can state the arguments using the formalism of path integrals. In Section 23.1 we found that cutting a tube out of \mathbb{R}^3 by removing the solenoid of the Aharonov-Bohm experiment introduced an ambiguous quantity δ that enters the propagator as $(e^{i\delta})^n$. The quantity δ was identified as $\delta = \oint e\mathbf{A} \cdot d\mathbf{x}/\hbar c$, and we see that the physical effect depends on $e^{i\delta}$ only. Thus fields for which δ is an integral multiple of 2π are equivalent to no field at all.

This suggests that if one were sufficiently perverse he could derive the homotopy factors for propagators by looking at gauge transformations

alone. Such perversity might be due to a reluctance to accept the "sum over paths" picture and a preference for seeing the path integral as (merely?) a poetic restatement of the Trotter product formula. In any case, gauge transformations can be used to obtain the homotopy factor for propagators because on multiply connected spaces it is possible to define smooth many valued functions. Such a derivation of the homotopy factor was given in

L. S. Schulman, Caustics and Multivaluedness: Two Results of Adding Path Amplitudes, in A. M. Arthurs, Ed., *Functional Integration and Its Applications*, (Proc. of London Conference, 1974) Oxford University Press, London, 1975

In field theory quantization the homotopy phase factor has also been used. See

N. K. Pak and H. C. Tze, Chiral Solitons and Current Algebra, *Ann. Phys.* **117**, 164 (1979)

23.4 EXTENSIONS OF SYMMETRIC OPERATORS

Since functional analysis is likely to be in the repertoire of the theoretical physicist, we do not provide an extensive background (as was done for homotopy theory), but only the minimum needed to develop a broader view of what was done in Section 23.3.

For simplicity restrict attention to the Hamiltonian

$$H = -\frac{1}{2} \frac{d^2}{dx^2}$$

on the circle S^1 . From the standpoint of a theory of square integrable functions on S^1 , $L^2(S^1)$ cannot be distinguished from $L^2(I)$, $I=[0, 1]$. The ambiguity encountered earlier in the path integral is reflected here in that while H is symmetric (say for functions vanishing at both 0 and 1) it is not self adjoint until boundary conditions are given. In particular H is not essentially self adjoint in that many self adjoint extensions are possible and with them different spectra for H occur.

Because the path integral is a method of producing the unitary operator $\exp(-itH)$ (which thus has a self adjoint generator) it necessarily extends H . Referring to earlier work it is seen that the domain of H consists of functions satisfying

$$f(1) = e^{i\varphi} f(0)$$

each φ defining a particular extension of H . However, using the general theory of extensions of symmetric operators one can find the deficiency indices of H to be (2, 2) so that the class of possible extensions is much larger than the one-parameter family given above.

As an example, the condition $f(1)=f'(0)$, $f'(1)=-f(0)$ defines a domain for H that makes it self adjoint. The spectrum of this operator is easy to calculate, and it is certainly different from the set $\{\frac{1}{2}(\varphi+2n\pi)^2 | n=0, \pm 1, \pm 2, \dots\}$ that one obtains from the path integral extensions.

The path integral extension thus uses more information about the space and the operator than just the Hilbert space structure. Some symmetry properties of the underlying coordinate space are retained through demands on properties of the propagator under rotations. In particular, the path integral extension appears to demand good behavior of the square root of H , $-id/dx$ which is the generator of translations.

Occasionally path integrals can be used to extend an operator even where no homotopy considerations are present. In studying diffraction by a smooth body, Buslaev found it useful to consider two copies of the space outside the body and let his paths go from one copy to the other. The relative sign with which direct and indirect contributions are taken determines the boundary conditions at the surface of the diffracting body, whether Dirichlet or Neumann. Similar considerations arise in the method of images.

Notes for Section 23.4

Extensions of symmetric operators, deficiency indices, and so on, are treated in

F. Riesz and B. Sz-Nagy, *Functional Analysis*, Ungar, New York, 1955

L. A. Lusternik and V. J. Sobolev, *Elements of Functional Analysis*, Wiley, New York, 1974 and many other books. The work of Buslaev is

V. S. Buslaev, Continuum Integrals and the Asymptotic Behavior of the Solutions of Parabolic Equations as $t \rightarrow 0$. Applications to Diffraction, in *Topics in Mathematical Physics*, Sh. Birman, Ed., Consultants Bureau, New York, 1967, p. 67

One of the topics treated in the paper of Wu and Yang (quoted in the notes to Section 23.3) suggests an interesting consequence for functional analysis. They use their approach to deduce the quantization of the magnetic charge of the Dirac magnetic monopole. The argument runs as follows: a magnetic monopole has a field

$$\mathbf{B} = g \frac{\mathbf{r}}{r^3}$$

with g the magnetic charge. A vector potential that gives rise to this field (through $\mathbf{B} = \nabla \times \mathbf{A}$) is

$$\mathbf{A} = \frac{g(1 - \cos \theta)}{r \sin \theta} \hat{\phi} \quad (23.37)$$

where $\hat{\phi}$ is the unit vector in the φ direction in spherical coordinates. This field \mathbf{A} is singular on the line $\theta=\pi$, often referred to as a “string” in this context. By a gauge transformation the string can be sent elsewhere, and an example of an alternative field is

$$\tilde{\mathbf{A}} = \frac{g(-1 - \cos \theta)}{r \sin \theta} \hat{\phi}$$

which has its string along $\theta=0$. The gauge transformation relating these is

$$\chi = 2g\varphi \quad \mathbf{A} - \tilde{\mathbf{A}} = \nabla \chi$$

where φ is the coordinate function. One expects the position of the string to be arbitrary because the field \mathbf{B} contained no reference to a special string direction. Now, argue Wu and Yang, when we perform the above gauge transformation the wave function of an electron in the field is multiplied (as in all gauge transformation—see Section 5) by $\exp(ie\chi/\hbar c)$. If one insists on single valued wave functions both before and after the gauge transformation then the phase factor must be unity for φ any integer multiple of 2π . Thus

$$\frac{2ge}{\hbar c} = \text{integer} \quad (23.38)$$

which is Dirac's quantization conditions for g . (As in the notes to Section 23.3 an interesting exercise is to put this argument in path integral language.)

The functional analysis question that I pose is, how does this look in Hilbert space? That is, you are handed the Hamiltonian $H=(\mathbf{p}-e\mathbf{A}/c)^2/2m$, with the \mathbf{A} of (23.37) above. You do not ask about the physics (i.e., the fact that the derived \mathbf{B} is spherically symmetric), and you eschew gauge transformations—all you notice is that \mathbf{A} is singular along the line $\theta=\pi$. You do not even demand single valuedness of wave functions since in principle this is no worse than talking about a discontinuity and on Hilbert space your only requirement is square integrability. The presumption, based on the discussion of this section, is that H or some operator related to it such as $\mathbf{p}-e\mathbf{A}/c$ will not be self adjoint unless the Dirac quantization condition (23.38) is satisfied.

TWENTY-FOUR

Quantum Mechanics on Curved Spaces

If you like excitement, conflict, and controversy, especially when nothing very serious is at stake, then you will love the history of quantization on curved spaces. Path integration represents only the most recent technique for approaching this problem. Which is not to say the issues are trivial—they are closely related to the operator ordering problems of quantum mechanics and like them do not yield easy, or unique, solutions. My own opinion on the source of all the confusion is that there is no physics at stake: no experimental result has ever been shown to depend on the details of curved space quantization. This is why people (myself included) continue to get signs and factors of 2 wrong in their results and why from time to time new papers appear rediscovering (up to factors and signs) old formulas that have in the meantime been forgotten. The fundamentalist view expressed here on the nature of physics is not in accord with my personal predilections, but the observed facts on the history of the topic at hand unfortunately support that view.

Consider a manifold M with coordinates $q^i, i=1, \dots, n$ defined on it and with a Riemannian metric g_{ij} . If the metric is assumed to define the kinetic energy and if we allow for potentials (and vector potentials) the Lagrangian for motion on the manifold is

$$L = \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j + a_i \dot{q}^i - v \quad (24.1)$$

where the summation convention is adopted, a_i is the vector potential, v is the potential, and dots denote differentiation with respect to t . The dots could as well refer to some other path parameter, such as proper time or the “fifth parameter” of Fock and Feynman, and t could be included

among the $\{q_i\}$. This in fact is the usual situation for relativistic applications.

The Hamiltonian associated with L is well defined in the context of classical mechanics and is

$$H = \frac{1}{2}g^{ij}(p_i - a_i)(p_j - a_j) + v \quad (24.2)$$

where g^{ij} is the inverse of g_{ij} ($g^{ij}g_{jk} = \delta_k^j$) and

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = g_{ij}\dot{q}^j + a_i \quad (24.3)$$

The source of the ambiguity in the quantization of L or H can be described in various ways. The metric g_{ij} is a function of q . Generally speaking, p will go over to $\partial/\partial q$, so that an expression $g(p-a)(p-a)$ already has ordering problems. For path integral quantization the dilemma is best understood by recalling Section 4 and the quantization of vector potentials. The product

$$a(q)\dot{q} dt \quad (24.4)$$

will give different results, depending on whether it is taken to be $a(q(t_i))\Delta q$, $a(q(t_{i+1}))\Delta q$, $a(q(\frac{1}{2}t_i + \frac{1}{2}t_{i+1}))\Delta q$, or some other value (a mesh t_1, \dots, t_N for the path integral is assumed). The difference between these various objects is of the order of $a'(q)(\Delta q)^2$, which in turn is of order $a'(q)\Delta t$ in the path integral. This issue and its relation to the Ito integral is discussed fully in Section 5. For the Lagrangian (24.1) the problem is even more acute as g_{ij} multiplies a term of order $(\Delta q)^2/\Delta t$, which is of order unity, so the point at which g is evaluated will certainly affect the final form of the propagator.

Although we do not go into operator ordering questions at all it is worth observing that it is at this stage—in deciding where to evaluate functions—that the same ambiguities enter path integral quantization. There is no free lunch.

Our only prejudice regarding the form of the quantized Hamiltonian is simplicity, and we require that in the absence of potentials

$$H = -\frac{1}{2}\Delta = -\frac{1}{2}g^{-1/2}\frac{\partial}{\partial q^i}\left(g^{ij}g^{1/2}\frac{\partial}{\partial q^j}\right) \quad (24.5)$$

with $g = \det g_{ij}$. Again, although our primary interest here is not canonical

quantization, it is worth mentioning that the canonical operators take the form

$$\hat{q}^i = q^i \text{ (i.e., multiplication)}, \quad \hat{p}_i = -i \left(\frac{\partial}{\partial q^i} + \frac{1}{2} \frac{\partial}{\partial q^i} (\log g) \right) \quad (24.6)$$

so that for H to be merely the Laplacian (when $v=0, a_i=0$) it must be

$$\hat{H} = \frac{1}{2} g^{-1/4} \hat{p}_i g^{ij} g^{1/2} \hat{p}_j g^{-1/4} = \frac{1}{2} \hat{p}_i g^{ij} \hat{p}_j - \frac{1}{2} g^{1/4} \Delta g^{-1/4} \quad (24.7)$$

The basic object with which one works in path integration is the propagator for infinitesimal times; if this is right, then by iteration (or by whatever way it is you prefer to say that

$$\lim_{n \rightarrow \infty} \left[1 + \frac{(x+y_n)}{n} \right]^n = e^x \quad \text{if } \lim y_n = 0 \quad) \quad (24.8)$$

one can get to the finite time propagator. It was infinitesimal time propagation that we used in Section 4 to check that path integration recovered the Schrödinger equation. It is infinitesimal time propagation that must be used here to check any putative propagator. We seek then a $K(q'', t''; q', t')$ such that

$$H'' K = i\hbar \frac{\partial K}{\partial t} \quad \text{for } t'' - t' \rightarrow 0 \quad (24.9)$$

Our first guess is motivated by the fact that $\hbar \rightarrow 0$ and $|t'' - t'| \rightarrow 0$ limits give essentially the same approximation (of course, this lies behind the form of the infinitesimal propagator that we have been using throughout). One might try

$$K = D^{1/2} e^{iS/\hbar} \quad (24.10)$$

where S is the classical action for the path from (q', t') to (q'', t'') and D is an appropriate scalar built from the second derivatives of S (see (12.46)).

The main oddity of curved space path integration is that the guess (24.10) is wrong. B. DeWitt discovered that you needed an additional factor $\exp[i(\hbar^2 R/12)t/\hbar]$ where R is the curvature of M (to be defined

shortly).* Aside from the novelty of finding a disagreement between the limits $\Delta t \rightarrow 0$ and $\hbar \rightarrow 0$, one wants to know, does this matter? Is this term experimentally observable? The question came up in a conference in 1957, and in a discussion among DeWitt, Wheeler, and (I believe) Feynman no hope was offered for experimental verification. For a space of constant curvature there is clearly no effect, as the term $\hbar^2 R$ just leads to a shift of all energy levels and an overall irrelevant phase factor. One might have hoped that certain constrained coordinates—perhaps a relative coordinate in a molecule—would effectively live on a space of variable curvature, but the participants in the discussion estimated the uncertainties due to the constraint as yielding greater energies than the effect sought. Perhaps gravitation will give us the system on which $R\hbar^2$ can be seen, but at present the effect is purely in the realm of theory.

Having denied the validity of (24.10) and explained that the correct K is obtained by applying to it $H - i\hbar\partial/\partial t$ and looking at first-order terms in Δt , we now present the correct K but will not go through the tedious process of verification.

For the Lagrangian (24.1), let $S_c(q'', t''; q', t')$ be the exact classical action between the indicated endpoints (i.e., $\int L dt$ along the classical path; the path is assumed unique for sufficiently small $t'' - t'$). Define the van Vleck determinant as (dropping the subscript c on S)

$$D_{ij} = - \frac{\partial^2 S}{\partial q''^i \partial q'^j}$$

$$D = D(q'', t''; q', t') = \det D_{ij} \quad (24.11)$$

The curvature is built from the metric in the usual way:

$$[jk, i] = \frac{1}{2}(g_{ij, k} + g_{ik, j} - g_{jk, i})$$

$$R_{ikjl} = \frac{1}{2}(g_{ij, kl} - g_{il, kj} - g_{kj, il} + g_{kl, ij})$$

$$+ g^{mn}([ij, m][kl, n] - [kj, m][il, n])$$

$$R_{ij} = -g^{kl}R_{ikjl}, \quad R = g^{ij}R_{ij} \quad (24.12)$$

*A note of caution is in order here on factors $R/12$, $R/6$, and signs assigned thereto. One often writes the propagator K without the $D^{1/2}$ that appears in (24.10), and then $R/6$ must be used instead of $R/12$. This is because on n -dimensional curved space

$$D(q'', t''; q', t') = (\Delta t)^{-n} g''^{1/2} g'^{1/2} \left[1 + \frac{1}{6} R'_{ij} \Delta q^i \Delta q^j \right]$$

+ terms that drop out in the functional integral

where we have used notation defined below. Thus the $R_{ij}/6$ in D also contributes a term $R/12$. There is also the possibility that for some reason it is preferable to use a Hamiltonian different from (24.5) (perhaps by virtue of some canonical definition), and then additional pieces of $\pm R/12$ can easily creep in.

In terms of these quantities the propagator is given by

$$K(q'', t''; q', t') = (2\pi i \hbar)^{-n/2} g''^{-1/4} D^{1/2}(q'', t''; q', t') g'^{-1/4} \\ \times e^{i\hbar R t'/12} \exp[iS(q'', t''; q', t')/\hbar] \quad (24.13)$$

with $g'' = g(q'')$, $g' = g(q')$, $t = t'' - t'$. For this propagator the volume element for integration is

$$g(q)^{1/2} dq^1 \cdots dq^n \quad (24.14)$$

As for the flat space case, S can be expanded as a function of coordinate differences

$$\Delta^i \equiv q''^i - q'^i, \quad \Delta t = t'' - t' \quad (24.15)$$

but the details of the expansion depend on where $g_{ij}(q)$ is evaluated. S is simplest for the “midpoint” $q = (q'' + q')/2$, $t = (t'' + t')/2$ so that g_{ij} and functions thereof in the following equation are understood to be evaluated at the “midpoint.” The approximation to the classical action is

$$S(q'', t''; q', t') = \frac{1}{\Delta t} \left\{ \frac{1}{2} g_{ij} \Delta^i \Delta^j + \frac{1}{48} (g_{ij,kl} - 2g^{mn}[ij, m][kl, n]) \Delta^i \Delta^j \Delta^k \Delta^l \right\} \\ + a_i \Delta^i - v \Delta t \quad (24.16)$$

with errors of order $\Delta^5/\Delta t$, Δ^3 , $\Delta \cdot \Delta t$, and $(\Delta t)^2$.

At long last we have all the tools needed for the path integral for spin, that is, for the path integral of a system whose external position coordinate moves in \mathbf{R}^3 and whose internal spin coordinate moves on $SO(3)$.

We shall consider only free motion of the spin coordinate (no external magnetic field) so the \mathbf{R}^3 motion is uncoupled and can be handled as a function of initial and final points for short times. Actually it is easier to obtain S for finite times and not to have to resort to (24.16) at all.

Following the approach of Section 23, the path integral is done on the covering space of $SO(3)$, namely $SU(2)$. Afterwards we take appropriate sums of $SU(2)$ propagators. The choice of Lagrangian on $SU(2)$ is governed by simplicity. Thinking in terms of the top, the simplest is the spherical top; in group theoretical terms simplicity suggests the natural metric on $SU(2)$, that which is invariant under both left and right translations. These

notions coincide and the Lagrangian

$$\left(\frac{1}{I}\right)L = \frac{1}{2}(\dot{\theta}^2 + \dot{\varphi}^2 + \dot{\psi}^2 + 2\cos\theta\dot{\varphi}\dot{\psi}) \quad (24.17)$$

(θ , φ , and ψ standard Euler angles, I a “moment of inertia”) gives rise to an invariant distance

$$\Gamma = 2\cos^{-1}\left(\frac{1}{2}\text{Tr}U_2U_1^{-1}\right) \quad (24.18)$$

Here $U_1, U_2 \in SU(2)$. This assertion is easily verified using the relation of Euler angles to $SU(2)$ matrix elements

$$U = \exp\left(\frac{-i\varphi\sigma_3}{2}\right)\exp\left(\frac{-i\theta\sigma_2}{2}\right)\exp\left(\frac{-i\psi\sigma_3}{2}\right) \quad (24.19)$$

(σ_i are the Pauli matrices) and the fact that geodesics in the natural metric are translations of the one parameter subgroups:

$$U(t) = \exp\left(-it\omega\hat{n}\cdot\frac{\sigma}{2}\right)U(0) \quad (24.20)$$

(the constant ω is inserted for later convenience). For the spherical top, free motion consists of rotation about a fixed axis with magnitude and direction fixed by the boundary conditions. By rotational invariance we can consider motion about the z axis, beginning at φ_1 and ending at φ_2 ($\theta=0$ throughout), taking time T ; with $(1/I)L(dt)^2$ taken to be the metric this obviously gives

$$\text{distance} = \varphi_2 - \varphi_1 + 4n\pi \quad (24.21)$$

where n is the number of whole revolutions (these boundary conditions yield an infinity of classical paths for each pair of end points; the 4 multiplying the $n\pi$ occurs because on $SU(2)$ φ must increase by 4π for U to return to its original value, and the boundary conditions are actually given in terms of U , not φ). The quantity Γ for the same boundary conditions satisfies

$$\begin{aligned} \cos\frac{1}{2}\Gamma &= \frac{1}{2}\text{Tr}\left[U_2U_1^{-1}\right] = \frac{1}{2}\text{Tr}\exp\left[-i\frac{(\varphi_2-\varphi_1)\sigma_3}{2}\right] \\ &= \cos\left[\frac{\varphi_2-\varphi_1}{2}\right] \end{aligned} \quad (24.22)$$

which proves our assertion on the relation of (24.17) and (24.18). The classical action is just $S = \int_0^T L dt$, and for boundary conditions $U(0)=U_1$, $U(T)=U_2$ we have

$$S = \frac{I\Gamma^2}{2T} \quad (24.23)$$

with Γ given by (24.18). The expression relating Γ to the Euler angles is

$$\begin{aligned} \cos \frac{1}{2}\Gamma &= \cos \frac{1}{2}(\theta_2 - \theta_1) \cos \frac{1}{2}(\varphi_2 - \varphi_1) \cos \frac{1}{2}(\psi_2 - \psi_1) \\ &\quad - \cos \frac{1}{2}(\theta_2 + \theta_1) \sin \frac{1}{2}(\varphi_2 - \varphi_1) \sin \frac{1}{2}(\psi_2 - \psi_1) \end{aligned} \quad (24.24)$$

The factor $(g')^{-1/4} D^{1/2} (g'')^{-1/4}$ can be calculated in a straightforward way from S by taking appropriate derivatives. Such a straightforward derivation ignores the geometric significance of this quantity and obscures the generalization to path integrals on other spaces; more will be made of this issue later, but for now it is enough to realize that even if one is unaware of the more profound features of $g'^{-1/4} D^{1/2} g''^{-1/4}$ it is still possible to evaluate it. Finally, the curvature R can also be obtained by straightforward calculation and is $\frac{3}{2}$. Assembling all the pieces that go into the propagator we have (an I^{-1} must multiply the curvature when L is proportional to I)

$$K = \left(\frac{I}{2\pi i\hbar T} \right)^{3/2} \frac{\Gamma}{2 \sin \frac{1}{2}\Gamma} \exp \left(\frac{i\hbar T}{8I} \right) \exp \left(\frac{iI\Gamma^2}{2\hbar T} \right) \quad (24.25)$$

K is here the propagator on $SU(2)$. Its arguments are U_1 and U_2 at $t=0$ and $t=T$ respectively and Γ is given by (24.18). There are of course many solutions to (24.18), and for $T \rightarrow 0$ the one that should appear in (24.25) is the smallest in absolute value so that $K \rightarrow \delta$ -function. This way of specifying Γ can be inconvenient, since as the final point moves on $SU(2)$ it may be necessary to switch the selected Γ . In principle we are only interested in the $T \rightarrow 0$ propagator. In this limit all solutions Γ of (24.18) except the smallest are insignificant. There is then no harm in taking for K the sum of all these, and one thereby eliminates bookkeeping problems as different values of $U(T)$ are considered. The $SU(2)$ propagator is thus taken to be

$$K_{SU(2)}(U_2, T; U_1, 0) = \left(\frac{I}{2\pi i\hbar T} \right)^{3/2} \sum_{n=-\infty}^{\infty} \frac{\Gamma + 4n\pi}{2 \sin \frac{1}{2}\Gamma} \exp \left\{ \frac{i\hbar T}{8I} + \frac{iI(\Gamma + 4n\pi)^2}{2\hbar T} \right\} \quad (24.26)$$

where Γ is some definite (but arbitrary) solution of (24.18). There is no question of relative phases entering the sum here, as $SU(2)$ is simply connected. The $SO(3)$ propagator is obtained according to the prescription of Section 23. The fundamental group is \mathbb{Z}_2 (the unique group of two elements), and its two representations are the trivial representation and the faithful representation. There are therefore two possible $SO(3)$ propagators. In expressing these we shall not stick to the notation of Section 23, which would here produce a nightmare of subscripts and asterisks. The $SO(3)$ propagator will have as its arguments $SU(2)$ elements (which can cause no ambiguity) and the two preimages of the covering projection will be written $+U$ and $-U$. Bringing a (\pm) superscript to designate the two possible propagators we have

$$K_{SO(3)}^{(\mp)}(U_2, T; U_1, 0) = K_{SU(2)}(U_2, T; U_1, 0) \pm K_{SU(2)}(-U_2, T; U_1, 0) \quad (24.27)$$

From (24.26) and the definition of Γ this becomes

$$K_{SO(3)}^{(\mp)} = \left(\frac{I}{2\pi i\hbar T} \right)^{3/2} \exp\left(\frac{i\hbar T}{8I} \right) \sum_{n=-\infty}^{\infty} \frac{\Gamma + 2n\pi}{2 \sin \frac{1}{2}\Gamma} (\pm)^n \exp\left(\frac{iI(\Gamma + 2n\pi)^2}{2\hbar T} \right) \quad (24.28)$$

(the reason for the inversion of signs in the definition of K in (24.27) is now evident: the factor $\sin \frac{1}{2}\Gamma$ also changes sign as Γ increases by 2π).

It is of some interest to compare (24.28) to the propagator computed by conventional Schrödinger equation techniques. The surprising result (which we next verify) is that $K_{SO(3)}$ of (24.28) is not only correct for $T \rightarrow 0$ but is *exact* for all T . Geodesic motion on $SO(3)$ is in this sense as “simple” as free motion on \mathbb{R}^3 .

Calculation of the propagator makes use of (7.6), involving the eigenfunctions of the Schrödinger equation on $SO(3)$. Because it is essentially the Laplacian on the group, the wave functions are just the representation matrices. Again using $SU(2)$ elements for coordinates, the wave functions are labeled by j, m, k ; $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$, $-j \leq m, k \leq j$. The normalized eigenstates and eigenvalues are

$$\left(\frac{2j+1}{8\pi^2} \right)^{1/2} D_{mk}^{j*}(U), \quad E_{jmk} = \frac{\hbar^2}{2I} j(j+1) \quad (24.29)$$

The Green's function is obtained by summing either all integral or all half integral j : if *all* j were included in a single sum a 2π rotation would change

the relative phase of integral and half integral spin wave functions causing a physical change (this is a superselection rule). The two possible Green's functions are

$$G^{(\pm)}(U_2, T; U_1, 0) = \sum_{j, m, k} \frac{2j+1}{8\pi^2} D_{mk}^{j*}(U_2) D_{mk}^j(U_1) \exp\left(-\frac{i\hbar T}{2I} j(j+1)\right) \quad (24.30)$$

with the upper sign being the sum over integral j and the lower sign the sum over half integral j . The fact that D is a unitary representation of the group plays an important simplifying role, and for any j we have

$$\begin{aligned} \sum_{mk} D_{mk}^{j*}(U_2) D_{mk}^j(U_1) &= \sum_{mk} D_{mk}^{j*}(U_2) D_{km}^{j*}(U_1^{-1}) \\ &= \sum_m D_{mm}^{j*}(U_2 U_1^{-1}) = \text{Tr} D^{j*}(U_2 U_1^{-1}). \end{aligned} \quad (24.31)$$

Since the trace is invariant under similarity transform we can take D^j diagonal. This will be the case if the z axis is taken along the direction \hat{n} defined by U_2 and U_1 through (24.20). In that case $D_{mm}^j(U_2 U_1^{-1}) = \exp(im\Gamma)$, with the Γ of (24.18). It follows that

$$\text{Tr} D^{j*}(U_2 U_1^{-1}) = \frac{\sin(j + \frac{1}{2})\Gamma}{\sin \frac{1}{2}\Gamma} \quad (24.32)$$

and

$$\begin{aligned} G^{(\pm)} &= \frac{-2}{\sin(\frac{1}{2}\Gamma)} \frac{1}{8\pi^2} \frac{\partial}{\partial\Gamma} \sum_j \cos\left[\left(j + \frac{1}{2}\right)\Gamma\right] \exp\left(-\frac{i\hbar T}{2I} j(j+1)\right) \\ &\equiv \frac{-2}{\sin(\frac{1}{2}\Gamma)} \frac{1}{8\pi^2} \frac{\partial}{\partial\Gamma} g_{\pm}, \quad g_+ \text{ is for } j=0, 1, \dots \\ &\qquad\qquad\qquad g_- \text{ is for } j=\frac{1}{2}, \frac{3}{2}, \dots \end{aligned} \quad (24.33)$$

where

$$\begin{aligned} g_+ &= \frac{1}{2} \theta_2\left(\frac{1}{2}\Gamma, -\frac{c}{\pi}\right) e^{ic/4} = \frac{1}{2} e^{i\Gamma/2} \theta_3\left(\frac{1}{2}\Gamma - \frac{1}{2}c, -\frac{c}{\pi}\right), \\ g_- &= \frac{1}{2} e^{ic/4} \left[\theta_3\left(\frac{1}{2}\Gamma, -\frac{c}{\pi}\right) - 1 \right] \end{aligned} \quad (24.34)$$

and with $c = \hbar T/2I$. The functions g_{\pm} are Jacobi theta functions, evoking a strong parallel to the much simpler case of a free particle in S^1 .

In fact the parallel extends beyond formal resemblance. The purpose of the foregoing exercise [beginning after (24.28)] was to compare the path integral propagator [(24.26) or (24.28)] to the conventional one [(24.34)]. $K_{SO(3)}^{(\pm)}$ are easily expressed as theta functions, and their equality with $G^{(\pm)}$ follows from the “fundamental theorem for theta functions” [(23.13)] which was also the tie for the free particle.

This nearly completes the task of producing a path integral for spin. The usual spin theory writes an evolution equation for a particle having definite spin. Our propagator on the other hand is applied to a state having various spins, either all integral or half integral. From our theory the usual formalism can be recovered simply by projection.

In an external magnetic field the classical Lagrangian will contain a coupling of field to angular velocity and ultimately to a propagator which can couple different m values [cf. (24.29)] but not different total spins. We do not go into a detailed verification of this assertion but rely instead on the work of Bopp and Haag and Rosen: they did a conventional quantization of the spherical top including the case of an external field.

Our results on $SO(3)$ quantization have led to two suggestive results. First, although $K_{SO(3)}$ of (24.28) was put forward as valid only for $|t'' - t'| \rightarrow 0$ it turned out, by comparison with G , to be *exact*. Second, the propagator itself is simply related to the Jacobi theta function, and the passage from the energy expansion (G) to the path expansion (K) is via the fundamental theorem for theta functions. A reasonable conjecture based on the similarity to the free particle on S^1 is that somehow geodesic motion with the natural metric on the group manifold is “free” and $SO(3)$ ought not be unique in this respect.

It turns out that the sum over classical paths (the analogue of “ K ”) is exact for a large class of spaces including the group manifolds of simple Lie groups. Moreover, the theta function, now in an easily generalized form as a function of several variables, is again simply related to the propagator. These remarkable results were proved by Dowker. He identified the quantity $g'^{1/2} D^{-1} g''^{1/2}$, a function of which appears in the propagator (24.13), as Ruse’s invariant. He further related this to the characters of the representations of the group (on whose group manifold we are interested in path integrating), and by tying together various other strands on group representations and their relation to partial differential equations was able to arrive at his conclusions.

NOTES

The following is a partial bibliography of curved space path integration:

Wolfgang Pauli, *Pauli Lectures on Physics, Vol. 6, Selected Topics in Field Quantization*, C.P. Enz, Ed., MIT Press, Cambridge, Mass., 1973, Ch. 7

- C. Morette (Morette-DeWitt), *Phys. Rev.* **81**, 848 (1951)
- B. S. DeWitt, *Rev. Mod. Phys.* **29**, 377 (1957)
- S. F. Edwards and Y. V. Gulyaev, Path Integrals in Polar Coordinates, *Proc. Roy. Soc. Lond. A* **279**, 229 (1964)
- L. S. Schulman, A Path Integral for Spin, *Phys. Rev.* **176**, 1558 (1968)
- A. M. Arthurs, *Proc. Roy. Soc. Lond. A* **318**, 523 (1970)
- D. W. McLaughlin and L. S. Schulman, *J. Math. Phys.* **12**, 2520 (1971)
- J. S. Dowker, Covariant Schrödinger Equations, in *Functional Integration and Its Applications*, A. M. Arthurs, Ed., Oxford, 1975

Additional papers on curved space path integration are legion, and we do not attempt to list them all. Some are correct; some are less correct. Some have original features; some are less rich in this praiseworthy property.

The demonstration of the exactness of the sum over classical paths for the $SO(3)$ propagator, as well as a conjecture that this property should hold for Lie groups in general appears in my Ph.D. thesis (Princeton University, 1967), much of which appears in *Phys. Rev.* **176**, 1558 (1968). Dowker's proof of the exactness of the classical paths as well as the evaluation of the propagator in terms of theta functions appear in

- J. S. Dowker, *J. Phys. A* **3**, 451 (1970)
 J. S. Dowker, *Ann. Phys.* **62**, 361 (1971)

In the past three sections we have obtained the propagator K on multiply connected spaces by projecting from a covering space— K is a sum of terms, one from each copy of the physical space and they are combined so as to provide particular boundary conditions. A similar picture, albeit as an approximation, is possible for the WKB propagator when several classical paths are available. Labeling the paths with α , we have, as in (3.7),

$$G_{WKB}(x, t; y) \sim \sum_{\alpha} \sqrt{\frac{\partial^2 S_{\alpha}}{\partial x \partial y}} \exp[iS_{\alpha}(x, t; y)/\hbar].$$
 Generally the functions $S_{\alpha}(x, t; y)$, which are solutions of the Hamilton-Jacobi equation, are branches of a single multivalued function whose branch points occur at focal points (for example, by page 126, near a simple focus $S \sim (\Delta b)^{3/2}$). Hence there is a Riemann surface \mathcal{R} on which S is single valued in x . Now for the analogy: do the path integral on \mathcal{R} and for any x get K by summing over contributions from each of the sheets covering x . Note that since WKB fails at foci, K at caustics is not entirely determined by S in that region.

TWENTY-FIVE

Relativistic Propagators and Black Holes

By means of a simple trick it is easy to write down propagators for relativistic spinless particles; when spinning particles are encountered the formalism runs into difficulty, but that is a familiar story.

In this section we describe the “simple trick” and give a brief account of a recent application, namely the quantum emission of energy from black holes. This application actually takes in quite a bit in the way of special techniques of path integration, and besides requiring the form of the propagator on curved space it also involves an analytic continuation of the classical action so as to get a kind of barrier penetration. Unfortunately though, despite the use of curved space propagators, the application turns out to be to a metric of zero curvature and hence does not test the curvature term discussed in Section 24.

Consider first a single Klein-Gordon particle of charge e on flat (Minkowski) space time ($g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$). In the presence of an external electromagnetic field described by a vector potential $A^\mu(x)$ it satisfies the equation

$$\left(i \frac{\partial}{\partial x^\mu} - eA_\mu \right)^2 \psi(x) = m^2 \psi(x) \quad (25.1)$$

Before going further there is an apology to be made for (25.1). Trying to think of the Klein-Gordon equation—or the Dirac equation—as a one-particle equation is fraught with all sorts of conceptual pitfalls, negative energy states, holes, and other ideas that lose their theological character

only after one has understood the subject in a more pedestrian way through field theory.

For our present purposes, (25.1) is taken as a one-particle equation. By using a “fifth parameter,” introduced by Fock, Feynman managed to make the second-order (in time) Klein-Gordon equation look formally like the first-order Schrödinger equation.

Let $\varphi(x^\mu, \lambda)$ be a function of five variables, satisfying the equation

$$i \frac{\partial \varphi}{\partial \lambda} = -\frac{1}{2} \left(i \frac{\partial}{\partial x^\mu} - e A_\mu \right)^2 \varphi \quad (25.2)$$

Because the functions A_μ are independent of λ , (25.2) has solutions of the form $\varphi(x, \lambda) = \exp(im^2\lambda/2)\psi(x)$ with $\psi(x)$ satisfying (25.1). The close resemblance of (25.2) to the Schrödinger equation means that we can now work by blind mathematical analogy and represent the propagator for (25.2) as a path integral for a classical particle moving on a four-dimensional trajectory parametrized by a fifth parameter: $x^\mu(\lambda)$. The Lagrangian for this motion is

$$L = \frac{1}{2} \left(\frac{dx^\mu}{d\lambda} \right)^2 + e \frac{dx^\mu}{d\lambda} A_\mu \quad (25.3)$$

It follows that if one wants the propagator for (25.1) (which one often does) the path integral propagator for (25.2) need only be Fourier transformed (in λ) in complete analogy with the step that takes one from $G(x, t; y)$ to $G(x, y; E)$. Thus

$$G(x''^\mu, x'^\mu; m) = \int_0^\infty d\lambda \exp\left(-\frac{im^2\lambda}{2}\right) \sum_{\substack{x^\mu(\lambda)=x''^\mu \\ x^\mu(0)=x'^\mu}} \exp\left(i \int_0^\lambda L(x(\lambda')) d\lambda'\right) \quad (25.4)$$

(The procedure here is formally identical with that described in Section 20 where a path integral for electromagnetic waves is presented.)

Notwithstanding our earlier tendency toward disparagement, the idea of “negative energy states” as particles moving backward in time arises naturally in the formulation just developed. For a free particle ($A^\mu = 0$) the propagator for (25.2) is

$$G(x'', x'; \lambda) = (2\pi^2\lambda^2 i)^{-1} \exp\left[-\frac{i(x'' - x')^2}{2\lambda}\right] \quad (25.5)$$

Suppose that the boundary values are such that $x''^0 - x'^0 = t'' - t' < 0$, that is, the “final” (in λ) time is “earlier” (in t) than the “initial” (in λ) time. Then G , thought of as a wave function, is an eigenfunction of the energy operator

$$i \frac{\partial G}{\partial t''} \sim i \frac{\partial}{\partial t''} e^{-i(t'' - t')^2/2\lambda} = \frac{t'' - t'}{\lambda} e^{-i(t'' - t')^2/2\lambda} \quad (25.6)$$

with the negative eigenvalue $(t'' - t')/\lambda$. One can also get a nice picture of electron-positron annihilation by considering the propagator for two points x'' and x' having the same $t = x''^0 = x'^0$ and for which at some later time the electromagnetic field A takes large values. Then we can expect significant contributions to the propagator from classical paths which start (i.e., $\lambda = 0$) at x' , propagate “forward” (i.e., $\partial t/\partial\lambda > 0$) to the future, scatter off A , and are reflected “backward” (i.e., $\partial t/\partial\lambda < 0$, λ increases always) and return to x'' . See Fig. 25.1. Mathematically this is no more mysterious than having a particle propagate in space with some momentum, hit a barrier, and return with opposite momentum. As for conservation laws, in the approximation of a rigid potential barrier the potential absorbs the momentum change. Similarly, interpreting the process in Fig. 25.1 as

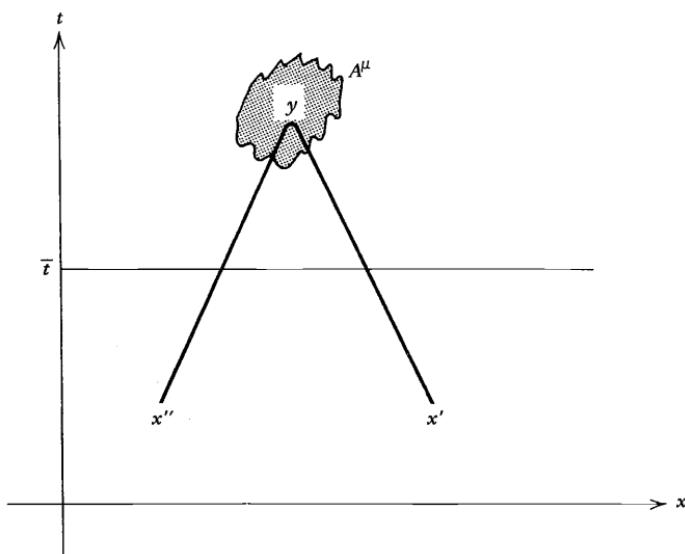


Fig. 25.1 Particle “starts” at x' and at y is scattered (due to the presence of A^μ) “back” to x'' . Interpreted differently: a particle left x' and its antiparticle leaves from x'' . They meet and annihilate at y and their energy and momentum are absorbed by the electromagnetic field A^μ .

particle-antiparticle annihilation, the external electromagnetic field absorbs any extra energy and momentum.

This is also a convenient way of reaching the conclusions sought when discussing the Klein paradox. The burden of those discussions is that you cannot just throw away negative energy states; that is, if you try to restrict yourself to a Hilbert space of positive frequency functions alone (thereby hoping to avoid problems of interpretation), then the dynamics forces negative energy states onto the scene. This is obvious if one considers $G(x'', x'; \lambda)$ with $x''^0 > x'^0$. If there are large fields at later times, then not only will the propagator have contributions from paths for which $dt/d\lambda > 0$ for all λ but also from paths that go further into the future than x''^0 and then bounce back after hitting a large A .

Yet another concept for which this formalism aids the intuition is scattering "off the mass shell." Although the propagator for mass m particles is obtained by means of the Fourier transform, (25.4), there will be paths that contribute (for nontrivial A^μ) for which $(dx^\mu/d\lambda)(dx_\mu/d\lambda)$ is not always equal to m^2 . We will not belabor this point but only remark that there is nothing more mysterious here than there is in off-energy-shell contributions in the nonrelativistic theory.

Together with this plethora of delights to the intuition, the present formalism also serves to remind us of its own limitations as a putative one-particle theory. Obviously, if one were to put a spacelike surface at time \bar{t} in Fig. 25.1, one would see *two* particles, not one. (In fact, I think it is Feynman who once speculated that all the universe's electrons and positrons might be just one electron with many, many reversals of $dt/d\lambda$ in the far future and past). So because we have prejudices regarding the reasonableness of looking at spacelike surfaces (prejudices not paralleled for spatial planes in the nonrelativistic theory) we are forced to consider annihilation and creation of our "single" particle. Another problem, one whose cure requires more than the mere introduction of field theory formalism, arises when one allows particles to interact with each other, not just with external fields. Then one can ask whether or not the two sections of the broken line in Fig. 25.1 ought to interact with each other.

There is a mathematical point that must be made concerning the present formalism, about whose physical significance I am uncertain. Although there is a formal similarity between the evolution equations for the relativistic and nonrelativistic cases there is a difference in the nature of "acceptable" wave functions. In nonrelativistic quantum mechanics we require of course that $\int |\psi(\mathbf{x}, t)|^2 d^3x < \infty$. It would be inappropriate (I think) to require $\int |\varphi(\mathbf{x}, t; \lambda)|^2 d^3x dt < \infty$ for the solutions of (25.2). The probability interpretation of the Klein-Gordon wave function is problematic, but however one succeeds in arriving at such an interpretation it

should not include a requirement that the particle disappear in the distant past and future.

The methods described here work for Dirac particles to the extent that anything involving spin can be made to work in a path integral context. One “squares” the Dirac equation to obtain

$$\left(i \frac{\partial}{\partial x_\mu} - A^\mu \right)^2 \psi - \frac{1}{2} \sigma_{\mu\nu} F^{\mu\nu} \psi = m^2 \psi \quad (25.7)$$

where $F^{\mu\nu}$ is the electromagnetic field and $\sigma_{\mu\nu}$ is (essentially) a commutator of γ matrices. If one again replaces m^2 by $-2i\partial/\partial\lambda$ the resulting equation is formally similar to the Schrödinger equation for a particle with spin and can be handled (clumsily) exactly as done for the Hamiltonian in (22.2).

We turn next to the use of relativistic propagators on curved space-times. The physical problem concerns the emission of radiation from black holes. Since the first thing one learns about black holes—as an explanation of their name—is that they absorb everything, the possibility of emission may be surprising. However, it is only classically that emission is forbidden, and the emission to be derived below is the result of a quantum tunneling process.

Actually even before the emission mechanism described below was worked out, there was reason to expect radiation from black holes. This was because black holes had certain properties best understood in terms of thermodynamics, and when one tried to push the thermodynamic analogy as far as possible, the need for some sort of emission arose. Specifically, Bekenstein, beginning from the fact that when black holes collide the area of the event horizon can never decrease, was led to identify (some multiple of) the area of the event horizon as an entropy so that the nondecrease of area was an expression of the second law of thermodynamics. He then formulated a generalized second law, in which the sum of the black hole entropy and the usual entropy never decreases. Going further, there is also an analogue of the first law in that nearby equilibrium states of a black hole are related by

$$dM = \frac{\kappa}{8\pi} dA + \Omega dJ \quad (25.8)$$

where M is the mass of the black hole and Ω and J are angular velocity and angular momentum. A is the area and κ the surface gravity, equal to $1/4M$. Comparing this to the usual first and second laws of thermodynamics one is led to consider $\kappa dA/8\pi$ as the analogue of TdS and thus—following the inexorable chain of analogies—one is led to associate

a temperature with the black hole, this temperature being some multiple of κ . But now if one wants to take all these analogies seriously there must be some emission of energy from the black hole. For imagine the black hole—described by some temperature—in equilibrium with an ordinary object at the same numerical value of temperature. If they are indeed to be in equilibrium then the energy flow from one to the other must be equal. Since, as in ordinary thermodynamic considerations, there is no reason for this energy flow to be zero, we must assume some energy flow from the black hole.

The emission mechanism can be variously described. Prosaically one can think of the powerful gravitational fields near the black hole as producing particle-antiparticle pairs, with one member of the pair falling into the hole and the other escaping and interpreted by an outside observer as “radiation from the black hole.” This has nothing to do with quantization of the gravitational field. Gravity here is an external classical field and particle pair production is in principle no different from production by an unquantized “external” electromagnetic field.

A more graphic description of the production process is possible in terms of the concepts presented earlier in this section. The particle trajectory is parametrized by a “fifth parameter.” The particle-antiparticle pair whose creation was prosaically described above is thought of as a single particle starting from some future time inside the black hole, propagating backward in time to a point outside the black hole (such classical trajectories *do* exist), there scattering off strong gravitational fields, the scattering being so traumatic that $dt/d\lambda$ (using earlier notation) is reversed and the particle goes forward in time to the observer stationed outside the hole. Moreover, the trajectory that “began” (i.e., lowest value of λ) at some future time in the hole could be continued in the hole to earlier λ (and later t) until it hit the $r=0$ singularity within the black hole. Hence one might consider the particle observed on the outside to have “originated” at the $r=0$ singularity. See Fig. 25.2.

The calculation of the propagators for the processes just described will turn out to involve a very pretty analytic continuation. But even a minimal presentation of that calculation requires a bit of detail concerning the Schwarzschild metric. With t = time and (r, θ, φ) spherical spatial coordinates the metric is

$$\begin{aligned} g_{tt} &= 1 - \frac{2M}{r} & g_{rr} &= -\left(1 - \left(\frac{2M}{r}\right)\right)^{-1} \\ g_{\theta\theta} &= -r^2 & g_{\varphi\varphi} &= -r^2 \sin^2\theta \end{aligned} \quad (25.9)$$

all other components zero. M is a parameter identified with the mass of the

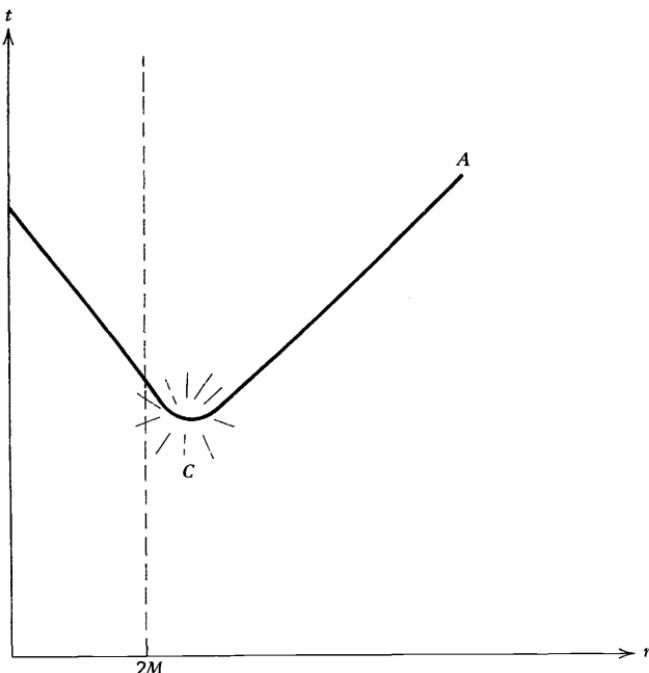


Fig. 25.2 Particle "originating" as negative energy state ($dt/d\lambda < 0$) at black hole singularity and propagating to an outside observer A after "reflecting" off strong gravitational fields at C . t =time, r =radial coordinate.

source of this gravitational field. As is well known, the points $r=2M$ do *not* represent singularities (despite the form of g_{rr}) in the sense that geodesics (which do not depend on the choice of coordinates) show no untoward behavior there. The only singularity is at $r=0$. The coordinate choice most useful for our purpose was introduced by Kruskal, and it has the advantage that when all parameters are extended as far as possible (giving in some sense two copies of the original Schwarzschild space) all geodesics can be continued indefinitely except those ending on the singularity $r=0$.

In the Kruskal coordinates θ and φ are unchanged but r and t are transformed to

$$v = \left[\frac{r}{2M} - 1 \right]^{1/2} \exp\left(\frac{r}{4M}\right) \sinh\left(\frac{t}{4M}\right)$$

$$u = \left[\frac{r}{2M} - 1 \right]^{1/2} \exp\left(\frac{r}{4M}\right) \cosh\left(\frac{t}{4M}\right) \quad (25.10)$$

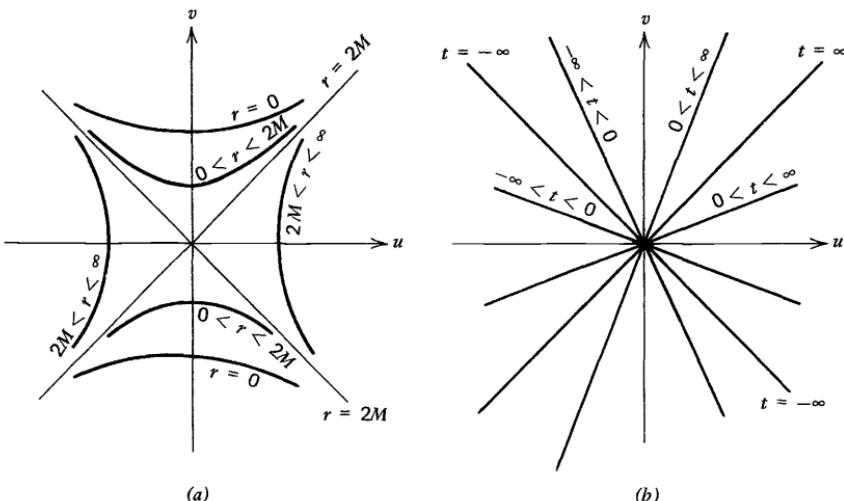


Fig. 25.3 Lines of constant r (a) and t (b) in Kruskal coordinates. Each point (u, v) represents the surface of a sphere (for $0 < \theta < \pi, 0 < \varphi < 2\pi$).

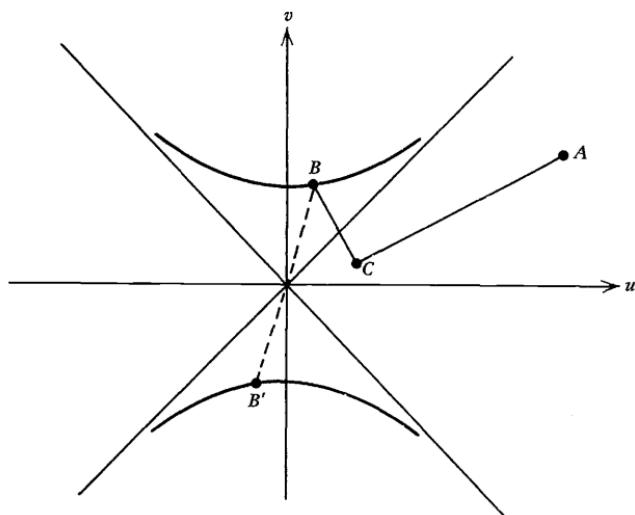


Fig. 25.4 Particle trajectory BCA leading to the observation, at A , of a particle "emitted" by the black hole. Analytic continuation relates the amplitude for this to the amplitude for particles emitted at A to reach B' .

In Fig. 25.3 we illustrate surfaces of constant r (hyperbolas) and constant t (straight lines) in the $u-v$ plane. In Fig. 25.4 is shown one of the trajectories that contribute to black hole emission, as described above. The particle that ultimately reaches A is scattered off the strong gravitational fields at C , having "originated" (in the sense of earliest values of λ) at B . In the prosaic description first employed one would say a pair is created at C , one member going to B , one to A .

Just as was described earlier for Minkowski space, the propagator for a particle to go from x_B to x_A can be written

$$G(x_A; x_B) = \int_0^\infty e^{-im^2\lambda/2} G(x_A, x_B, \lambda) d\lambda \quad (25.11)$$

with $G(x_A, x_B; \lambda)$ given by the path integral in the Schwarzschild metric

$$G(x_A, x_B; \lambda) = \sum_{x(\cdot) \atop x(0)=x_A \atop x(\lambda)=x_B} \exp\left(i \int_0^\lambda \frac{1}{2} g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} d\lambda\right) \quad (25.12)$$

The observer at A further restricts attention to particles of some fixed energy E and in some particular state. The amplitude for this will be called f , which is just the Fourier transform of G of (25.11), and is

$$f(\mathbf{R}_A, \mathbf{R}_B) = \int_{-\infty}^\infty dt e^{-iEt} G(0, \mathbf{R}_A; t, \mathbf{R}_B)$$

or, using the symmetry under interchange of x_A and x_B ,

$$f(\mathbf{R}_A, \mathbf{R}_B) = \int_{-\infty}^\infty dt e^{-iEt} G(t, \mathbf{R}_B; 0, \mathbf{R}_A) \quad (25.13)$$

Now everything is defined. What can one say about G ?

It turns out that the main thing one wants to say (which is to relate black hole temperature and emission) can be said on the basis of G 's analyticity properties alone. Hartle and Hawking show that (in the semi-classical approximation), as a function of complex t , G is analytic in a strip of width $4\pi M$ below the real t axis. They also show that just above and just below this strip there are singularities due to null geodesics, but we shall not need this information. The contour of integration in (25.13) can therefore be deformed downward by $4\pi Mi$ with no change in f . Going to a new variable $t + 4\pi iM \rightarrow t$, f becomes ($\kappa = 1/4M$)

$$f(\mathbf{R}_A, \mathbf{R}_B) = e^{-\pi E/\kappa} \int_{-\infty}^\infty dt e^{-iEt} G\left(t - \frac{i\pi}{\kappa}, \mathbf{R}_B; 0, \mathbf{R}_A\right) \quad (25.14)$$

The beauty of this change of variables is that in the Kruskal coordinates the "time" $t - i\pi/\kappa$ corresponds to real physical u and v (or t and r), and so the original physical process has been related to another one with a factor $\exp(-\pi E/\kappa)$ connecting the amplitudes.

From (25.10) it is clear that $t \rightarrow t - i\pi/\kappa$ simply sends $u \rightarrow -u$ and $v \rightarrow -v$. Hence the propagator describes a particle emitted at B' (Fig. 25.4), the reflection of B in the origin of the $u-v$ coordinates, and received at A . But by time reversal invariance this amplitude is the same (up to a sign) as that for a particle to be emitted at A and received at B' . The absolute value squared of this quantity is just the absorption probability for a black hole. Hence from (25.14) and its interpretation we conclude that

(emission probability for particle of energy E)

$$= \exp\left(\frac{-2\pi E}{\kappa}\right) \times (\text{absorption probability for particle in same state}) \quad (25.15)$$

This relation is characteristic of black bodies of temperature $T = \kappa/2\pi k_B$ (k_B = Boltzmann constant) and thereby justifies the assigning of a temperature to a black hole.

We shall not go into the proof of the analyticity properties of G . The methods used by Hartle and Hawking involve continuation to pure imaginary λ to get a diffusion equation rather than a Schrödinger equation. The same kind of continuation is done for t , changing the hyperbolic wave equation with its D'Alembartian to a more tractable object involving the Laplace-Beltrami operator. In their deductions they sometimes use the path integral language, in effect making use of the fact that asymptotic approximations for solutions of the diffusion or Schrödinger equation can be given in terms of geodesics of the metric, that is, the classical paths. Other general asymptotic properties of the solutions of partial differential equations are also used. Full exposition would take us yet farther afield, and the interested reader is referred to the original works.

NOTES

Feynman's development of the relativistic propagator for Klein-Gordon and Dirac particles is given in

R. P. Feynman, Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction, *Phys. Rev.* **80**, 440 (1950), App. A

R. P. Feynman, An Operator Calculus Having Applications in Quantum Electrodynamics, *Phys. Rev.* **84**, 108 (1951), App. D

Following (25.7) we indicated that a relativistic path integral for spinning particles could be developed using a product integral as in Section 22.1. It would not be idle to ask whether a continuous spin model could be incorporated into a relativistic path integral, further pursuing the ideas of Section 22.2. In working on this question some years ago I managed to produce a spin model and relativistic path integral, but ran into trouble when this spinning particle was allowed to interact with an external field. The spin model is described in

L. S. Schulman, Relativistic Spin: Tops and Wave Equations, *Nuclear Phys.* **B18**, 595 (1970)

The spin coordinates are (as in the nonrelativistic case) the group manifold of $SO(3)$, but now these coordinates cannot be clearly separated from the spatial coordinates (essentially because boosts and rotations do not commute relativistically). For the noninteracting particle a path integral can be written down for the propagator, but I did not succeed in developing a theory with interactions. My efforts on this problem are described in my 1967 Princeton thesis, but were not published.

Black hole thermodynamics is developed in

J. D. Bekenstein, Black Holes and Entropy, *Phys. Rev. D* **7**, 2333 (1973)

J. D. Bekenstein, Generalized Second Law of Thermodynamics in Black Hole Physics, *Phys. Rev. D* **9**, 3292 (1974)

Emission by black holes is treated in

S. W. Hawking, Black Hole Explosions? *Nature* **248**, 30 (1974)

S. W. Hawking, Particle Creation by Black Holes, *Commun. Math. Phys.* **43**, 199 (1975)

The handling of this problem through the use of path integrals appears in

J. B. Hartle and S. W. Hawking, Path-Integral Derivation of Black-Hole Radiance, *Phys. Rev. D* **13**, 2188 (1976)

where can be found the proofs of analyticity mentioned in the text. A general background to the ideas of black hole thermodynamics is given by

J. D. Bekenstein, Black Hole Thermodynamics, *Phys. Today*, January 1980, p. 24

Further information on the Kruskal coordinates is in the book

J. L. Anderson, *Principles of Relativity Physics*, Academic Press, New York, 1967

The ultimate in combining the Feynman quantization principle with relativity is to use that principle to quantize the gravitational field itself. The idea is to allow "all" 3-geometries, including those that do not satisfy the classical field equations. The amplitude for a given 3-geometry is $\exp(iS/\hbar)$ where S is a functional on the geometry, in particular an action whose extrema satisfy the classical field equations. Now the rest of the theory is easy: just sum these amplitudes. But of course it is not so easy and the intrinsic difficulties of quantizing the gravitational field make themselves

felt. An early work in this field is

C. W. Misner, Feynman Quantization of General Relativity, *Rev. Mod. Phys.* **29**, 497 (1957)

Further attempts are described in

J. A. Wheeler, Superspace and the Nature of Quantum Geometrodynamics, in C. M. DeWitt and J. A. Wheeler, Eds. *Battelle Rencontres*, Benjamin, New York, 1968

C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, Freeman, San Francisco, 1971

More recently Hawking has used this approach combined with analytic continuation to obtain a "Euclideanized" form of the quantum field theory. He does not restrict his sum to 3-geometries, and defining boundary conditions for the sum is touchy. There are also convergence problems because in the analytically continued amplitude $\exp(-S/\hbar)$ there are metrics for which S is large and negative. The stationary phase approximation recovers the classical field equations, and the small oscillation integral ("one loop approximation") gives thermal gravitons. Hawking also presents a derivation of black hole thermodynamics from the quantized theory. All this can be found in

S. W. Hawking, The Path-Integral Approach to Quantum Gravity, in S. W. Hawking and W. Israel, Eds., *General Relativity*, Cambridge, London, 1979

See also

G. W. Gibbons, S. W. Hawking, and M. J. Perry, Path Integrals and the Indefiniteness of the Gravitational Action, *Nucl. Phys. B* **138**, 141 (1978)

From the technical path integral viewpoint this approach shares with other field theory quantizations the defect (?) that very rough fields enter the "path" sum. As discussed in Section 32.1 not only will the fields entering the sum lack derivatives, but it would take several smoothings (by integration) to make them as "good" as Brownian motion paths. For gravity this may not be so serious a drawback: in terms of an action I having dimensions of length the amplitude is $\exp(iI/L_0)$ with $L_0 = (\hbar G/c^3)^{1/2}$, the Planck length (G is the gravitational coupling constant). Since $L_0 \sim 1.6 \times 10^{-33}$ cm, the bad behavior of the field may well be on a scale so remote from observation that any submicroscopic pathologies the quantization allows will have been smoothed and rendered harmless at large distances like 10^{-13} cm.

TWENTY-SIX

Applications to Statistical Mechanics

One of the principal uses of functional integration is the calculation of the partition function in statistical mechanics. More than one book could be written on this subject, and in fact several monograph-sized review papers have come out devoted to various applications. Hence our treatment will be far from comprehensive.

The basic fact underlying most, but not all, applications is that the partition function, the central theoretical quantity of statistical mechanics, is

$$Z = \text{Tr} \exp(-\beta H) \quad (26.1)$$

with β inverse temperature ($\beta = 1/k_B T$, k_B = Boltzmann constant). Path integration on the other hand is devoted to the evaluation of

$$G(x, t; y) = \left\langle x \left| \exp\left(\frac{-itH}{\hbar}\right) \right| y \right\rangle \quad (26.2)$$

the kernel of the evolution operator. Hence if one has a way of calculating G , then by doing a trace and an analytic continuation he has obtained the partition function.

Consider first the simple application of this relation to an ideal gas at (inverse) temperature β . Each of the, say, N particles has the one-particle Hamiltonian $H = p^2/2m + V(x)$, but they do not interact with each other. Then

$$\begin{aligned} Z &= \text{Tr} \exp\left(-\beta \sum_{i=1}^N H_i\right) = [\text{Tr} \exp(-\beta H)]^N \\ &= \left[\text{Tr} \exp\left(\frac{-i(\beta\hbar/i)H}{\hbar}\right) \right]^N = \left[\int dx G(x, -i\beta\hbar; x) \right]^N \end{aligned} \quad (26.3)$$

The classical limit of Z is immediately obtained through the grossest possible approximation to G ,

$$G(x, t; y) = \sqrt{\frac{m}{2\pi i\hbar t}} \exp\left[\frac{i}{\hbar} \left(\frac{m}{2} \frac{(x-y)^2}{t} - V(x)t \right) \right] \quad (26.4)$$

This corresponds to an evaluation of the classical action on the straight line from y to x , with no intervening steps [compare (1.24)]. Substituting and changing variables ($t \rightarrow -i\beta\hbar$) as in (26.3) gives

$$Z^{1/N} = \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}} \int \exp(-\beta V) dx \quad (26.5)$$

which is precisely the correct classical result, even including the correct volume element ($2\pi\hbar$) to use when partitioning phase space to count states.

It can also be useful to write Z itself as a functional integral. From (26.3) we have

$$\begin{aligned} \xi \equiv Z^{1/N} &= \int dx G(x, -i\beta\hbar; x) \\ &= \sum_{\text{paths}} \exp\left[\frac{i}{\hbar} \int_0^{-i\beta\hbar} dt' \left(\frac{m}{2} \left(\frac{dx}{dt'} \right)^2 - V(x) \right) \right] \end{aligned} \quad (26.6)$$

By a change of variables $\tau = it/\hbar$ (or it may be preferable to look on this as an analytic continuation) the integral takes the form

$$\xi = \sum \exp\left[- \int_0^\beta dt' \left(\frac{m}{2\hbar^2} \left(\frac{dx}{dt'} \right)^2 + V(x) \right) \right] \quad (26.7)$$

which can also be written (for time independent V)

$$\xi = E(\exp(-\beta V)) \quad (26.8)$$

using the notation of expectation value over the sample space of Brownian motion paths as introduced in Section 9. Note that what corresponds to the diffusion coefficient in (26.7) is \hbar^2 so that for $\hbar \rightarrow 0$ uncertainty in position due to this "diffusion" vanishes, as one would expect.

By a different change of variables, $\tau = it$, the argument of the exponential in (26.7) can be made to look like the Hamiltonian, and this form is occasionally useful.

The Wigner-Kirkwood expansion provides quantum corrections to (26.5). The first term in this expansion is easily obtained by considering $G(x, t; x)$ for small t which, recalling the replacement $t \rightarrow -i\beta\hbar$, means small \hbar or large temperature. The classical path from $(x, 0)$ to (x, t) is $x(t') \equiv x$, $0 < t' < t$, and the correction comes by considering fluctuations about this constant path. Hence, letting $\xi(t') = x(t') - x$ be the integration variable in the functional integral we have

$$G(x, t; x) = \sum_{\substack{\xi(0)=0 \\ \xi(t)=0}} \exp \left\{ \frac{i}{\hbar} \int_0^t dt' \left[\frac{1}{2} m \left(\frac{d\xi}{dt'} \right)^2 - V(x + \xi) \right] \right\} \quad (26.9)$$

Since fluctuations in x are small, that is, $\xi^2 = O(t)$, we expand V to second order in ξ and note that the resulting effective Lagrangian is quadratic, hence is known from our previous work. Write therefore

$$\begin{aligned} V(x + \xi) &= V(x) + \xi V'(x) + \frac{1}{2} \xi^2 V''(x) + O(\xi^3) \\ &= V - \frac{\frac{1}{2}(V')^2}{V''} + \frac{1}{2} V'' [\xi + V'/V'']^2 + O(\xi^3) \end{aligned} \quad (26.10)$$

so that (26.9) is approximated as a harmonic oscillator with $m\omega^2/2 = V''/2$, whose propagator will be called G^0 . Hence

$$G(x, t; x) = \exp \left[-\frac{i}{\hbar} \int_0^t \left(V - \frac{(V')^2}{2V''} \right) dt' \right] G^0 \left(\frac{V'}{V''}, t; \frac{V'}{V''} \right) \quad (26.11)$$

By our results on the harmonic oscillator (6.3) this is

$$\begin{aligned} G(x, t; x) &= \exp \left[-\frac{it}{\hbar} \left(V - \frac{(V')^2}{2V''} \right) \right] \sqrt{\frac{m\omega}{2\pi i\hbar \sin \omega t}} \\ &\times \exp \left[\frac{im\omega}{\hbar \sin \omega t} \left(\frac{V'}{V''} \right)^2 (\cos \omega t - 1) \right] \end{aligned} \quad (26.12)$$

Since $|t|$ is assumed small (and we now see that it must be small compared

to $\sqrt{m/V''}$) this becomes

$$G(x, t; x) = \exp\left[-\frac{it}{\hbar}\left(V - \frac{(V')^2}{2V''}\right)\right] \sqrt{\frac{m}{2\pi i\hbar t}} \sqrt{\frac{1}{1 - \frac{\omega^2 t^2}{6}}} \\ \times \exp\left\{\frac{im\omega^2 t^2/2}{\hbar t\left(1 - \frac{\omega^2 t^2}{6}\right)} \left(\frac{V'}{V''}\right)^2 \left(-1 + \frac{\omega^2 t^2}{12}\right)\right\} \quad (26.13)$$

Recalling that $m\omega^2 = V''$, the peculiar term $(V')^2/2V''$ cancels, and keeping terms t and t^3 in the exponent, G is

$$G(x, t; x) = \left(1 + \frac{V''\beta^2 \hbar^2}{12m}\right) \sqrt{\frac{m}{2\pi i\hbar t}} \exp\left(-\frac{itV}{\hbar}\right) \exp\left[\frac{-it^3(V')^2}{24\hbar m}\right] \quad (26.14)$$

Returning to the calculation of ξ , t is replaced by $-i\beta\hbar$, the exponent involving t^3 expanded, and an integral over x performed. Within the integral over x an integration by parts is performed, replacing V'' by $\beta(V')^2$. Finally

$$\xi = \int dx \frac{1}{\hbar} \sqrt{\frac{m}{2\pi\beta}} e^{-\beta V} \left[1 - \frac{\beta^3 \hbar^2}{24m} \left(\frac{dV}{dx}\right)^2\right] \quad (26.15)$$

which constitutes the first two terms (orders \hbar^0 and \hbar^2) of the Wigner-Kirkwood expansion.

This demonstration appears to require that $V'' \neq 0$, but in the final result V'' does not appear. In fact should V'' vanish, the calculation is even easier, and instead of using the known formula for the harmonic oscillator Green's function, (6.3), we could have used the corresponding formula for a particle in a uniform field, (6.5).

Our restriction to ideal gases was a matter of expository convenience. In the Hamiltonian that we have written, $p^2/2m + V(x)$, the variables p and x could be thought of as N -dimensional (N degrees of freedom). The new form of (26.5) is immediate:

$$Z = \left(\frac{m}{2\pi\beta\hbar^2}\right)^{N/2} \int d^N x \exp(-\beta V(x_1, \dots, x_N)) \quad (26.16)$$

For V a sum of one-particle potentials the previous equation is recovered.

The Wigner-Kirkwood expansion can also be easily generalized and aside from having to use $(m/2\pi\beta\hbar^2)^{N/2}$ one only need reinterpret dx and $(\partial V/\partial x)^2$.

NOTES

Any attempt at a comprehensive bibliography on applications of path integration to statistical physics would be absurd, and I shall here content myself with review papers and books. Among the latter...

R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965

R. P. Feynman, *Statistical Mechanics*, Benjamin, Reading, Mass., 1972

M. Kac, *Probability and Related Topics in Physical Sciences*, Interscience, New York, 1959

Some review papers are

S. G. Brush, Functional Integrals and Statistical Physics, *Rev. Mod. Phys.* **33**, 79 (1961)

S. K. Ma, Introduction to the Renormalization Group, *Rev. Mod. Phys.* **45**, 589 (1973)

F. W. Wiegel, Path Integral Methods in Statistical Mechanics, *Phys. Rep. (Phys. Lett. C)* **16**, 57 (1975)

TWENTY-SEVEN

Coherent State Representation

In field theory, the free field can often be represented as a collection of oscillators, but the “coordinate” and “momentum” variables of the oscillators do not have physical significance. Of greater interest are the creation and annihilation operators, and a functional integral related to these operators is more useful for field theory calculations.

The usual path integral is obtained by writing the propagator $\exp(-iHt)$ as $[\exp(-iHt/N)]^N$ and inserting a resolution of the identity in terms of coordinate state basis functions between the terms in the product. Similarly, the phase space path integral (Section 31) can be generated with the insertion of both $\mathbf{1} = \int dq |q\rangle\langle q|$ and $\mathbf{1} = \int dp |p\rangle\langle p|$ at different times. By using a basis of eigenfunctions of the annihilation operator yet another path integral emerges, although as we shall see it is in some ways the same as the phase space path integral. These eigenfunctions are called coherent states and are particularly suited to many-body theory and to the path integration of field theories.

The basic arithmetic for this form of the functional integral already appears for a first quantized one dimensional system, that is, the simplest case of quantum mechanics. Hence we shall do this case first and in detail.

Let the coordinate and momentum of a particle be given by x and p respectively. Let the Hamiltonian be H . For convenience we define

$$H_0 = \frac{p^2}{2} + \frac{x^2}{2} \quad (27.1)$$

which may or may not be the same as H . Later a frequency ϵ will be introduced, but for now we keep the formulas simple.

Define the annihilation (or lowering) operator

$$a = \frac{x + ip}{\sqrt{2}} \quad (27.2)$$

and its conjugate the creation (or raising) operator

$$a^\dagger = \frac{x - ip}{\sqrt{2}} \quad (27.3)$$

By virtue of $[x, p] = i$ these satisfy

$$[a, a^\dagger] = 1 \quad (27.4)$$

There now follows the whole slew of the usual results, some of which we list:

$$\begin{aligned} N &= a^\dagger a = (\text{number operator}) = H_0 - \frac{1}{2} \\ |n\rangle &= (\text{normalized eigenstate of } N) = (n!)^{-1/2} (a^\dagger)^n |0\rangle \\ N|n\rangle &= n|n\rangle \quad n = 0, 1, 2, \dots \\ \langle x|n\rangle &= (\pi^{1/2} 2^n n!)^{-1/2} H_n(x) e^{-x^2/2} \\ H_n(x) &= (-1)^n e^{x^2} \left(\frac{\partial}{\partial x} \right)^n e^{-x^2} = (n \text{th Hermite polynomial}) \\ e^{-y^2 + 2xy} &= \sum_{n=0}^{\infty} \frac{H_n(x)y^n}{n!} \quad \text{generating function} \\ a|n\rangle &= \sqrt{n}|n-1\rangle \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \end{aligned} \quad (27.5)$$

The coherent states are the eigenstates of the non-Hermitian operator a . It is easy to verify that these states have the form

$$|z\rangle = N(z) \exp(z a^\dagger) |0\rangle \quad (27.6)$$

where z is any complex number, $N(z)$ is a normalization constant, and $|0\rangle$ is the vacuum or ground state according to the notation of (27.5). By the properties of a

$$\begin{aligned} a|z\rangle &= N(z)a \sum_{n=0}^{\infty} \frac{z^n}{n!} (a^\dagger)^n |0\rangle = N(z) \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} a|n\rangle \\ &= N(z) \sum_{n=1}^{\infty} \frac{z^n}{\sqrt{(n-1)!}} |n-1\rangle = z|z\rangle \end{aligned} \quad (27.7)$$

The state $\langle z|$ is the adjoint of $|z\rangle$ and is given by $N(z)^* \langle 0 | \exp(z^*a)$. The inner product of two coherent states is

$$\langle z'|z\rangle = N(z')^* N(z) \langle 0 | \exp(z'^*a) \exp(za^\dagger) | 0 \rangle \quad (27.8)$$

We can expand the exponent as in (27.7), but it is easier to use the identity

$$e^{A+B} = e^A e^B e^{-\frac{1}{2}[A, B]} = e^B e^A e^{-\frac{1}{2}[B, A]} \quad (27.9)$$

which holds whenever $[A, B]$ commutes with both A and B . Hence changing the order of the operators in (27.8) introduces a factor $\exp(z'^*z)$, while the operator $\exp(z'^*a)$ acting on $|0\rangle$ simply gives $|0\rangle$; similarly for $\langle 0 | \exp(za^\dagger)$. Therefore

$$\langle z'|z\rangle = N(z')^* N(z) \exp(z'^*z) \quad (27.10)$$

It follows that as normalization we can take $N(z) = \exp(-|z|^2/2)$ and finally

$$|z\rangle = \exp\left(-\frac{|z|^2}{2} + za^\dagger\right) |0\rangle, \quad \langle z'|z\rangle = \exp\left(z'^*z - \frac{1}{2}(|z|^2 + |z'|^2)\right) \quad (27.11)$$

The coherent states are also “minimum uncertainty wave packets” and their time development under H_0 is trivial. We list a number of properties

$$\begin{aligned} |z\rangle &= \exp\left(-\frac{|z|^2}{2}\right) \sum_{n=0}^{\infty} z^n (n!)^{-1/2} |n\rangle \\ \langle z|x|z\rangle &= \sqrt{2} \operatorname{Re} z, \quad \langle z|p|z\rangle = \sqrt{2} \operatorname{Im} z \\ \exp(-iH_0 t)|z\rangle &= \exp\left(-\frac{it}{2}\right) |z \exp(-it)\rangle \\ \langle w|\exp(-iH_0 t)|z\rangle &= \exp\left[-\frac{it}{2} + w^*ze^{-it} - \frac{(|w|^2 + |z|^2)}{2}\right], \quad z, w \in \mathbb{C} \\ \langle x|z\rangle &= \pi^{-1/4} \exp\left[-\frac{|z|^2}{2} - \frac{1}{2}(x^2 + z^2) + \sqrt{2} xz\right] \end{aligned} \quad (27.12)$$

The last line is the coherent state wave function in the coordinate representation, and this function is essentially the generating function for the Hermite polynomials. One also recognizes in $\langle x|z\rangle$ the minimum uncertainty wave packet, and the explicit information on time dependence

indicates that there is no spreading of this packet in time when in a harmonic oscillator potential with frequency $\omega = 1$.

The important formulas for our purpose involve the resolution of the identity. The resolution is not unique for the very good reason that the set of states $\{|z\rangle |z \in \mathbb{C}\}$ is overcomplete. It should not be surprising that the non-Hermitian operator a has yielded more eigenstates than are needed to serve as a basis for the Hilbert space. (In fact, defining $K = \{|n+im\rangle |n, m = \text{integer}\}$, a complete set can be obtained by leaving out any one (not more, not less) state from K . This leads to some very singular—and pretty—mathematics. See the notes.) A particularly convenient form is the following

$$\mathbf{1} = \int \frac{d^2 z}{\pi} |z\rangle \langle z| \quad (27.13)$$

where $d^2 z$ is integration in the complex plane: $d^2 z = d(\text{Re } z)d(\text{Im } z)$. Equation 27.13 can be proved by expanding $|z\rangle$ in harmonic oscillator eigenstates and using

$$\int d^2 z e^{-|z|^2} z^* z^m = \pi n! \delta_{nm} \quad (27.14)$$

Exercise: Check the assertion concerning overcompleteness by finding a function $v(z)$ such that

$$\int d^2 z v(z) |z\rangle = 0$$

By the parenthetical remark above there must also be coefficients v_{nm} such that

$$|0\rangle = \sum_{\substack{n, m \\ \text{but not } n=m=0}} |n+im\rangle v_{nm}$$

Find them.

A pleasant digression at this stage would be to study the Hilbert spaces of analytic functions developed by Bargmann. By and large we shall resist this temptation, except for the following brief remarks.

Starting with the original Hilbert space, say $L^2(\mathbb{R})$, we can consider, for each $|\psi\rangle \in L^2(\mathbb{R})$ the complex function

$$f(z) = e^{|z|^2/2} \langle z^* | \psi \rangle \quad (27.15)$$

By expanding $\langle z^* |$ in harmonic oscillator eigenstates it is clear that f has an everywhere convergent power series expansion and is not just analytic but entire. For $f(z)$ and for a function $g(z)$ defined analogously in terms of a state $|\varphi\rangle$ the integral

$$\frac{1}{\pi} \int d^2z e^{-|z|^2} f^*(z) g(z) \quad (27.16)$$

is equal to the inner product $\langle \psi | \varphi \rangle$, a fact that follows directly from (27.13) (with z^* in place of z).

Now turn everything around and consider a vector space of entire analytic functions of one or more complex variables. Taking (27.16) (or its multidimensional generalization) as an inner product yields a Hilbert space (\mathcal{F}_n , n = number of complex variables) of those entire functions for which the norm is finite. Equation 27.15 is then interpreted as a unitary transformation relating $L^2(\mathbb{R})$ and \mathcal{F}_1 (or $L^2(\mathbb{R}^n)$ and \mathcal{F}_n). It is obvious that on \mathcal{F}_1 the operator a^\dagger is represented by multiplication by z and nearly so that a on \mathcal{F}_1 is just d/dz .

One of the elegant applications of this approach is in finding the representations of $SU(2)$. A matrix $U \in SU(2)$ is a unitary matrix of determinant 1 acting on \mathbb{C}^2 . This action induces a transformation on the Hilbert space \mathcal{F}_2 in the usual way:

$$(A_U f)(\xi) = f(U^T \xi), \quad \xi = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2 \quad (27.17)$$

where T denotes transpose. $\{A_U\}$ is obviously a unitary representation of $SU(2)$. The profit in all this formalism is realized by noting that A_U leaves invariant the subspace of polynomials in \mathcal{F}_2 which are homogeneous in z_1 and z_2 . Thus each subspace

$$\begin{aligned} Q_j &= \left\{ f(\xi) = \sum a_{nm} z_1^n z_2^m \mid n+m=2j, n \geq 0, m \geq 0, \xi = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\} \\ &= \{f \in \mathcal{F}_2 \mid f(\lambda \xi) = \lambda^{2j} f(\xi)\} \\ j &= 0, \frac{1}{2}, 1, \dots \end{aligned} \quad (27.18)$$

is left invariant under $\{A_U\}$ and therefore provides a finite-dimensional (of dimension $2j+1$) representation of $SU(2)$. The infinitesimal generators of $\{A_U\}$ are also easy to obtain:

$$\begin{aligned} M_1 + iM_2 &= z_1 \frac{d}{dz_2} & M_1 - iM_2 &= z_2 \frac{d}{dz_1} \\ M_3 &= \frac{1}{2} \left(z_1 \frac{d}{dz_1} - z_2 \frac{d}{dz_2} \right) \end{aligned} \quad (27.19)$$

They satisfy $\mathbf{M} \times \mathbf{M} = i\mathbf{M}$. A fourth operator is conveniently defined as

$$N = \frac{1}{2} \left(z_1 \frac{d}{dz_1} + z_2 \frac{d}{dz_2} \right) \quad \text{so that } M^2 = M_1^2 + M_2^2 + M_3^2 = N(N+1) \quad (27.20)$$

Applying the transformations (27.17) to the spaces Q_j , one gets the representations of $SU(2)$ in a jiffy; not only do convenient expressions for the representation matrices ("D" matrices) emerge, but this is also a good starting point for calculation of $3j$ and $6j$ symbols. That the representations on the spaces Q_j are irreducible can be shown using Schur's lemma applied to the matrices M_i (restricted to Q_j). This formalism does not seem, however, to provide any shortcut for the demonstration of completeness, namely that there are no other irreducible representations of $SU(2)$.

If the generators M_i are rewritten in terms of a_k and a_k^\dagger ($k=1, 2$ corresponding to z_1 and z_2), then one sees that the boson creation and annihilation operators of field theory provide representations of $SU(2)$ in a natural way. To some extent this lies behind the frequent appearance of $SU(2)$ as a symmetry group in field theory, even where the $SU(2)$ group has nothing to do with rotations of (physical 3-) space. Thus isospin and pseudospin operators provide $SU(2)$ representations. However, the appearances of $SU(2)$ are even more common than is suggested by our formalism, since that group emerges even for fermion field operators.

The path integral in the coherent state representation is either easy or very easy, depending on which of two forms one prefers. First, the very easy result, (for the harmonic oscillator) can be obtained through evaluation of the finite time propagator:

$$\begin{aligned} G(z_f, t; z_i) &= \langle z_f | e^{-iH_0 t} | z_i \rangle = e^{-it/2} \langle z_f | z_i e^{-it} \rangle \\ &= \exp \left[-\frac{1}{2} (|z_f|^2 + |z_i|^2) + z_f^* z_i e^{-it} - \frac{it}{2} \right] \end{aligned} \quad (27.21)$$

Given a propensity for doing Gaussian integrals, one could use $G(z_f, t; z_i)$ and $\langle x | z \rangle$, (27.12), to derive the harmonic oscillator propagator that was obtained earlier in this volume.

Another form of the propagator, looking more like a path integral, has also proved useful. Let $\epsilon = t/(n+1)$ and insert the resolution of the identity n times between the states z_f and z_i . Without assuming the harmonic oscillator form for H we obtain ($z_{n+1} = z_f$, $z_0 = z_i$)

$$G(z_f, t; z_i) = \pi^{-n} \int d^2 z_1 \dots d^2 z_n \prod_{j=1}^{n+1} \langle z_j | e^{-iH\epsilon} | z_{j-1} \rangle \quad (27.22)$$

In general the Hamiltonian can be written

$$H(\alpha^*, \beta) = \frac{\langle \alpha | H | \beta \rangle}{\langle \alpha | \beta \rangle}, \quad |\alpha\rangle, |\beta\rangle \text{ coherent states} \quad (27.23)$$

which yields for (27.22)

$$G = \pi^{-n} \int \left(\prod_{i=1}^n d^2 z_i \right) \prod_{j=1}^{n+1} \langle z_j | z_{j-1} \rangle \left[1 - i\epsilon H(z_j^*, z_{j-1}) + O(\epsilon^2) \right] \quad (27.24)$$

The inner product $\langle \alpha | \beta \rangle$ can be expressed as

$$\begin{aligned} \langle \alpha | \beta \rangle &= \exp \left[-\frac{1}{2} (|\alpha|^2 + |\beta|^2) + \alpha^* \beta \right] \\ &= \exp \left[-\frac{1}{2} [\alpha^* (\alpha - \beta) - (\alpha^* - \beta^*) \beta] \right] \end{aligned} \quad (27.25)$$

By dividing and multiplying by ϵ , and by calling

$$\frac{z_j - z_{j-1}}{\epsilon} = \frac{dz_j}{dt} \quad (27.26)$$

we obtain

$$\begin{aligned} G &= \pi^{-n} \int \prod_{i=1}^n d^2 z_i \exp \left\{ i\epsilon \sum_{j=1}^{n+1} \left[\frac{-1}{2i} \left(z_j^* \frac{dz_j}{dt} - \frac{dz_j^*}{dt} z_{j-1} \right) - H(z_j^*, z_{j-1}) \right] \right. \\ &\quad \left. + \sum_{j=1}^{n+1} O(\epsilon^2) \right\} \end{aligned} \quad (27.27)$$

Neglect of $O(\epsilon^2)$ terms is by now standard procedure; pending further discussion we replace z_{j-1} by z_j , introducing errors of $O(\epsilon(z_j - z_{j-1}))$. Dropping these errors, yields a path integral expression for G , involving sums over the complex valued functions $z(\tau)$ with $z(0) = z_i$, $z(t) = z_f$. The path integral is

$$\begin{aligned} G(z_f, t; z_i) &= \lim_{n \rightarrow \infty} \pi^{-n} \int d^2 z_1 \dots d^2 z_n \\ &\quad \times \exp \left\{ i \int_0^t d\tau \left[\frac{i}{2} \left(z^* \frac{dz}{d\tau} - \frac{dz^*}{d\tau} z \right) - H(z^*, z) \right] \right\} \end{aligned} \quad (27.28)$$

with the derivatives and integral in (27.28) interpreted as limits of sums.

No one (to my knowledge) has made a serious investigation of the neglected terms $\epsilon(z_j - z_{j-1})$. My own guess is that they can contribute and that this contribution will be related to the operator ordering problem in quantum mechanics. Thus (27.28) and (27.23) *appear* to give a unique propagator for any Hamiltonian—but this cannot be right, since, as many have found to their dismay, path integrals do not settle ambiguities in the ordering of operators. The reason $\epsilon(z_j - z_{j-1})$ does not disappear is that the curves $z(\tau)$ that contribute to the “sum over paths” are more singular than the Brownian motion paths $x(\tau)$ that enter the usual path integrals ($x(\tau)$ is continuous but nowhere differentiable: $(\Delta x)^2 \sim \Delta t$). There are two ways to look at the bad behavior of $z(\tau)$. For one thing, states $|z\rangle$, $|z'\rangle$ ($z \neq z'$) have finite overlap, so that the quantity

$$\langle z' | \exp(-iHt) | z \rangle$$

does not go to zero as $t \rightarrow 0$, in contrast, say, to

$$\langle x' | \exp(-iHt) | x \rangle$$

($|x\rangle$ = position eigenstate) which becomes a δ -function in that limit. Another view of this problem is by analogy to the phase space path integral. In a sense the integration element d^2z behaves like $dp dx$ (recall $a = (x + ip)/\sqrt{2}$) and the integrand in (27.28) is the analogue of $\frac{1}{2}(p dx - x dp) - H dt$. A well known (see Section 31) feature of the phase space path integral is that the paths in phase space are discontinuous so that a term $\epsilon \Delta p$ need not go to zero any faster than ϵ .

Notwithstanding the difficulties just discussed, we shall examine the properties of the path integral (27.28). First we view the argument of the exponent as an action and see what sort of classical mechanics results. If z and z^* are treated as independent variables, complex equations of motion are obtained. A more familiar form is obtained through the substitution

$$z = \frac{1}{\sqrt{2}} (\xi + i\pi) \quad (27.29)$$

for real numbers ξ and π . Then the action becomes

$$S = \int d\tau \left\{ \frac{1}{2} \left[\pi \frac{d\xi}{d\tau} - \xi \frac{d\pi}{d\tau} \right] - h(\pi, \xi) \right\} \quad (27.30)$$

with $h(\pi, \xi) = H[(\xi - i\pi)/\sqrt{2}, (\xi + i\pi)/\sqrt{2}]$. Variation with respect to ξ

and π yields the usual Hamiltonian equations of motion. As an example of complex equations of motion, consider $H_0 = a^\dagger a + \frac{1}{2}$. The classical equations of motion are

$$\begin{aligned} i \frac{dz}{dt} &= \frac{\partial H_0}{\partial z^*} = \frac{\partial}{\partial z^*} \left[\frac{\langle z | H_0 | z \rangle}{\langle z | z \rangle} \right] = z \\ -i \frac{dz}{dt} &= z^* \end{aligned} \quad (27.31)$$

with solutions

$$z(t) = z(0)e^{-it} \quad z^*(t) = z^*(0)e^{it} \quad (27.32)$$

The principal application of the path integral (27.28) has been to field theory and in particular to the statistical mechanics of bosons—in the hope of understanding superfluid helium. The generalization to field theory and statistical mechanics is formally trivial, even though the mathematical questions raised by this generalization are difficult.

Let the boson field have one particle states labeled by an index k so that creation and annihilation operators a_k^\dagger and a_k are available. They satisfy the usual Bose commutation relations

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \quad \text{etc.} \quad (27.33)$$

($\delta_{kk'}$, may be either a Kronecker or Dirac delta function). If the modes are uncoupled (free bosons), the Hamiltonian is (ϵ_k is the energy or frequency of the k th mode)

$$H_0 = \sum_k \epsilon_k a_k^\dagger a_k \quad (27.34)$$

A “two-body interaction potential” may take the form

$$V = \lambda \sum_{k, k', q} v(q) a_{k+q}^\dagger a_{k'-q}^\dagger a_{k'} a_k \quad (27.35)$$

Let $H = H_0 + V$. The grand canonical partition function, with chemical potential μ , is ($N = \sum a_k^\dagger a_k$)

$$Z = \text{Tr } \exp(-\beta(H - \mu N)) \quad (27.36)$$

One possible complete set of states for the trace is $\{|n_{k_1}, n_{k_2}, \dots\rangle | k \text{ runs through single particle labels; } n_k = 0, 1, \dots\}$. Instead of this set, for each k

we can run through the coherent states. Consider an operator A acting in the subspace of a particular k ; its trace is (α, β label coherent states)

$$\begin{aligned}\text{Tr } A &= \sum_n \langle n | A | n \rangle = \sum_n \int \frac{d^2\alpha}{\pi} \langle n | \alpha \rangle \langle \alpha | A | n \rangle \\ &= \int \frac{d^2\alpha}{\pi} \langle \alpha | A | \alpha \rangle\end{aligned}\quad (27.37)$$

Clearly then the trace in (27.36) can be taken over the states $\{|z_{k_1}, z_{k_2}, \dots\rangle = |z_{k_1}\rangle|z_{k_2}\rangle\dots|z_{k_N}\rangle|z_{k_i}\rangle \in \mathbb{C}$, k runs through single particle labels}. The relation between the partition function and the Green's function is the usual

$$Z = \text{Tr } e^{-\beta(H - \mu N)} = \text{Tr } G(-i\beta) \quad (27.38)$$

If $\lambda=0$ ($V=0$) the modes are not coupled and the Green's function is just the product of Green's functions for individual modes. Earlier results for a single mode show the field theory Green's function to be

$$\begin{aligned}G(\{z_{fk}\}, t; \{z_{ik}\}) &= \langle z_{fk_1} \dots | \exp[-it(H_0 - \mu N)] | z_{ik_1} \dots \rangle \\ &= \int \left(\prod_k d^2 z_k(\tau) \right) \exp \left\{ i \int_0^t d\tau \sum_k \left[\frac{1}{2} i \left(z_k^* \frac{dz_k}{d\tau} - \frac{dz_k^*}{d\tau} z_k \right) - (\epsilon_k - \mu) z_k^* z_k \right] \right\} \quad (27.39)\end{aligned}$$

where we have used the notation $d^2 z_k(\tau)$ to stand for the “measure” effectively obtained by the limiting process

$$\lim_{n \rightarrow \infty} \pi^{-n} d^2 z_{k1} \dots d^2 z_{kn} \quad (27.40)$$

For the integral over τ we change variables to $s = i\tau$ whose upper limit (in calculating Z) is just β . The partition function is therefore

$$\begin{aligned}Z &= \int \prod_k \frac{d^2 z_k}{\pi} \int \prod_k d^2 z_k(s) \exp \left\{ - \int_0^\beta ds \sum_k \left[\frac{1}{2} \left(z_k^*(s) \frac{dz_k(s)}{ds} - \frac{dz_k^*(s)}{ds} z_k(s) \right) - (\epsilon_k - \mu) z_k^*(s) z_k(s) \right] \right\} \\ &\quad - \frac{dz_k^*(s)}{ds} z_k(s) \Big) + (\epsilon_k - \mu) z_k^*(s) z_k(s) \Big] \right\} \quad (27.41)\end{aligned}$$

The paths $z_k(s)$ in (27.41) begin and end ($s=0$ and β respectively) at z_k , the value of the integration variable in the first integral in (27.41). The path integral (27.41) can of course be done easily [note (27.21)] and yields the partition function for the ideal bose gas.

Exercise: Check this assertion. What is the expectation value of the number operator for the k th mode?

With the potential V [of (27.35)] one can hope to describe physical phenomena. So long as one sticks to the “easy” form of the propagator, (27.28), rather than the “very easy” form, a neat formal expression can be obtained. To get the partition function we can evaluate the trace using the same basis functions as before (but now Z does not factor into a product of integrals):

$$Z = \int \prod_k \frac{d^2 z_k}{\pi} \left(\prod_k \langle z_k | \right) \exp \left\{ -\beta [(H_0 - \mu N) + V] \right\} \left(\prod_k |z_k \rangle \right) \quad (27.42)$$

Again resolutions of the identity can be inserted, and we are led to examine terms of the form

$$\left(\prod_k \langle z_{k,j} | \right) \left\{ \sum_l (\epsilon_l - \mu) a_l^\dagger a_l + \lambda \sum_{k, k', q} v(q) a_{k+q}^\dagger a_{k'-q}^\dagger a_{k'} a_k \right\} \left(\prod_{k'} |z_{k'j-1} \rangle \right) \quad (27.43)$$

As is the case for one-body interactions, a_k (acting to the right) becomes $z_{k,j-1}$ and a_k^\dagger (appearing on the left) becomes $z_{k,j}^*$. The Hamiltonian that will appear in the integrand is therefore a function of many different z_k 's and does not have its diagonalized form. Other than this one (serious) complication, all goes through as above and Z is

$$\begin{aligned} Z = \int \prod_k \frac{d^2 z_k}{\pi} \int \prod_k d^2 z_k(s) \exp & \left\{ - \int_0^\beta ds \left[\sum_k \left[\frac{1}{2} \left(z_k^*(s) \frac{dz_k(s)}{ds} \right. \right. \right. \right. \\ & \left. \left. \left. \left. - \frac{dz_k^*(s)}{ds} z_k(s) \right) + (\epsilon_k - \mu) z_k^*(s) z_k(s) \right] \right] \\ & \left. + \lambda \sum_{k, k', q} v(q) z_{k+q}^*(s) z_{k'-q}^*(s) z_{k'}(s) z_k(s) \right] \right\} \end{aligned} \quad (27.44)$$

Some formal simplification of this expression is obtained if k is a momentum variable and $\epsilon_k = k^2/2$. Taking $\text{Re } z_k$ and $\text{Im } z_k$ as (respec-

tively) the Fourier coefficients of $\text{Re } \psi(\mathbf{r})$ and $\text{Im } \psi(\mathbf{r})$ (which in turn inherit an s dependence from z) (27.44) becomes

$$Z = \int \prod_{\mathbf{r}} d\psi(\mathbf{r}, s) \exp \left\{ - \int_0^\beta ds \int d^3 r \left[\frac{1}{2} \left(\psi^* \frac{\partial \psi}{\partial s} - \frac{\partial \psi^*}{\partial s} \psi \right) \right. \right. \\ \left. \left. - \psi^* \left(\frac{1}{2} \nabla^2 + \mu \right) \psi + \frac{1}{2} \lambda \int d^3 r' \psi^*(\mathbf{r}, s) \psi^*(\mathbf{r}', s) v(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}', s) \psi(\mathbf{r}, s) \right] \right\} \quad (27.45)$$

This form of Z is an attractive starting point for deriving the linked cluster expansion. Although we shall not discuss this, the interested reader is referred to the paper of Casher, Lurie, and Revzen mentioned in the notes below.

The real impact of the form (27.45) lies, unfortunately, in a direction that is not mathematically defensible. Suppose—and this supposition is surely false—that the functional integral can be approximated using Laplace's method, that is, the principal contribution to the integral is from that ψ which minimizes the argument of the exponent in (27.45). Let

$$\Lambda(\psi^*, \psi) = -\psi^*(\mathbf{r}, s) \left(\frac{1}{2} \nabla^2 + \mu \right) \psi(\mathbf{r}, s) \\ + \frac{1}{2} \lambda \int d^3 r' \psi^*(\mathbf{r}, s) \psi^*(\mathbf{r}', s) v(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}', s) \psi(\mathbf{r}, s) \quad (27.46)$$

then the variational equations are

$$\frac{\partial \psi}{\partial s} = -\frac{\partial \Lambda}{\partial \psi^*}, \quad \frac{\partial \psi^*}{\partial s} = -\frac{\partial \Lambda}{\partial \psi} \quad (27.47)$$

Since the functions $z_k(s)$ begin and end at the same value, the functions ψ have this property too. Thus if there is a ψ such that

$$\frac{\partial \Lambda}{\partial \psi} = 0 \quad (27.48)$$

then it provides a minimum, since $\partial \psi / \partial s = 0$ is consistent with the boundary conditions.

Anyone who has seen the Landau-Ginzburg equation for superfluid helium ought to be greeting the foregoing few equations as old friends. In the Landau-Ginzburg theory the wave function of the superfluid component is the complex order parameter and satisfies an equation exactly of the form (27.48) [or (27.47) for dynamical problems] with "Hamiltonian" looking very much like (27.46). (Usually $v(\mathbf{r}, \mathbf{r}')$ is taken to be a δ -function.)

We have therefore, in a sense, derived the Landau-Ginzburg theory and identified the order parameter as the coherent state wave function. A fundamental theoretical problem in the study of superfluidity is to connect the two-fluid phenomenology to an underlying microscopic dynamics. According to the lore, the superfluid density is the magnitude squared of the complex order parameter and that order parameter satisfies the Landau-Ginzburg equation. Were it not for the nagging mathematical difficulties encountered along the way, the present theory would have brought us close to justifying the Landau-Ginzburg equation and making the macro-micro connection.

Laplace's method is the weakest link in our reasoning. (In physics literature the terms Laplace's method, method of steepest descent, and method of stationary phase are often used interchangeably.) This simple approximation to the functional integral has given us a mean field theory—the Landau-Ginzburg theory—and in fact is a good example of the recurrent phenomenon that Laplace's method and mean field theory are in many instances the same approximation. What vitiates the Laplace method is that near the minimum selected by the differential equation (27.48) are many other minima. Each such minimum can be interpreted as a mode of the system. The large parameter for the asymptotic approximation is the volume V . Unfortunately as $V \rightarrow \infty$ the other minima crowd around the lowest (the spacing goes to zero as $V \rightarrow \infty$) and the integral can never be approximated by a single minimum. Moreover, there is the further possibility that this multitude of modes recouples and the integral looks like an infinite dimensional version of the degenerate critical points of Sections 15 and 16. This problem will occupy us elsewhere too.

But I do not wish to leave the reader in a state of depression. The mean field theory works. It covers a good range of physical phenomena, and its recovery from the functional integral represents a small but real step forward in statistical mechanics.

The functional integral in a coherent state basis is also useful for seeing how quantum statistics enter the partition function. We return to considering a single variable z and rewrite (27.28) with the change of variable $s = \tau/t$

$$\begin{aligned} G(z'', t; z') &= \int_{z(0)=z'}^{z(1)=z''} d^2 z(s) \exp \left\{ \int_0^1 ds \left[\frac{1}{2} \left(z \frac{dz^*}{ds} - z^* \frac{dz}{ds} \right) - itH(z^*(s), z(s)) \right] \right\} \\ &= \langle z'' | \exp(-itH) | z' \rangle \end{aligned} \quad (27.49)$$

First observe that for $t=0$ we still have a functional integral, but it must collapse to the identity, since G must become the matrix element $\langle z'' | z' \rangle$.

Expectations of more general objects than $\exp(-iHt)$ can also be obtained by inserting resolutions of the identity, and we have for example

$$\begin{aligned} \langle z'' | a \exp(-iHt) a^\dagger | z' \rangle \\ = \int d^2 z(s) \left\{ \exp \int_0^1 ds \left[\frac{1}{2} \left(z \frac{dz^*}{ds} - z^* \frac{dz}{ds} \right) - itH \right] \right\} z(s_1) z^*(s_2) \end{aligned} \quad (27.50)$$

where $1 \geq s_1 > s_2 \geq 0$ (but s_1 , and s_2 are otherwise arbitrary) and the arguments of all z 's in H are between s_1 and s_2 . Letting $t=0$ gives

$$\langle z'' | aa^\dagger | z' \rangle = \int d^2 z(s) z(s_1) z^*(s_2) \exp \left(\int_0^1 \frac{1}{2} \left(z \frac{dz^*}{ds} - z^* \frac{dz}{ds} \right) ds \right) \quad (27.51)$$

This expression and the corresponding one for $a^\dagger a$ are enough to yield the commutation relation for a and a^\dagger . Rather than go through the details we note that because most of the integrals in (27.51) collapse, a single resolution of the identity suffices. The expressions for $a^\dagger a$ and aa^\dagger are

$$\langle z'' | a^\dagger a | z' \rangle = \int \frac{d^2 z}{\pi} \langle z'' | z \rangle z''^* z' \langle z | z' \rangle \quad (27.52)$$

$$\langle z'' | aa^\dagger | z' \rangle = \int \frac{d^2 z}{\pi} \langle z'' | z \rangle |z|^2 \langle z | z' \rangle \quad (27.53)$$

In (27.52) the quantity $z''^* z'$ factors out of the integral. In (27.53) $|z|^2$ must be integrated. The identity

$$\begin{aligned} \int \frac{d^2 \eta}{\pi} |\eta|^2 \exp(-|\eta|^2 + \xi^* \eta + \eta^* \xi) \\ = (1 + \xi^* \xi) \exp(\xi^* \xi) \end{aligned} \quad (27.54)$$

(for complex numbers ξ, η, ζ) shows that, besides the term $z''^* z'$, (27.53) will have the additional contribution 1, or $\langle z'' | 1 | z' \rangle$. The difference between (27.53) and (27.52) is just the commutation relation. For the record we give the simpler identity

$$\int \frac{d^2 \eta}{\pi} \exp(-|\eta|^2 + \xi^* \eta + \eta^* \xi) = \exp(\xi^* \xi) \quad (27.55)$$

on which (27.52) is based.

We now have a consistent picture of the role played by the term

$$z^* \frac{dz}{ds} - z \frac{dz^*}{ds}$$

Given a Hamiltonian $H(p, q)$ the classical partition function is

$$Z_{\text{classical}} = \int dp dx e^{-\beta H} \quad (27.56)$$

The quantum partition function is obtained by an integral over d^2z —which we have noted earlier is closely related to $dp dx$. Thus

$$Z_{\text{quantum}} = \int \frac{1}{\pi} d^2z \exp \left\{ - \int_0^1 ds \left[\frac{1}{2} \left(z^* \frac{dz}{ds} - z \frac{dz^*}{ds} \right) + \beta H \right] \right\} \quad (27.57)$$

The s dependence of z , with the weighting factor $\exp[-\frac{1}{2} \int_0^1 ds (z^* dz/ds - z dz^*/ds)]$ manages to change the classical to the quantum expression for bosons.

Is there a similar way to handle fermions? There are hints that some other factor might do the job for Fermi statistics. There is also the work of Klauder (see notes) in which he developed parallel coherent state path integrals for both bosons and fermions. Fermi statistics enter in his picture when ambiguities arise in how to sum certain “unruly histories” (to use his term). An appropriate resolution of the ambiguity yields Fermi statistics. Actually it has turned out that field theory histories are far more unruly than anyone would have imagined at the time of Klauder’s work. This problem will be discussed when we treat field theory in general.

The present status of the fermion path integral problem is...in my opinion...open. In fact a simple, intuitive fermion theory, mathematically no worse than Feynman’s original path integral, would be one of the most welcome contributions to the field (see also Section 32.9).

NOTES

Much of the material in this section can be found in the following papers:

- A. Casher, D. Lurie, and M. Revzen, Functional Integration for Many-Boson Systems, *J. Math. Phys.* **9**, 1312 (1968)
- S. S. Schweber, On Feynman Quantization, *J. Math. Phys.* **3**, 831 (1962)
- J. S. Langer, Coherent States in the Theory of Superfluidity, *Phys. Rev.* **167**, 183 (1968)

Coherent states were first used in path integration in the paper

J. R. Klauder, The Action Option and a Feynman Quantization of Spinor Fields in Terms of Ordinary C-Numbers, *Ann. Phys.* **11**, 123 (1960)

More on coherent states can be found in

I. R. Senitzky, *Phys. Rev.* **95**, 1115 (1954)

R. J. Glauber, *Phys. Rev.* **131**, 2766 (1963)

A. Messiah, *Quantum Mechanics*, Wiley, New York, 1961

“Generalized” coherent states were defined by

A. M. Perelomov, Coherent States for Arbitrary Lie Group, *Commun. Math. Phys.* **26**, 222 (1972)

Their use as basis states in the path integral can be found in

J. R. Klauder, Path Integrals and Stationary-Phase Approximations, *Phys. Rev. D* **19**, 2349 (1979)

H. Kuratsuji and Toru Suzuki, Path Integral in the Representation of SU(2) Coherent State and Classical Dynamics in a Generalized Phase Space, *J. Math. Phys.* **21**, 472 (1980)

The Hilbert space of analytic functions that helps elucidate many of the questions concerned with overcomplete sets of basis functions is studied in

V. Bargmann, *Commun. Pure and Appl. Math.* **14**, 187 (1961)

V. Bargmann, On the Representations of the Rotation Group, *Rev. Mod. Phys.* **34**, 829 (1962)

The overcompleteness of the set $\{|n+im\rangle| n, m = \text{integers}\}$ was found by

A. M. Perelomov, *Teor. Mat. Fiz.* **6**, 213 (1971)

and some of the mathematical consequences are developed in

M. Boon and J. Zak, Discrete Coherent States on the von-Neumann Lattice, *Phys. Rev. B* **18**, 6744 (1978)

where it turns out that the Jacobi theta function, which we met earlier in connection with spin, plays an important role. Pseudospin, isospin, and other appearances of $SU(2)$ in many-body theory (rather than in connection with physical rotations) are discussed in

Harry J. Lipkin, *Lie Groups for Pedestrians*, North Holland, Amsterdam (and Wiley-Interscience), 1965, Ch. 2

J. J. Sakurai, *Invariance Principles and Elementary Particles*, Princeton University Press, Princeton, N.J., 1964, Sec. 9.3

P. W. Anderson, *Concepts in Solids*, Benjamin, New York, 1963

Papers that discuss, debate, and dispute the ordering problem in quantum mechanics and in particular the relevance of path integrals to that issue are

Leon Cohen, *J. Math. Phys.* **11**, 3296 (1970)

E. Kerner and W. Sutcliffe, *J. Math. Phys.* **11**, 391 (1970)

F. J. Testa, *J. Math. Phys.* **12**, 1471 (1971)

I. W. Mayes and J. S. Dowker, *J. Math. Phys.* **14**, 434 (1973)

M. M. Mizrahi, *J. Math. Phys.* **16**, 2201 (1975)

J. S. Dowker, Path Integrals and Ordering Rules, *J. Math. Phys.* **17**, 1873 (1976)

A. C. Hirshfeld, Canonical and Covariant Path Integrals, *Phys. Lett.* **67A**, 5 (1978)

While I agree with Dowker that path integrals will not fix operator ordering any more than any other kind of quantization, certain ordering rules may still emerge in a "more natural" or "simpler" fashion as a result of "more natural" or "simpler" ways of doing the path integral.

Derivation of the commutation relations from the coherent state path integral and various other applications have been considered by Revzen in the papers

M. Revzen, Functional Integrals in Statistical Physics, *Am. J. Phys.* **38**, 611 (1970)

M. Revzen and L. S. Schulman, *J. Stat. Phys.* **8**, 217 (1973)

TWENTY-EIGHT

Systems with Random Impurities

Natural systems never quite correspond to the theoretical idealizations physicists find convenient. Lattices are never perfectly periodic, metals never pure. Sometimes these deviations are unimportant and sometimes they themselves are the effect to be studied. Impurity problems can be usefully attacked using functional integrals because even for impurities, namely the departures from one sort of idealization, there is a theoretical idealization to simplify the description of their properties. Specifically, impurities are usually taken to be randomly distributed, and it is this randomness that finds expression as a functional integral.

As a first example we consider Bose-Einstein condensation. If N noninteracting bosons are in an otherwise empty box of volume V , then below a certain transition temperature a finite (i.e., much larger than $1/V$) fraction of them will be in the quantum mechanical ground state. This phase transition has been important in attempts to understand superfluidity in He^4 . Moreover, in some recent work an analogous Bose-Einstein condensation has been shown to occur in a large class of systems so that Bose-Einstein condensation may turn out to be the archetypal phase transition. This is not the place for speculations on this point; however, it is clear that the various idealizations involved in the simplest derivation of Bose-Einstein condensation are worth examining to see whether the phenomenon persists in a more realistic physical model. The most important idealization is that the particles are noninteracting, but we will not try anything so difficult as dropping that assumption. Rather, we shall allow the volume to have impurities randomly distributed in it and consider their effect on each boson separately. In other words, we are adding a one-particle interaction to the Hamiltonian.

Thus we have N noninteracting bosons in a cube of side L , volume $V=L^3$. Within the cube are M impurities, idealized as spheres of radius δ , whose centers are uniformly distributed throughout V (so that two spheres can overlap). Periodic boundary conditions are adopted for both particles and impurities; a sphere whose center is within δ of a face of the cube will “reenter” at the other side of the cube. The one-particle Hamiltonian for the bosons is therefore

$$h=h_0+U=\frac{p^2}{2m}+U \quad (28.1)$$

where m is mass and U is an infinite potential located throughout each impurity. Alternatively, we can take $h=h_0$ but require the wave function ψ to satisfy

$$\psi=0 \quad \text{on} \quad \partial S_k(\delta) \quad k=1, \dots, M \quad (28.2)$$

where $S_k(\delta)$ is the k th sphere (of radius δ) and $\partial S_k(\delta)$ is its boundary. The two descriptions are equivalent.

For a given configuration C of spheres let the spectrum of the Hamiltonian be given by $E_j^{(C)}$, $j=1, \dots$ with $E_j^{(C)} < E_{j+1}^{(C)}$. For noninteracting bosons the grand canonical partition function can be written down, and the mean number of particles in state j is

$$\bar{n}_j = \frac{1}{\exp[\beta(E_j^{(C)} - \mu)] - 1} \quad (28.3)$$

where the chemical potential μ is obtained by solving

$$N = \sum_j \bar{n}_j \quad (28.4)$$

and as usual $\beta=1/k_B T$ with k_B the Boltzmann constant. It is also convenient to define the fugacity, $\xi=\exp(\beta\mu)$. Note that for (28.3) to make sense (for $j=1$) we must have

$$\exp(\beta E_1^{(C)}) > \xi \quad (28.5)$$

By defining

$$Q(t) = \frac{1}{V} \sum_j \exp(-tE_j^{(C)}) \quad (28.6)$$

a series of simple manipulations on (28.4) leads to

$$N = V \sum_{l=1}^{\infty} \xi^l Q(l\beta) \quad (28.7)$$

or

$$\rho = \sum_{l=1}^{\infty} \xi^l Q(l\beta) \quad (28.8)$$

The usual Bose-Einstein condensation occurs because for sufficiently low temperatures the sum on the right hand side of (28.7), *excluding* contributions to Q from $E_1^{(C)}$, is bounded, while N can be made arbitrarily large. Hence the terms involving $E_1^{(C)}$, which together comprise \bar{n}_1 , must become large and in fact, yield a finite fraction of N . Let us see how this is manifested in the behavior of Q itself. Let Q_0 be the quantity defined in (28.6) with no impurities ($M=0$). Using the free particle energy spectrum for $\{E_j\}$, the sum in (28.6) becomes an integral and

$$VQ_0(t) = 1 + \frac{V}{(2\pi)^3} \int d^3k \exp\left(-\frac{\hbar^2 k^2 t}{2m}\right) = 1 + V \left(\frac{m}{2\pi\hbar^2 t}\right)^{3/2} \quad (28.9)$$

where the $E_1=0$ term in Q_0 is included separately as 1. Inserting Q_0 into (28.7) yields

$$N = \frac{\xi}{1-\xi} + V \left(\frac{m}{2\pi\hbar^2\beta}\right)^{3/2} \sum_{l=1}^{\infty} \frac{\xi^l}{l^{3/2}} \quad (28.10)$$

with the first term on the right corresponding to \bar{n}_1 . For ξ satisfying (28.5), which in this case means $\xi < 1$, the series on the right hand side of (28.10) converges and remains finite for $\xi = 1$. There will therefore be an N large enough and a $1/\beta$ small enough that the only way to satisfy (28.10) is to have the term $\xi/(1-\xi)$ contribute significantly. Note that ostensibly \bar{n}_1 is not proportional to V , while (for fixed density) the other terms in (28.10) are. For high temperature \bar{n}_1 plays a negligible role in (28.10). However, we have now seen that for low enough temperature \bar{n}_1 must be large and of order V . (Recalling that $\xi = \exp(\beta\mu)$, this shows that $\mu = O(1/V)$ so that the chemical potential gets very close to zero.) The state 1 is therefore macroscopically occupied and we have Bose-Einstein condensation.

There are various ways of showing that there continues to be a phase transition in the presence of impurities. One fact that played a role in

drawing conclusions from (28.9) was that the ground state energy is zero so that $Q_0(l\beta)$ does not drop off with l . Hence one approach to the impurity problem is to show that the ground state energy approaches zero in probability.

The method that we use to show the existence of a phase transition is to show that $\rho(\zeta)$, as given by (28.8), has a singularity for $\zeta \rightarrow 1$. Unfortunately, this is not the same as proving Bose-Einstein condensation, which would require the stronger result that

$$\lim_{V \rightarrow \infty} V(E_2^{(C)} - E_1^{(C)}) = \infty \quad (28.11)$$

in probability. When (28.11) obtains there is macroscopic occupation of the ground state, but we shall not prove that result. The remainder of our discussion of the topic is devoted to showing the divergence of the series (28.8).

The quantity $Q(t)$ has the structure of a partition function, and we therefore have a path integral—or more correctly in this case, Wiener integral—expression for it. Thus

$$Q(t) = \frac{1}{V} \text{Tr } e^{-tH} = \frac{1}{V} \int d^3r G(\mathbf{r}, -it; \mathbf{r}) \quad (28.12)$$

Recall from (1.24) the basic expression for G :

$$\begin{aligned} G(\mathbf{r}, T; \mathbf{r}) &= \lim_{N \rightarrow \infty} \int d^3x_1 \cdots d^3x_{N-1} \left(\frac{m}{2\pi\hbar i\epsilon} \right)^{3N/2} \exp \left\{ \frac{i\epsilon}{\hbar} \sum \left[\frac{m}{2} \left(\frac{\Delta \mathbf{x}_j}{\epsilon} \right)^2 - V \right] \right\} \\ &\quad (28.13) \end{aligned}$$

with $\epsilon = T/N$, $\Delta \mathbf{x}_j = \mathbf{x}_{j+1} - \mathbf{x}_j$, and $\mathbf{x}_0 = \mathbf{x}_N = \mathbf{r}$. Putting $T = -it$ and defining $\epsilon' = t/N$ we have

$$\begin{aligned} G(\mathbf{r}, -it; \mathbf{r}) &= \lim_{N \rightarrow \infty} \int d^3x_1 \cdots d^3x_{N-1} \left(\frac{m}{2\pi\hbar\epsilon'} \right)^{3N/2} \exp \left\{ -\frac{\epsilon'}{\hbar} \sum \left[\frac{m}{2} \left(\frac{\Delta \mathbf{x}}{\epsilon'} \right)^2 + V \right] \right\} \\ &\quad (28.14) \end{aligned}$$

This expression is just a Wiener integral. By comparing (28.14) to (9.15) it follows that

$$G(\mathbf{r} + \mathbf{x}, -it; \mathbf{r}) = E_{\mathbf{x}, t} \left(\exp \left(-\frac{1}{\hbar} \int_0^t V(x(\tau)) d\tau \right) \right) \equiv W(\mathbf{x}, t) \quad (28.15)$$

with the diffusion coefficient D implicit in $E_{\mathbf{x}_t}$ taken to be $D=\hbar/2m$ [see (9.11)]. To obtain this result we did not have to go through (28.13) [alias (1.24)], but could have argued from the analytic continuation of the partial differential equations satisfied by G and W [cf. (9.30)].

At this stage it is desirable to write Q or G as an expectation value, but we recall from Section 9 that $E_{\mathbf{x}_t}$ is not properly normalized, a fact reflected in

$$E_{\mathbf{x}_t}(1) = (4\pi Dt)^{-3/2} \exp\left(-\frac{\mathbf{x}^2}{4Dt}\right) \quad (28.16)$$

(The $\frac{3}{2}$ power appearing instead of the square root in (9.16) occurs because we are now working in three dimensions.) Let $\mathcal{E}_{\mathbf{x}_t}$ be the normalized expectation, so that if A is a function on the path space

$$\mathcal{E}_{\mathbf{x}_t}(A) = E_{\mathbf{x}_t}(A)/E_{\mathbf{x}_t}(1) \quad (28.17)$$

We require G of (28.15) for the case $\mathbf{x}=0$ so that

$$Q(t) = \frac{1}{V} \int d^3r (4\pi Dt)^{-3/2} \mathcal{E}_{0t} \left[\exp\left(-\int_0^t U(\mathbf{x}(\tau)) d\tau\right) \right] \quad (28.18)$$

with $D=\hbar/2m$. The “potential” U is infinite within the volume of each of the impurity spheres and zero elsewhere. Hence the expectation in (28.18) is simply the (Wiener) measure or probability of those Brownian motion paths that start and end at \mathbf{r} , but never enter any of the impurity regions. Thus

$$\begin{aligned} Q(t) &= \frac{1}{V} \int d^3r (4\pi Dt)^{-3/2} \\ &\times \text{Prob}\{\mathbf{r} + \mathbf{r}_V(\tau) \notin S_k(\delta), \\ &k=1, \dots, M, \quad 0 < \tau < t | \mathbf{r}_V(0) = \mathbf{r}_V(t) = 0\} \end{aligned} \quad (28.19)$$

where $\mathbf{r}_V(\tau)$ begins and ends at 0 in V and because of the boundary conditions on V it is assumed to reenter V on an opposite face in case it passes out of V (i.e., V is a torus).

There is a second kind of expectation value to be taken here, the average over configurations C , denoted $\langle \rangle_C$. The impurities are randomly distributed with a uniform distribution; this condition together with our treatment of V as a torus (through the reentry at opposite faces of both \mathbf{r}_V and the spherical impurities) imply that

$$\langle \text{Prob}\{\mathbf{r} + \mathbf{r}_V \notin S_k | \mathbf{r}_V(0) = \mathbf{r}_V(t) = 0\} \rangle_C \quad (28.20)$$

is independent of \mathbf{r} . Hence the integral $\int d^3r$ in (28.19) is trivial and merely cancels the V . It remains to evaluate (28.20) for $\mathbf{r}=0$.

For a given path $\mathbf{r}_V(\tau)$, $0 \leq \tau \leq t$, consider the union of the set of spheres of radius δ with centers at each point of $\mathbf{r}_V(\tau)$. That is, define

$$W_{\delta,t}(\mathbf{r}(\tau)) = \{\mathbf{x} \in V \mid |\mathbf{x} - \mathbf{r}(\tau)| < \delta \quad \text{for some } \tau, \quad 0 \leq \tau \leq t\} \quad (28.21)$$

where the distance $|\mathbf{x} - \mathbf{r}(\tau)|$ is understood to mean distance on the torus. W is a kind of sausage (the Wiener-Sausage??) and we denote its three-dimensional volume by $|W_{\delta,t}|$. For a given path $\mathbf{r}(\tau)$ it is clear that it will not enter any of the spheres $S_k(\delta)$ if those spheres lie outside $W_{\delta,t}(\mathbf{r}(\tau))$. The probability that this will happen for uniformly distributed impurities (hence the average over C) is obviously

$$\left[1 - \frac{|W_{\delta,t}(\mathbf{r}(\tau))|}{V} \right]^M \quad (28.22)$$

As usual it is the thermodynamic limit $V \rightarrow \infty$ that is of interest, and we shall assume that the number of impurities per unit volume goes to a finite limit, that is,

$$\nu = \frac{M}{V}, \quad \text{both } M \text{ and } V \rightarrow \infty \quad (28.23)$$

For this limit the expression (28.22) is simply

$$\exp[-\nu |W_{\delta,t}(\mathbf{r}(\tau))|] \quad (28.24)$$

with the path $\mathbf{r}(\tau)$ now an ordinary Brownian motion path. The probability of (28.20) is therefore the expectation over the space of Brownian motion paths of (28.24). Therefore

$$\langle Q(t) \rangle_{(C)} = (4\pi Dt)^{-3/2} \mathcal{E}_{0,t} \{ \exp[-\nu |W_{\delta,t}(\mathbf{r}(\tau))|] \} \quad (28.25)$$

with $\mathbf{r}(t) = \mathbf{r}(0)$.

Our purpose in evaluating Q is to study the series (28.8); the expression (28.25) however, deals with $\langle Q \rangle_{(C)}$, not Q . To establish that Q can be replaced by its average we now show that in the thermodynamic limit Q does not fluctuate. That is, the relation $\langle Q^2 \rangle_{(C)} = \langle Q \rangle_{(C)}^2$ is true, as are similar higher-order relations. The expression $\langle Q^2 \rangle_{(C)}$ will involve two integrals in place of the one $\int d^3r$ occurring (e.g.) in (28.19). There appear therefore two Brownian motion paths (which are themselves dummy variables). Because the integration volume is large ($V \rightarrow \infty$) these paths are

mostly far apart, and there is little overlap of their respective sets $W_{\delta,t}$. This turns out to be enough to establish that $\langle Q^2 \rangle_{(C)} = \langle Q \rangle_{(C)}^2$.

Our study of this system is therefore reduced to study of the series

$$\begin{aligned} \rho &= \sum_{l=1}^{\infty} \xi^l \langle Q(l\beta) \rangle_{(C)} \\ &= (4\pi D\beta)^{-3/2} \sum_{l=1}^{\infty} \xi^l \frac{\mathcal{E}_{0,l\beta}\{\exp(-\nu|W_{\delta,l\beta}|)\}}{l^{3/2}} \end{aligned} \quad (28.26)$$

The existence of a phase transition will now be demonstrated by showing that the series (28.26) has a singularity at $\xi = 1$.

To estimate $\mathcal{E}_{0,t}\{\exp(-\nu|W_{\delta,t}|)\}$ we use the inequality

$$\begin{aligned} \text{Prob}\{r_V(\tau) \notin S_k(\delta), k=1, \dots, M, 0 < \tau < t | r_V(t) = r_V(0) = 0\} \\ > \text{Prob}\{r_V(\tau) \in C(a), 0 < \tau < t | r_V(t) = r_V(0) = 0\} \\ \cdot \text{Prob}\{\text{centers of all } S_k(\delta) \text{ lie outside } C(a+\delta)\} \end{aligned} \quad (28.27)$$

where $C(b)$ is a cube of side b centered at 0. To me this inequality seems a bit gross, but surprisingly it is strong enough to give the phase transition; heuristics for this anomaly will be appended below. In any case the last probability of (28.27) is greater than

$$\left[1 - \frac{(a+\delta)^3}{V} \right]^M \quad (28.28)$$

The first probability on the right hand side of (28.27) is just the probability that $r(\tau)$ remains in $C(a)$ throughout. This is clearly related to the first passage time. However, we need not appeal to any results of the theory of first passage times, but merely observe that this probability is the same as

$$p = \mathcal{E}_{0,t} \left[\exp \left(- \int_0^t V_a(\mathbf{x}(\tau)) d\tau \right) \right] \quad (28.29)$$

where V_a is infinite outside $C(a)$ and zero inside. But $\mathcal{E}_{0,t}$ is related to a Green's function and we have

$$\begin{aligned} p &= \frac{E_{0,t} \left(\exp \left(- \int_0^t V_a d\tau \right) \right)}{E_{0,t}(1)} \\ &= (4\pi D t)^{3/2} G_a(0, -it; 0) = (4\pi D t)^{3/2} \sum e^{-\lambda_n t} |\psi_n(0)|^2 \end{aligned} \quad (28.30)$$

with ψ_n, λ_n eigenfunctions and eigenvalues of the Schrödinger equation for a particle in an infinite square well at $\pm a/2$. The three dimensions separate and for the one-dimensional well

$$\lambda_n = \frac{\hbar^2}{2m} n^2 \frac{\pi^2}{a^2} \quad \psi_n(0) = \begin{cases} \pm \sqrt{2/a} & n \text{ odd} \\ 0 & n \text{ even} \end{cases} \quad (28.31)$$

It follows that

$$\begin{aligned} p &= (4\pi Dt)^{3/2} \left[\frac{2}{a} \sum_{j=0}^{\infty} \exp \left[\frac{-\hbar^2(2j+1)^2 \pi^2 t}{2ma^2} \right] \right]^3 \\ &> \left(\frac{16\pi Dt}{a^2} \right)^{3/2} \exp \left(\frac{-3\hbar^2 \pi^2 t}{2ma^2} \right) \end{aligned} \quad (28.32)$$

Until now a has been arbitrary. We set

$$a = t^{1/5} \quad (28.33)$$

and combine the foregoing results to yield

$$\mathbb{E}_{0,l\beta} [\exp(-\nu|W_{\delta,l\beta}|)] > e^{-\nu(t^{1/5} + \delta)^3} t^{9/10} e^{-\frac{3\hbar^2 \pi^2}{2m} t^{3/5}} > e^{-ct^{3/5}} \quad (28.34)$$

for some constant c and $t = l\beta$. Therefore

$$\rho > \text{const.} \sum_{l=1}^{\infty} \frac{\xi^l}{l^{3/2}} e^{-cl^{3/5}} \quad (28.35)$$

For $\xi > 1$ the series in (28.35) diverges, while for $\xi < 1$ it converges, confirming the claim that the radius of convergence of the series in (28.26) is unity (since $|W_{\delta,l\beta}| \geq 0$, (28.26) obviously converges for $\xi < 1$). Moreover, since all terms in (28.26) are positive, the singularity is actually at $\xi = 1$.

I believe that the reason we could get away with the gross inequality (28.27) is that all that needs proving is that *somewhere* in the infinitely large V there is a large region with no impurities. Thus you can always find a big fluctuation in the uniform distribution of impurities.* The same feature is what makes the ground state energy $E_1^{(C)}$ approach zero in probability, a

*Although we have only proved an inequality, the correct behavior is in fact the same $t^{3/5}$ that results from our “gross” estimates. This is because the main contributors to the object we are estimating are the *large deviations*. It is not $\mathbb{E}[|W_{\delta,t}|]$ that we study, but rather $\mathbb{E}[\exp(-\nu|W_{\delta,t}|)]$ and the exponentiation gives large weight to those rare paths for which $|W_{\delta,t}|$ is very small. See Donsker and Varadhan (reference below) for more on this.

fact that plays a role in an alternate proof of the existence of a phase transition in a system of bosons with impurities.

A second approach to systems with impurities is to introduce a potential that is itself a random variable and then do a functional integration over the space of potentials, postulating some distribution function for the V 's. (By contrast, in our previous example the random impurities had a uniform distribution with no correlations between them. Because of translational invariance averaging over the impurities was trivial.) Quite a bit of work has been done on this problem, and we shall only give a brief introduction without getting into the physics.

Consider an electron in a solid. In the absence of impurities we assume for simplicity its motion to be free so that the Hamiltonian is

$$H = -\frac{1}{2} \nabla^2 + V(\mathbf{r}) \quad (28.36)$$

where V is the impurity potential and is itself a random variable whose distribution we shall take to be Gaussian. Specifically for A a function on the space of potentials

$$\langle A \rangle = \mathcal{N} \int dV(\cdot) \exp \left[-\frac{1}{2} \int d^3r d^3r' V(\mathbf{r}) K(\mathbf{r} - \mathbf{r}') V(\mathbf{r}') \right] A[V(\cdot)] \quad (28.37)$$

where K is the kernel that characterizes the distribution and \mathcal{N} a normalization so that $\langle 1 \rangle = 1$. It follows that

$$\langle V(\mathbf{r}) \rangle = 0$$

(taking $A[V(\cdot)] = V(\mathbf{r})$ for a specific value of \mathbf{r}). The correlation function for the potential is given by

$$\langle V(\mathbf{r}) V(\mathbf{r}') \rangle = w(\mathbf{r} - \mathbf{r}') \quad (28.38)$$

where w satisfies

$$\int d^3r' K(\mathbf{r} - \mathbf{r}') w(\mathbf{r}' - \mathbf{r}'') = \delta(\mathbf{r} - \mathbf{r}'') \quad (28.39)$$

Proof of (28.38) involves Gaussian integrals of a familiar kind. One can also evaluate the integral by the formal trick of ignoring the fact that the kernel in (28.37) depends only on $\mathbf{r} - \mathbf{r}'$ and letting the kernel be instead

$K(\mathbf{r}, \mathbf{r}')$. One then notes that

$$\begin{aligned} \mathcal{N} \int dV(\cdot) V(\mathbf{r}) V(\mathbf{r}') \exp \left[-\frac{1}{2} \int K(\mathbf{r}, \mathbf{r}') V(\mathbf{r}) V(\mathbf{r}') d^3 r d^3 r' \right] \\ = -\mathcal{N} \frac{\partial}{\partial K(\mathbf{r}, \mathbf{r}')} \int dV(\cdot) \exp \left[-\frac{1}{2} \int KV V \right] \\ = -\frac{\partial}{\partial K(\mathbf{r}, \mathbf{r}')} \log \int dV(\cdot) \exp \left[-\frac{1}{2} \int KV V \right] \end{aligned} \quad (28.40)$$

(The normalization is just the inverse of the Gaussian integral, hence the log.) This is a procedure often used in statistical mechanics for the evaluation of correlation functions. The Gaussian integral in (28.40) is the determinant of K , and the derivative with respect to a given matrix element is the minor, hence an element of the inverse matrix of K . Hence (28.38) and (28.39).

The physical problem here is to find the expected density of states, which involves averaging over the Green's function. Formally

$$G(E) = \frac{-i}{H-E-i\epsilon} = P \frac{-i}{H-E} - i(+i\pi)\delta(H-E) \quad (28.41)$$

(using $(x \pm ie)^{-1} = P(1/x) \mp i\pi\delta(x)$) so that

$$\rho(E) = \frac{1}{\pi} \operatorname{Re} G(\mathbf{r}, \mathbf{r}; E) \quad (28.42)$$

is the density of states at \mathbf{r} . Since the distribution function for V , K is translationally invariant (depending only on $\mathbf{r}-\mathbf{r}'$) the average of ρ will in fact be independent of \mathbf{r} . The quantity of physical interest is therefore

$$n(E) = \langle \rho(E) \rangle = \frac{1}{\pi} \operatorname{Re} \left\langle \int_0^\infty dt e^{iEt} G(\mathbf{r}, t; \mathbf{r}) \right\rangle \quad (28.43)$$

(with $\operatorname{Im} E > 0$). When G in (28.43) is written as a path integral it is seen that only the potential term is subject to the bracket. (Note that two independent functional integrals are being done here: one for $\mathbf{r}(t)$, the other for V). Hence

$$n(E) = \frac{1}{\pi} \operatorname{Re} \int_0^\infty dt e^{iEt} \int d\mathbf{r}(\cdot) \exp \left(i \int_0^t ds \frac{1}{2} m \dot{\mathbf{r}}^2 \right) \left\langle \exp \left(i \int_0^t V(\mathbf{r}(s)) ds \right) \right\rangle \quad (28.44)$$

But the bracket in (28.44) can be evaluated, since it is just another Gaussian integral. In principle it is no worse than evaluating $\int dy \exp(-y^2 + by)$, but a slight complication arises because it is not quite in the form where we could simply apply, say, (6.41) to it. For a quick formal proof it is convenient to define a function $f(\mathbf{r}'; \mathbf{r}(\cdot))$ such that

$$\int d^3 r' f(\mathbf{r}'; \mathbf{r}(\cdot)) V(\mathbf{r}') = \int_0^t ds V(\mathbf{r}(s))$$

and it follows that

$$\left\langle \exp\left(i \int_0^t V(\mathbf{r}(s)) ds\right) \right\rangle = \exp\left[\frac{1}{2} \int ds ds' w(\mathbf{r}(s) - \mathbf{r}(s'))\right] \quad (28.45)$$

At this stage all characteristics of the potential are expressed in terms of the correlation function, and when (28.45) is inserted in (28.44) we have a path integral with a “nonlocal interaction” similar to what results from the polaron calculation. Whether or not it is worth bringing $n(E)$ into the form above depends on what can be done with the path integral over r with a particular choice of w . Some authors have taken w to be a Gaussian and for $w \sim \exp(-(\mathbf{r} - \mathbf{r}')^2/L^2)$ the expansion of this exponential leads to yet another Gaussian functional integral for the evaluation of $n(E)$. We shall not pursue this further; some of the literature is noted below. I think though that one should commend the resourcefulness of those who succeed in making all problems reduce to Gaussian integrals.

NOTES

Our treatment of phase transitions for bosons in the presence of impurities follows very closely the following papers:

M. Kac and J. M. Luttinger, *J. Math. Phys.* **14**, 1626 (1973)

M. Kac and J. M. Luttinger, *J. Math. Phys.* **15**, 183 (1974)

A number of points touched very lightly in the text are treated in more detail in these papers.

In the text, we were content with an inequality for $|W_{\delta,t}|$, (28.34). In fact it is true that

$$\lim_{t \rightarrow \infty} t^{-3/5} \log \mathcal{E}_{0,t} [\exp(-\nu |W_{\delta,t}|)]$$

exists and is nonzero. This is shown in

M. Donsker and S. R. S. Varadhan, Asymptotics for the Wiener Sausage, *Comm. Pure. Appl. Math.* **27**, 525 (1975)

Random potentials are important in many physical situations. Some path integral treatments are listed below:

S. F. Edwards and V. V. Gulyaev, *Proc. Phys. Soc.* **83**, 495 (1964)

J. Zittarz and J. S. Langer, *Phys. Rev.* **148**, 741 (1966)

R. Jones and T. Lukes, *Proc. Roy. Soc. A* **309**, 457 (1969)

V. Bezak, *Proc. Roy. Soc. Lond. A* **315**, 339 (1970)

M. Fibich and E. Helfand, *Phys. Rev.* **183**, 265 (1969)

For the evaluation of $n(E)$ with assumptions on w such that Gaussian integrals result, see

G. J. Papadopoulos, *J. Phys. A* **7**, 183 (1974)

A. Maheshwari, *J. Phys. A* **8**, 1019 (1975)

More recent work can be found in

E. P. Gross, *J. Stat. Phys.* **17**, 265 (1977)

Functional integral approaches to impurities that involve many-body theories arise from the Anderson model. In this case the functional integral involves operators in the integrand, a significant complication. See the notes to Section 22 for references.

For manipulations resembling those in our (28.40), see

R. A. Ferrell, Field Theory of Phase Transitions, in *Contemporary Physics*, Vol. 1, International Atomic Energy Agency, Vienna, 1969

The relevance of Bose-Einstein condensation for many phase transitions is suggested by the use of "infrared bounds" by

J. Fröhlich, B. Simon, and T. Spencer, *Phys. Rev. Lett.* **36**, 804 (1976)

Critical Droplets, Alias Instantons, and Metastability

The title of this section could be dramatized a bit. It could have been called “An Essential Singularity in First-Order Phase Transitions,” or “First Steps in Singular Perturbation Theory,” or even (as similar work has been called) “The Fate of the False Vacuum.” In fact we look at the anharmonic oscillator, and conclusions about more grandiose systems are mostly implication by analogy.

The basic problem is that there is a local minimum in an energy function(al), and the system, be it more or less grandiose, tends to remain near the minimum for some time until it finally finds its way to distant but lower energy states. The simplest example occurs in one-particle quantum mechanics where a local minimum in a potential can lead to unstable states. The lifetime can often be gotten by WKB barrier penetration calculations, and from a more systematic theoretical view the state is a resonance or pole in the S matrix. Another example of the basic problem occurs in statistical physics where metastable states (like supercooled steam) are often thought of as local minima in the free energy, and the condensation to the stable ground state is supposed to go via the nucleation of critical droplets. Here both the calculational schemes and the general theoretical framework are far less secure. Finally, problems of this sort have arisen in quantum field theory and it is there that the term “instanton” was coined to describe what corresponds to the critical droplet of statistical physics.

What we do in this section is solve the statistical mechanics problem for a system in one space dimension, which is essentially the same as doing

quantum mechanics. However, instead of the traditional barrier penetration formulas we employ instanton methods. In the references are given both the sources of our development and the places to look for the applications.

Consider a system of spins in one space dimension with periodic boundary conditions, nearest neighbor coupling, and a self interaction. If the spacing of the spins goes to zero and other continuum limits are taken appropriately, the partition function for the system is

$$Z(\alpha) = \int d\varphi(x) \exp \left\{ - \int_0^L \left[\frac{1}{2} \left(\frac{d\varphi}{dx} \right)^2 + \epsilon\varphi^2 + \alpha\varphi^4 \right] dx \right\} \quad (29.1)$$

where L is the circumference of the ring, φ is the spin field, $d\varphi(x)$ is the measure for the functional integral (which when taken together with $\exp(-\frac{1}{2} \int \varphi^2 dx)$ is Wiener measure), ϵ and α are parameters, and the ϵ and L dependence of Z is suppressed. The quantity ϵ is taken positive throughout. The thermodynamic limit is $L \rightarrow \infty$, and the infinite volume free energy per unit volume is

$$\psi(\alpha) \equiv - \lim_{L \rightarrow \infty} L^{-1} \log Z(\alpha) \quad (29.2)$$

For $\alpha > 0$ the potential energy

$$V(\varphi) = \epsilon\varphi^2 + \alpha\varphi^4 \quad (29.3)$$

is positive for all real φ and the functional integral is perfectly well behaved. Of course, should one expand $\exp(-\int \alpha\varphi^4 dx)$ in powers of α and interchange order of summation and functional integration, the resulting series is expected to diverge—that we saw earlier in Section 13. In the present section we examine what happens when the sign of α is changed. Then V continues to have a minimum at $\varphi=0$, but for large values of φ V becomes negative and indeed the object on the right hand side of (29.1) does not exist.

For this reason the best one can hope for is an analytic continuation of $\psi(\alpha)$ from $\text{Re } \alpha > 0$ to $\text{Re } \alpha < 0$. Discovery of what this analytic continuation is turns out to involve working with the functional integral in (29.1) as if it meant something, even for $\text{Re } \alpha < 0$. I do not know how such a procedure can be justified, but we do it anyway.

First, what is the standard method of handling (29.1)? From (29.2) it is clear that we are just looking at an example of the Feynman-Kac formula (with L replacing T), and by comparing (29.1) to the work of Section 7 it follows that ψ is just the ground state energy of the Hamiltonian $\frac{1}{2} p^2 + V$.

Call this $E_0(\alpha)$. For $\alpha > 0$, $E_0(\alpha)$ is perfectly well defined. For $\alpha < 0$ there is no stable ground state, but for $|\alpha|$ small there is an approximately stationary state whose lifetime can be estimated by barrier penetration formulas. Then, again using standard ideas, the decay rate is related to the imaginary part of the energy and thus the analytical continuation of $E_0(\alpha)$ is achieved and an imaginary part assigned to E_0 for negative α .

The method that we use looks at the extrema of the integrand in (29.1). Thus our techniques are essentially those used in Section 13 on the WKB method. Naively one is aiming for a Laplace method—that is, extrema that are true maxima with real integration variables; for the analytic continuation, however, it will turn out that such naivete will be compromised. Write

$$S(\varphi) = \int_0^L dx \left[\frac{1}{2} \left(\frac{d\varphi}{dx} \right)^2 + \epsilon \varphi^2 + \alpha \varphi^4 \right]$$

Variation of S (i.e., $\delta S = 0$) yields

$$\frac{d^2\varphi}{dx^2} = 2\epsilon\varphi + 4\alpha\varphi^3 \equiv -\frac{\partial}{\partial\varphi} U(\varphi) \quad (29.4)$$

The function $U(\varphi)$ is just $-V(\varphi)$, with V given in (29.3). (Contrary to the reader's probable first reaction, it actually makes things clearer to define an object that is just minus an object already defined.)

First consider constant solutions, that is, $d\varphi/dx = 0$. Obviously $\varphi = 0$ is always a solution. For nonconstant solutions (29.4) is essentially a problem in classical mechanics, with x a "time" variable and φ having the role of "position." The periodic boundary conditions (that one uses for a partition function) require that we consider only paths for which $\varphi(0) = \varphi(L)$. Moreover, one can, so it is claimed, restrict attention to paths that start and end near $\varphi = 0$; other paths remain away from zero for finite fractions of the total "time," hence give action contributions that differ by multiples of L from those of the $\varphi(0) \sim 0$ paths.* For $\alpha > 0$ there are no paths that

*A bit of caution is required here, since there are periodic nonstationary paths that start and end at nonzero φ but have action not of order L (they have just the right initial momentum to get to $\varphi = 0$ and stay there for a "time" of order L and then fall back to the initial point). To see why they can be neglected recall that our procedures throughout depend on the Feynman-Kac formula. As pointed out in Section 7 for the evaluation of (29.1) in the limit $L \rightarrow \infty$ it is not essential to use $\text{Tr } \exp(-LH) = \int du G(u, -iL; u)$ (with $u = \varphi(0) = \varphi(L)$) but rather G at some particular spatial arguments can be taken or even the integral over some smaller spatial range can be used. Thus we can impose the additional requirement that $\varphi(0) = \varphi(L) = 0$. This presents no difficulty for the instanton paths discussed below, since they are $O(\exp(-L \cdot \text{const.}))$ at $x = 0$ and $x = L$.

start and end at $\varphi=0$ except for $\varphi(x)\equiv 0$. Hence in the language used earlier in this book there is a single classical path and it remains only to do the Gaussian integrals associated with $\delta^2 S$ evaluated along this path. Those integrals are just what one gets for the harmonic oscillator, since the φ^4 term does not contribute to $\delta^2 S$ for the path $\varphi\equiv 0$.

Now consider $\alpha < 0$. Again we look at (29.4) to find the extrema of S that should give the major contribution to Z . (This step in our analysis must be designated “inspired.” On the one hand, it leads to beautiful results; on the other, its mathematical justification—involving a meaningless object—is wanting.) The constant solution $\varphi(x)\equiv 0$ is still valid. There are additional constant solutions $\varphi = \pm \sqrt{-\epsilon/2\alpha}$,^{*} but their action differs by terms of order L from that of $\varphi\equiv 0$.[†] Now, however, there are solutions satisfying $\varphi(0)=\varphi(L)\sim 0$ which are not constant. From Fig. 29.1, which illustrates $U(\varphi)$ for $\alpha < 0$, the nature of these solutions is clear. At $x=0$, the particle starts very close to $\varphi=0$. Even though it is going to move away from $\varphi=0$ (it is not the trivial solution) it must take very long to do so (since $L\rightarrow\infty$). Hence it starts with little energy and spends almost all the time L getting away from $\varphi=0$ and falling back up the hill as $x\rightarrow L$. At some intermediate time though, it builds up speed and falls down past $\varphi = +\sqrt{-\epsilon/2\alpha}$ (or $\varphi = -\sqrt{-\epsilon/2\alpha}$) and up the other side. It happens that for (29.4) an explicit solution is available:

$$\varphi_z(x) = \pm \left(\frac{\epsilon}{-\alpha} \right)^{1/2} \operatorname{sech} \left[\sqrt{2\epsilon} (x-z) \right] \quad (29.5)$$

where z is arbitrary. The freedom to choose z essentially anywhere in the interval $[0, L]$ arises from the smallness of sech as $|x-z|$ becomes of order L . In principle the function entering the functional integral should be periodic. Making φ_z periodic adds little to S because of φ_z ’s smallness near

*With the $\varphi(0)=\varphi(L)$ boundary condition these paths do not appear but they *nearly* do. That is, there is a path that “hurries” to $\varphi=\sqrt{-\epsilon/2\alpha}$ (or $-\sqrt{-\epsilon/2\alpha}$), remains there almost until $t=L$, and hurries back to zero. It happens that the constant path $\varphi = \pm \sqrt{-\epsilon/2\alpha}$ terms are negligible, but if they were not the “near” solutions would have to be considered.

[†]Terms whose action differs by $O(L)$ drop out because we are looking at $\psi \sim (1/L) \log Z$. Using obvious notation ψ is

$$\begin{aligned} \frac{1}{L} \log Z &= \frac{1}{L} \log (Z_{\varphi=0} + Z_{\varphi=\sqrt{-\epsilon/2\alpha}}) \\ &= \frac{1}{L} \log \{ Z_{\varphi=0} [1 + O(\exp(-\text{const} \cdot L))] \} \\ &= \frac{1}{L} \log Z_{\varphi=0} + \frac{1}{L} \exp(-\text{const} \cdot L) \end{aligned}$$

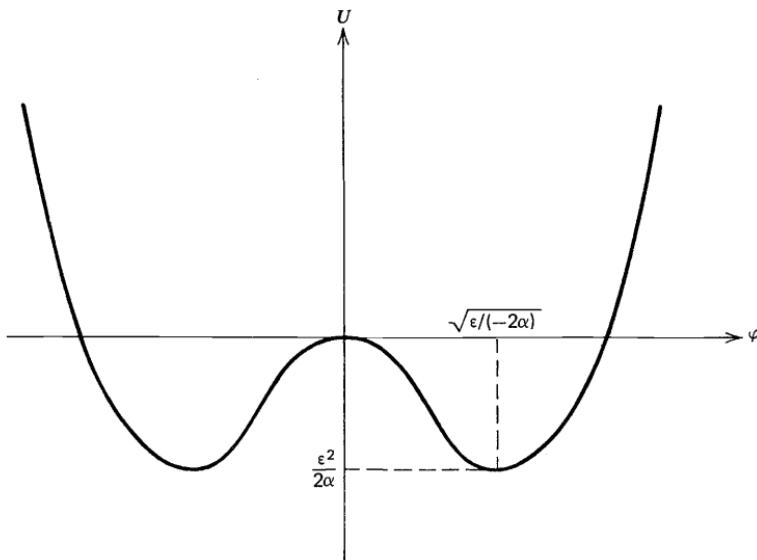


Fig. 29.1 $U(\varphi) = -\varepsilon\varphi^2 - \alpha\varphi^4$ for $\alpha < 0$.

0 and L ; hence the contribution of the modified φ_z is essentially the same as that of (29.5) and the distinction can be ignored. The action associated with the path can be evaluated directly

$$S(\varphi_z) = \frac{(2\varepsilon)^{3/2}}{-3\alpha} \quad (29.6)$$

Before analyzing these solutions more carefully we indicate the general structure of the partition function for $\alpha < 0$. Write

$$Z = Z_0 + Z_1 + Z' \quad (29.7)$$

Z_0 is the $\varphi \equiv 0$ contribution. It does not differ much from its $\alpha > 0$ counterpart and we could easily write its approximate value by doing the Gaussian integrals for $\delta^2 S$ evaluated along $\varphi \equiv 0$. However, since all we really want is the ratio of this quantity to other similar quantities we shall forebear. Z_1 is the contribution from the paths φ_z . Finally, Z' refers to everything else, about which more later.

The next task—and this is where things start getting interesting—is the small oscillation integral for the solution φ_z , that is, the Gaussian

integrals associated with $\delta^2 S$ evaluated along φ_z . Letting $\eta = \varphi - \varphi_z$ we have

$$S(\varphi) = S(\varphi_z) + \int_0^L dx \left[\frac{1}{2} \left(\frac{d\eta}{dx} \right)^2 + \epsilon \eta^2 + 6\alpha \varphi_z^2(x) \eta^2 \right] \quad (29.8)$$

As in our work on the WKB approximation the second variation will be diagonalized by means of an auxiliary differential equation having the form of a Schrödinger equation:

$$-\frac{1}{2} \frac{d^2 \eta_j}{dx^2} - 6(-\alpha) \varphi_z^2(x) \eta_j + \epsilon \eta_j = \lambda_j \eta_j \quad (29.9)$$

Using the normalized eigenfunctions of (29.9) an arbitrary deviation from φ_z can be written

$$\eta(x) = \sum_j a_j \eta_j(x) \quad (29.10)$$

and the corresponding action is

$$S(\varphi) = S(\varphi_z) + \sum_j a_j^2 \lambda_j \quad (29.11)$$

The integration variable for the functional integral started as $d\varphi$. In using Laplace's method we implicitly have gone over to $d\eta$ and with (29.10) we make the further change of variables to integration over the modes. The overall normalization constant for the functional integral (which we call \mathcal{N} and which we shall have no call to evaluate) will be adjusted so that the new integration volume is

$$\prod_j \left(\frac{da_j}{\sqrt{\pi}} \right)$$

In this way integrals of Gaussians involving the action (29.11) will simply yield $1/\sqrt{\lambda}$. By good fortune, when the solution (29.5) is inserted in (29.9) the Schrödinger equation that results

$$-\frac{1}{2} \frac{d^2 \eta_j}{dx^2} - 6\epsilon \operatorname{sech}^2 [\sqrt{2\epsilon} (x-z)] \eta_j + \epsilon \eta_j = \lambda_j \eta_j \quad (29.12)$$

is one for which some information is available. There are just two bound

states with eigenfunctions and eigenvalues:

$$\begin{aligned}\eta_0(x) &= \sqrt{\frac{3}{4}} (2\epsilon)^{1/4} \operatorname{sech}^2[(x-z)\sqrt{2\epsilon}] & \lambda_0 &= -3\epsilon \\ \eta_1(x) &= \sqrt{\frac{3}{2}} (2\epsilon)^{1/4} \frac{\sinh[(x-z)\sqrt{2\epsilon}]}{\cosh^2[(x-z)\sqrt{2\epsilon}]} & \lambda_1 &= 0\end{aligned}\quad (29.13)$$

All the interesting troubles associated with this problem come from these two states. The remaining states have $\lambda_j > 0$, and while it is not trivial to evaluate the product of the λ 's (which must be done for $\det \delta^2 S$), there is no conceptual difficulty to overcome.

From (29.13) it follows that the integral over a_0 diverges

$$\frac{1}{\sqrt{\pi}} \int da_0 e^{3\epsilon a_0^2} \quad (29.14)$$

while the argument of the exponent for the a_1 integration is just zero. There is physics in each divergence. To see this we shall have to interpret some of our results. The extremum φ_z is evidently—from (29.13)—a saddle of S and not a minimum. In the space of periodic functions on $[0, L]$, $\varphi \equiv 0$ is a local but not absolute minimum. Because $\alpha < 0$, paths with very large φ will have much smaller (algebraically) values of S . In passing from the valley of the local minimum to the depths of paths with much smaller S one must first rise and then descend. Among paths in function space that rise and descend there is one (or more) that rises as little as possible—this will be like the road through a mountain pass or col. It too will be an extremum of S and is in fact the path φ_z that we have found. A col has the property that mostly the peaks rise around it (corresponding to $\lambda_j > 0$ for $j \neq 0, 1$). However, there is at least one direction, corresponding to passage from the valley to the global depths, for which the col is a maximum. Because we are in function space, this “direction” corresponds to a mode, to a destabilizing fluctuation about the path φ_z . Finally, let us describe this in statistical mechanics language. The function $\varphi \equiv 0$ is a state of uniform zero magnetization. Because $\alpha < 0$, states with extremely large φ are in fact states of lower free energy (in this simplified model there is no description of the form of the stable state at large φ). Presumably there is some distant global minimum whose nature does not concern us. Perhaps a proper treatment of this minimum could justify the manipulations we have performed on (29.1) with $\alpha < 0$; if the system fluctuates away from $\varphi \equiv 0$ it will fall to the large φ values. The mountain pass represents the cheapest—in the currency of free energy—way of accomplishing this fluctuation.

This is exactly the role of the critical droplet in the theory of metastability. The mode η_0 represents increase or decrease in the size of the droplet. Both increase and decrease are destabilizing, the one sending the system to the large φ global minimum of S , the other sending the system back to $\varphi=0$. The state φ_z is therefore the main barrier the system must pass through in its decay from the metastable state $\varphi \equiv 0$, and we can reasonably look to the integral over a_0 to provide a lifetime for the metastable state. In field theory the variable x is time and the localization of the excitation is a localization in time. Hence φ_z is called an “instanton” solution.

From this picture we can also get an interpretation for η_1 . Mathematically, S is translationally invariant. Thus there is a solution $\varphi_z(x)$ for any z in the interval 0 to L . This problem has of course been with us since the arbitrary parameter z first appeared in φ_z , and we noted that rather than finding a single solution we had a one-dimensional family of them. The zero eigenvalue of (29.12) is just another manifestation of the translation invariance, and the wave function η_1 is just (proportional to) $d\varphi_z/dz \sim (\varphi_{z+\Delta} - \varphi_z)/\Delta$. In statistical mechanics this corresponds to the fact that the critical droplet can appear anywhere throughout the volume.

It follows that we have been double counting. An integral over z , to add up all the instanton extrema, would accomplish the same result as integrating over a_1 . In any case, it is clear that any way of taking the translational invariance into account ought to introduce a factor L (“the volume”). We shall do the integral over z , and not over a_1 . However, because the function η_1 was normalized to unity, the functional integral had an integral da_1 , not dz . The relation between them is obtained by considering a path of the form $\eta = \varphi_{z+dz} - \varphi_z$. This path is entirely in the “direction” η_1 and therefore

$$\varphi_{z+dz} - \varphi_z = \frac{d\varphi_z}{dz} dz = \eta_1 da_1$$

The Jacobian of the transformation from a_1 to z is thus the ratio of the norms of the functions. Therefore the integration over a_1 becomes

$$\begin{aligned} \frac{1}{\sqrt{\pi}} \int da_1 &= \frac{1}{\sqrt{\pi}} \int dz \left\| \frac{d\varphi_z}{dz} \right\| = \frac{1}{\sqrt{\pi}} \left[\int \left(\frac{d\varphi_z}{dz} \right)^2 dx \right]^{1/2} \int dz \\ &= \frac{(2\epsilon)^{3/4}}{(-3\alpha)^{1/2}} \frac{L}{\sqrt{\pi}} \end{aligned} \quad (29.15)$$

In our final result the term in (29.15) must be multiplied by 2 to allow for the instanton's heading either left or right.

The integrations over da_j , $j > 1$ give $(\Pi'_j \lambda_j)^{-1/2}$ where the prime indicates that $j=0$ and 1 are excluded from the product, and where overall normalization factors have been neglected. The only thing we shall want to know is the ratio of this product [which appears in Z_1 of (29.7)] to the corresponding product for Z_0 . That calculation is relegated to an appendix.

With everything else out of the way we have no choice but to confront the integral (29.14) with its manifest divergence. Well—one says at this point—we weren't *really* interested in (29.1) but only in its analytic continuation. And now it seems that all the bad behavior of that integral has concentrated itself in a single one-dimensional integral

$$J = \frac{1}{\sqrt{\pi}} \int da_0 e^{3\epsilon a_0^2} \quad (29.16)$$

But this is something that we know how to analytically continue.

Let us review what this continuation is. Let

$$F(\mu) \equiv \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\mu z^2} dz \quad (29.17)$$

For $\operatorname{Re} \mu > 0$, $F(\mu)$ is equal to the function $G(\mu)$ where

$$G(\mu) \equiv \frac{1}{\sqrt{\mu}} \quad (29.18)$$

Analytic continuation of F to $\operatorname{Re} \mu < 0$ is simple: just let $F = G$ everywhere. (Of course there is then a branch point in both F and G at $\mu = 0$.) However ... for $\operatorname{Re} \mu < 0$ the integral on the right hand side of (29.17) is no longer how F is defined. There is also a way of doing the analytic continuation in which the continued function is still an integral. In (29.17) consider the change of contour from the real z axis to $z = e^{-i\theta}x$, $-\infty < x < \infty$. Originally (29.17) was defined for $-\pi/2 < \arg \mu < \pi/2$. Now it is defined for $-\pi/2 < \arg \mu - 2\theta < \pi/2$. Hence for $\theta = \pi/2$ F has been continued to the negative half plane. (Note that contributions from contours at infinity are zero.) The continued function, written as an integral, looks like

$$F(\mu) = \frac{1}{\sqrt{\pi}} \int_{-i\infty}^{i\infty} e^{-\mu z^2} dz = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{\mu x^2} i dx = i \sqrt{\frac{1}{-\mu}} \quad (29.19)$$

It is in the sense of this continuation that the integral J of (29.16) is to be interpreted as

$$J = \frac{1}{\sqrt{\pi}} \int_0^{i\infty} da_0 e^{3\epsilon a_0^2} \quad (29.20)$$

(or $\int_0^{-i\infty}$ giving the other continuation of the square root). The integral along the real a_0 axis has become a vertical line from 0 to + (or $-$) $i\infty$. Notice that the integral—presumably just an analytic continuation—does not run from $-i\infty$ to $+i\infty$ as might have been suggested by (29.19). This change, which introduces a factor $\frac{1}{2}$ in the final result, is worth commenting on. To use statistical mechanics language, φ_z is a critical droplet and the mode a_0 is the unstable direction in the function space of microscopic configurations. Increase of a_0 (from 0) is growth of the critical droplet and decrease in free energy. Decrease of a_0 (to negative values) represents return of the system to the metastable minimum $\varphi \equiv 0$. Consequently the range of integration $(-\infty, 0]$ gives contributions of the metastable ground state and the infinity obtained from the integration $-\infty < a_0 < 0$ is fake: the integral should not really be Gaussian and for large negative values of a_0 one is simply picking up metastable ground state contributions. They should properly be included in the integral around $\varphi \equiv 0$ and—what is more relevant to our present concern—they only contribute to the real part of Z , perhaps changing it slightly. The principal embarrassment is the range $[0, \infty)$ for which the free energies get much deeper (in this model infinitely deeper, and in a model with the stable state included, deeper on the order of L). Hence it is the integral from 0 to $+\infty$ that needs shifting, and this can be rotated to the positive or negative imaginary a_0 axis, with the result written above. See Fig. 29.2. In some sense the passage to the imaginary axis represents outgoing wave boundary conditions, but I do not know of any way to make the analogy precise.

By (29.20)

$$J = \pm \frac{i}{2} \sqrt{\frac{1}{3\epsilon}} \quad (29.21)$$

and the free energy has acquired an imaginary part.

Because the negative energy mode is of such significance in this problem, we remark that we could have deduced its existence even without the solution $\eta_0(x)$ given above, a fact worth knowing if one works with some other potential and is not fortunate or clever enough to produce explicit solutions. The reasoning goes as follows. Translational invariance guarantees us a zero energy mode. Because the instanton starts and ends at

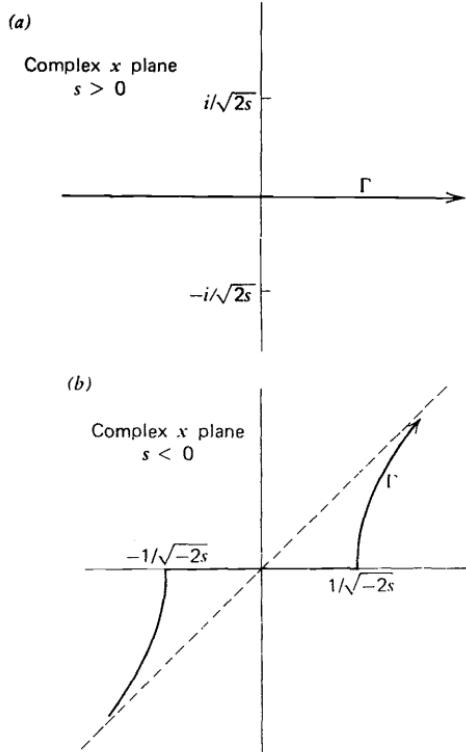


Fig. 29.2 The analytic continuation leading to (29.20) with its lower integration limit taken to be 0 can be understood in terms of the integral $f(s) = \int_{\Gamma} dx \exp(-x^2 - sx^4)$. This integral also sheds light in a general way on the analytic continuation of the functional integral (29.1). (a) For $\text{Re } s > 0$, Γ is taken to be the real line. By a change of variable $f(s) = (1/\sqrt{s}) \int_{-\infty}^{\infty} \exp[-(1/s)(y^2 + y^4)] dy$, so that for small s the main contribution comes from $y=0$ and the other extrema (at $y = \pm i/\sqrt{2}$) of the integrand play no role. (b) For $\text{Re } s < 0$, the real line is an unacceptable contour but as discussed in the text $f(s)$ can be analytically continued by rotating the contour so that its asymptotic direction satisfies $\text{Re}(sx^4) > 0$ [see also the discussion following (13.20)]. For $s < 0$ taking Γ to go asymptotically to 45 and 225° satisfies this. For $|s| \rightarrow 0$ the evaluation of the integral is best accomplished by passing through the critical points, which are 0 and $x = \pm(-2s)^{-1/2}$. This leads to the contour shown in the figure, where the contour's departure from the real axis is effected by a 90° turn at the critical points in keeping with the method of steepest descents. For small $|s|$, $\text{Re } f(s)$ is not much changed by the passage to negative s since the nonzero extrema are far from zero. Had we expanded the integrand about $x = (-2s)^{-1/2}$ (the analogue of the instanton) it would have looked as if the real part coming from $x < (-2s)^{-1/2}$ has a growing Gaussian, but we now see that this is just the tail end of the integrand that was decreasing from $x=0$. This is why in the text we call the infinity from the range $-\infty < a_0 < 0$ a "fake." The imaginary part of f comes from the part of Γ that heads into complex x , and when the integral is written in terms of the variable $a = x - (-2s)^{-1/2}$ the main contribution (from the right side of the contour) comes from the quadratic term in a with the integral starting from $a=0$ and going to $\text{Im } a > 0$. The critical point at $x = -(-2s)^{-1/2}$ is treated in the same way.

zero (warning: in some applications the instantons are *not* periodic) and is smooth, its derivative vanishes somewhere. Hence the zero energy mode, which is the derivative of the instanton, must vanish somewhere. But then, thinking of it as a solution of a certain Schrödinger-like equation, it is a solution with a node and therefore not the ground state. Consequently there is at least one state of lower, hence negative, energy which has no node.

To summarize what has been obtained so far, we have

$$\begin{aligned} Z &= Z_0 + Z_1 + Z' \\ Z_0 &= \int_{(0,0)}^{(0,L)} d\eta \exp \left\{ - \int_0^L \left[\frac{1}{2} \left(\frac{d\eta}{dx} \right)^2 + \epsilon \eta^2 \right] dx \right\} = \mathcal{N} \prod_j \lambda_j^{(0)-1/2} \\ Z_1 &= \mathcal{N} \left(\pm \frac{i}{2} \sqrt{\frac{1}{3\epsilon}} \right) \left[\frac{2(2\epsilon)^{3/4} L}{(-3\alpha)^{1/2} \sqrt{\pi}} \right] \exp \left(- \frac{(2\epsilon)^{3/2}}{(-3\alpha)} \right) \prod'_j \lambda_j^{-1/2} \end{aligned} \quad (29.22)$$

The normalization constant \mathcal{N} appearing above is the same in both equations. The quantities $\lambda_j^{(0)}$ correspond to the modes for the $\varphi \equiv 0$ state. The prime on the product in Z_1 indicates that $j=0$ and 1 are excluded.

Finally, what has been left out? This is the term Z' , and it is mainly comprised of multi-instanton terms (as far as anyone knows). Besides the solutions φ_z already considered there are those that make two or more trips away from $\varphi=0$. Because the instanton is so highly localized, corrections to the action [and deviations from being exact solutions of (29.4)] are negligible and the action for multi-instanton excitation is additive. The blowup mode around each instanton gives a separate contribution, but for the zero mode (due to translational invariance) a bit of care must be exercised. One sums over possible locations of (say) the centers of the instantons, and the instantons behave therefore like identical particles. Thus rather than a simple L^n for the n instanton mode one obtains $L^n/n!$.

The result is that the contribution to the partition function due to Z_1 and Z' looks exactly like the contribution of a dilute gas of excitations. "Dilute" because there is no mutual interaction, this corresponding to our taking the action to be additive. Consequently if we write

$$Z = Z_0 \left(1 + \frac{Z_1}{Z_0} + \frac{Z'}{Z_0} \right)$$

we observe that Z'/Z_0 consists of terms of the form $(Z_1/Z_0)^k/k!$, $k=2,3,\dots$. (Of course, adding action contributions means multiplying the

partition function, hence the term $(Z_1/Z_0)^k/k!$ for k instantons.) But this is just the exponential series, and we have finally

$$Z = Z_0 \exp\left(\frac{Z_1}{Z_0}\right) \quad (29.23)$$

Our main interest throughout has been the imaginary part of ψ which, because Z_0 is real, has the form

$$\begin{aligned} \text{Im } \psi(\alpha) &= -\frac{1}{L} \lim \text{Im} \left(\frac{Z_1}{Z_0} \right) \\ &= \pm \frac{1}{2} \sqrt{\frac{1}{3\epsilon\pi}} \left[\frac{2(2\epsilon)^{3/4}}{(-3\alpha)^{1/2}} \right] \exp \left[\frac{-(2\epsilon)^{3/2}}{(-3\alpha)} \right] \frac{\Pi' \lambda_j^{-1/2}}{\Pi \lambda_j^{(0)-1/2}} \end{aligned} \quad (29.24)$$

Bringing our expression for the ratio of the product of continuum states from the appendix we have

$$\text{Im } \psi(\alpha) = \pm \frac{2^{7/4} \epsilon^{5/4}}{(-\pi\alpha)^{1/2}} \exp \left[-\frac{(2\epsilon)^{3/2}}{(-3\alpha)} \right] \quad (29.25)$$

Thus ψ has a cut along the negative real α axis with a highly singular discontinuity along the cut. This is generally referred to as an essential singularity associated with first-order phase transitions, although it is not clear that the technical definition of essential singularity (i.e., a meromorphic function with a Laurant expansion containing arbitrarily large negative powers of the variable) is satisfied.

Is there an essential singularity, or indeed any singularity, at first-order phase transitions? Nobody knows for sure. In systems with long range interactions (e.g., Van der Waals, Kac-Hemmer-Uhlenbeck with appropriate limits) the free energy is analytic across the transition. This is because the critical droplet, or other condensing excitation, has infinite energy. For systems with short range forces the calculation given above may persuade one that there is a singularity, but one must be critical of the assumption of noninteracting instantons. This is because as $\alpha \rightarrow 0$ the singularity is due to larger and larger droplets, so that it is not obvious that they can be imagined all to be sufficiently far from each other for the action to be simply additive. Fisher once invented a model with rather strange interactions but with the merit that separate droplets do not interact at all. This shows an essential singularity as the system condenses.

However, according to results of Sewell in going from the condensed to the "gaseous" state the free energy remains analytic. There are also results from percolation theory indicating that there is an essential singularity in that phase transition.

The calculation of this section could be—and has been—performed for many other potentials, but the essential results and techniques are the same. We chose this example because of its large number of explicitly solved components, thanks to Langer's 1967 tour de force.

APPENDIX: SMALL OSCILLATIONS ABOUT THE INSTANTON

By (29.24) we require the ratio

$$A = \frac{\prod_j' \lambda_j^{-1/2}}{\prod_j \lambda_j^{(0)-1/2}} \quad (29.26)$$

where the notation is defined following (29.22). The unprimed product is (up to a normalization constant) the determinant of the infinite quadratic form $\delta^2 S$ for oscillations about $\varphi \equiv 0$. If the primed product were multiplied by the two missing (by definition of the prime) eigenvalues it too would be a determinant. But we know from Section 6 that these determinants satisfy the Jacobi equations about the respective extrema; hence A of (29.26) is essentially the ratio of solutions of the Jacobi equations. So it seems that all we need to know is solutions of the Jacobi equation. The only cloud on the horizon is that one of the eigenvalues of $\delta^2 S$ for the instanton mode is reputed to be zero, casting some doubt on its use as a divisor. Fortunately, though, its reputation for nullity misses the truth by a term on the order of $\exp(-L)$, and this is enough to make division well defined.

To be precise, let $f^{(0)}$ and f_z be the solutions of

$$\frac{1}{2} \frac{d^2 f^{(0)}(x)}{dx^2} - \epsilon f^{(0)}(x) = 0 \quad (29.27)$$

$$\frac{1}{2} \frac{d^2 f_z(x)}{dx^2} - \epsilon f_z(x) + 6\epsilon \operatorname{sech}^2 [\sqrt{2\epsilon} (x-z)] f_z(x) = 0 \quad (29.28)$$

with boundary conditions

$$\begin{aligned} f^{(0)}(0) &= f_z(0) = 0 \\ \frac{df^{(0)}(0)}{dx} &= \frac{df_z(0)}{dx} = 1 \end{aligned} \quad (29.29)$$

The eigenvalues λ_j (dropping the explicit z dependence) are defined by (29.9) (with of course the boundary condition $\eta_j(0)=\eta_j(L)=0$). The quantities $\lambda_j^{(0)}$ are defined as eigenvalues for an equation just like (29.9) but with no $\alpha\varphi_z^2$ term.

The statement made above concerning A can now be precisely formulated as

$$A^2 = \frac{f_0(L)}{f_z(L)/\lambda_0\lambda_1} \quad (29.30)$$

Before going further, it seems reasonable that A not depend on z , so we fix z to be $L/2$ and henceforth drop z as a subscript.

Now about the nullity of λ_1 ... This was based on the explicit solution (29.13), but one should notice that $\eta_1(x)$ misses being zero at the end points by $O(e^{-L/2})$. Hence we expect that λ_1 will be just a bit over zero because the actual eigenfunction is forced to be zero at the edges. Before evaluating this though it is best to evaluate $f(L)$, since some of the techniques for this enter in obtaining λ_1 .

To dispose of the easy part, we observe that

$$f^{(0)}(L) = \frac{1}{\sqrt{2\epsilon}} \sinh L\sqrt{2\epsilon} \sim \frac{1}{2\sqrt{2\epsilon}} e^{L\sqrt{2\epsilon}} \quad (29.31)$$

Now $\eta_1(x)$ [this is the η_1 of (29.13) with $z=L/2$] satisfies (29.28) *exactly*, but it does not quite satisfy the boundary conditions (29.29). Of course the second-order differential equation (29.28) has a second linearly independent solution, call it ξ_1 , and some linear combination can be found to satisfy (29.29). Let

$$f(x) = u\eta_1(x) + v\xi_1(x) \quad (29.32)$$

(f is the same as $f_{L/2}$). The Wronskian of η_1 and ξ_1 is a constant which we take to be 1:

$$\eta_1 \frac{\partial \xi_1}{\partial x} - \xi_1 \frac{\partial \eta_1}{\partial x} = 1 \quad (29.33)$$

The conditions (29.29) at $x=0$ require

$$0 = u\eta_1(0) + v\xi_1(0) \quad (29.34)$$

$$1 = u \frac{d\eta_1(0)}{dx} + v \frac{d\xi_1(0)}{dx} \quad (29.35)$$

leading to

$$u = -\xi_1(0) \quad v = \eta_1(0) \quad (29.36)$$

However, the requirement of constancy in (29.33) gives us all the information we need for ξ_1 . From its explicit form, for $|x - L/2|$ large,

$$\begin{aligned} \eta_1(x) &\sim N \exp\left(-\kappa\left|x - \frac{L}{2}\right|\right) \operatorname{sign}\left(x - \frac{L}{2}\right) \\ N &= 2\sqrt{\frac{3}{2}} (2\epsilon)^{1/4}, \quad \kappa = \sqrt{2\epsilon} \end{aligned} \quad (29.37)$$

From (29.28), away from $L/2$ the linearly independent solutions of (29.28) are

$$\exp\left(\pm\kappa\left(x - \frac{L}{2}\right)\right)$$

so that ξ_1 must have the asymptotic behavior

$$\xi_1(x) \sim N' \exp\left(+\kappa\left|x - \frac{L}{2}\right|\right) \quad (29.38)$$

It is (29.33) that sets the sign of ξ_1 for large $|x - L/2|$. It also gives us the constant N' ($N' = 1/2\kappa N$) and with it u and v . From u and v $f(L)$ immediately follows and is

$$f(L) = \frac{1}{\kappa} \quad (29.39)$$

Only the eigenvalues λ_0 and λ_1 remain to be determined. But λ_0 is just -3ϵ which is accurate enough for present purposes. (In fact the absolute value of λ_0 cancels entirely from the calculation, since inserting it in (29.30) simply cancels its appearance in (29.22), an appearance earned by virtue of (29.21). All we really needed to know about λ_0 was that it is negative.) The exact λ_1 is defined as the eigenvalue of

$$Hf + \lambda f \equiv \frac{1}{2} \frac{d^2 f}{dx^2} - Q(x)f + \lambda f = 0 \quad (29.40)$$

with $Q = \epsilon - 6\epsilon \operatorname{sech}^2[\kappa(x - L/2)]$ and with $f(0)$ and $f(L)$ strictly zero. Equation 29.40 can be rewritten as an integral equation in the same way that the Lippmann-Schwinger equation is obtained from the Schrödinger

equation in quantum mechanics. Let G be the Green's function for the "unperturbed" operator H , that is,

$$H_x G(x, y) = \delta(x - y)$$

A solution for G satisfying the boundary condition $G(0, y) = 0$ is

$$G(x, y) = 2\theta(x - y)[\eta_1(y)\xi_1(x) - \eta_1(x)\xi_1(y)] \quad (29.41)$$

Using G , (29.40) becomes (with the term λf as the "perturbation")

$$\begin{aligned} f(x) &= \psi(x) - \lambda \int_0^L G(x, y)f(y) dy \\ &= \psi(x) - \lambda \int_0^x G(x, y)f(y) dy \end{aligned} \quad (29.42)$$

where ψ is any function satisfying $H\psi = 0$. What we seek is a value of λ such that there is an f satisfying (29.42) for which both $f(0) = 0$ and $f(L) = 0$. This requires $\psi(0) = 0$; moreover since we are only interested in λ , the overall normalization of ψ is irrelevant and the vanishing of $\psi(0)$ is all we need to specify. Taking therefore

$$\psi(x) = \eta_1(x) + c\xi_1(x) \quad (29.43)$$

it follows that

$$c = -\frac{\eta_1(0)}{\xi_1(0)} = -\left(\frac{N}{N'}\right)e^{-\kappa L} \quad (29.44)$$

By (29.42), to lowest order in λ f is the same as ψ . Hence in the term Gf we replace f by ψ ; this is the analogue of the "first Born approximation." Setting $x = L$ in (29.42) and applying the requirement $f(L) = 0$ gives us the following equation:

$$0 = \eta_1(L) + c\xi_1(L) - \lambda \int_0^L 2[\eta_1(y)\xi_1(L) - \eta_1(L)\xi_1(y)][\eta_1(y) + c\xi_1(y)] dy \quad (29.45)$$

Equation 29.45 is a condition that λ must satisfy. After some algebra and the use of (29.37) and (29.38), (29.45) becomes

$$\lambda = \frac{\eta_1(L)}{\xi_1(L) \left[\int_0^L \eta_1^2(y) dy - c(\eta_1(L)/\xi_1(L)) \int_0^L \xi_1^2(y) dy \right]} \quad (29.46)$$

By definition of η_1 , the first integral in the denominator is unity. By (29.38) the integral $\int \xi_1^2$ is of order $\exp(\kappa L)$. However, the expression $c\eta_1(L)/\xi_1(L)$ is of order $\exp(-2\kappa L)$, so that the second integral in the denominator can be dropped. Using the known values of N and N' , (29.46) becomes

$$\lambda = 24\epsilon e^{-\kappa L} \quad (= \lambda_1) \quad (29.47)$$

This finally gives the value of A as

$$A^2 = \frac{(1/2\kappa) e^{\kappa L} 24\epsilon e^{-\kappa L} (-3\epsilon)}{-1/\kappa} = 36\epsilon^2 \quad (29.48)$$

NOTES

The basic reference for this section is

J. S. Langer, *Ann. Phys.* **41**, 108 (1967)

See also

H. A. Kramers, *Physica* **7**, 284 (1940)

J. Lothe and J. P. Hirth, *Phys. Rev.* **115**, 543 (1959)

R. Landauer and J. A. Swanson, *Phys. Rev.* **121**, 1668 (1961)

There has been quite a bit of work on the problem in recent years. Not only am I unable to be comprehensive in listing references, but the following can only serve as a point at which to begin detective work in searching for such references.

S. Coleman, The Uses of Instantons, in *The Whys of Subnuclear Physics*, Erice 1977, A. Zichichi, Ed., Plenum, New York, 1979.

M. Stone, *Phys. Rev. D* **14**, 3568 (1976)

S. Coleman, *Phys. Rev. D* **15**, 2929 (1977)

C. Callan and S. Coleman, *Phys. Rev. D* **16**, 1762 (1977)

M. Stone, *Phys. Lett.* **67B**, 186 (1977)

The analytic continuation and associated singularities developed in this section can be used to deduce the behavior of the high order terms in the perturbation expansion [in powers of "α" of (29.1), the coefficient of φ^4]. This is work of Bender, Wu, Banks Brezin, Zinn-Justin, Lipatov, and others, and an exposition as well as references can be found in

J. C. Collins and D. E. Soper, *Ann. Phys.* **112**, 209 (1978)

A critique of the use of instantons in quantum chromodynamics has been given by

A. Patrascioiu, "Farewell to Instantons," lecture, 1979

A. Patrascioiu, *Phys. Rev. D* **19**, 3800 (1979)

A physical picture of the critical droplet can be gained from

K. Huang, *Statistical Mechanics*, Wiley, New York, 1963

as well as many other texts on statistical mechanics. The droplet model of phase transitions is described in

M. E. Fisher, *Physics* **3**, 255 (1967)

The work of Sewell dealing with Fisher's model is

G. L. Sewell, *Ann. Phys.* **97**, 55 (1976)

Metastability in systems with a sort of long range force is shown by

O. Penrose and J. L. Lebowitz, *J. Stat. Phys.* **3**, 211 (1971)

Essential singularities in percolation have been found by

H. Kunz and B. Souillard, *Phys. Rev. Lett.* **40**, 133 (1978); *J. Stat. Phys.* **19**, 77 (1978)

The asymptotic behavior of the coefficients of the low temperature expansion for the Ising model has been studied and shows strong evidence for the appearance of *some* singularity. Moreover, that singularity has been found to be consistent with the form predicted by the droplet model.
References are

G. A. Baker, Jr., and D. Kim, Ising Model Phase Boundary, *J. Phys. A* **13**, L 103 (1980)

M. J. Lowe and D. J. Wallace, Instantons and the Ising Model Below T_c , *J. Phys. A* **13**, L 381 (1980)

For another view of metastability, or at least another way of trying for analytic continuation, see

C. M. Newman and L. S. Schulman, *J. Math. Phys.* **18**, 23 (1977)

R. J. McCraw and L. S. Schulman, *J. Stat. Phys.* **18**, 293 (1978)

A rigorous study of the analytic properties of anharmonic oscillator energy eigenvalues is

B. Simon, *Ann. Phys.* **58**, 76 (1970)

Renormalization and Scaling for Critical Phenomena

At second order phase transitions (critical points) there occur long range cooperative phenomena, and the important features of the transition are determined not by short range interparticle forces but by the way in which collective behavior sets in among the particles. Prominent among the "important features" are the critical exponents that characterize the behavior of various quantities as the critical point is approached. For example, as T (temperature) approaches T_c (the critical temperature) from below, the order parameter (say ϕ) vanishes as some power: $\phi \sim |T - T_c|^\beta$. The power is the critical exponent beta. One of the goals of a theory of second order phase transitions is to calculate the critical exponents and, it is hoped, to show them to have "universal" values, namely to depend only on, say, the dimension of the space and the number of components of the order parameter—but not on the details of the interaction.

The way of realizing these hopes has been to "scale," to smear out details of the system and look at it on larger scales. Here then is how Wilson brought in the functional integral: he found that a good way to eliminate short range behavior from a problem was to write the partition function as a functional integral, to transform away from spatial variables so as to be summing over an infinity of modes of the system and then to integrate over the short wavelength modes leaving only the long wavelengths which one believes are what determine the critical behavior. Next one rescales the spatial variables so that the wavelengths of the remaining modes span the same range as did the previous modes. This elimination of the shortest wavelength modes and subsequent rescaling is applied to the

system repeatedly. What then characterizes a system that is exactly at a critical point is that after such a scale transformation the infinite system looks exactly as it did before. Another way to see this is to consider the correlation length ξ defined by

$$g(r) \sim \exp\left(-\frac{r}{\xi}\right) \cdot (\text{power of } r) \quad (30.1)$$

with g being the two particle correlation function at distance r . If all distances are rescaled by a factor 2 then we expect that $\xi \rightarrow \xi/2$. At a critical point ξ is infinite, and the rescaled system is also at a critical point. One can also see from this what happens if the original system was near but not exactly at the transition. Then ξ is large and $\xi/2$ not quite so large. Hence rescaling moves the system away from the critical point. It then turns out that the rate at which the system moves away from the critical point is closely related to the critical exponents.

This rather abstract exposition may be clearer when we have considered a specific example below. If at this stage it all seems very obscure, perhaps it is a good thing in that it reflects the history of the subject. In the late 1960's people already knew that scaling was the way to understand critical phenomena, and in the famous paper of Kadanoff et al. (references below) there is a highly suggestive discussion of what a scale transformation ought to do and how to get critical exponents from it. Nevertheless, it was not until Wilson actually gave an example through the use of functional integration that things really got moving and actual calculations could be performed. Later Niemeijer and Van Leeuwen introduced coordinate space scale transformations more closely resembling those originally contemplated, but that is not our concern here.

The system that we study is the spherical model. This is not quite so direct an investigation of nature as looking at water or at the Curie point in iron; however, since in any event "true" reality is not subject to physical inquiry we reconcile ourselves to looking at a model that is admittedly distant from what one ordinarily considers to be reality. This system has been studied by Ma (see references below) and enough details have been worked out for it to serve as a good pedagogical example. We consider an n component real valued field ϕ defined on a d -dimensional space. A specific configuration of the field has energy

$$H = \int d^d x \left\{ a \sum_{i=1}^n \sum_{j=1}^d \partial_j \phi_i(x) \partial_j \phi_i(x) + U \left[\left(\sum_{i=1}^n \phi_i(x) \right)^2 \right] \right\} \quad (30.2)$$

where $\phi_i(x)$ is the i th component of the field at the point x in the

d -dimensional space. U is a power series

$$U(\rho) = t_1 \rho + \frac{1}{2} t_2 \rho^2 + \frac{1}{6} t_3 \rho^3 + \dots \quad (30.3)$$

We use the following notation:

$$\phi^2 = \sum_{i=1}^n \phi_i(x) \phi_i(x), \quad (\nabla \phi)^2 = \sum_{i=1}^n \sum_{j=1}^d \partial_j \phi_i(x) \partial_j \phi_i(x) \quad (30.4)$$

so that H is more simply written as

$$H = \int d^d x [a(\nabla \phi)^2 + U(\phi^2)] \quad (30.5)$$

The partition function is then the functional integral

$$Z = \int \prod_{i=1}^n d\phi_i(\cdot) \exp(-H) \quad (30.6)$$

Note that the factor “ β ” (inverse temperature) ordinarily appearing in the expression for the partition function has been absorbed into H . Similarly all other thermodynamic parameters are expressed through the coefficients a, t_1, t_2, \dots . A problem inherent in (30.6) is that the functional integral is over a field depending on d variables, with d allowed to be greater than 1. This is a serious complication in that the nature of the functions contributing to the sum is different from what we got for the paths using Wiener measure. However, we ignore this point here; it is taken up in Section 32.1.

Our goal is to study Z in the limit of large n , a limit in which a good deal is known about the model.

We shall never really do the functional integral but rather only integrate out its shortest wavelength components and rescale. To this end we introduce the Fourier components ϕ_{ik} through

$$\phi_i(x) = L^{-d/2} \sum_{k < \Lambda} \exp(ik \cdot x) \phi_{ik} \quad (30.7)$$

where L^d is the total volume of the system and some short distance cutoff is assumed (e.g., this may be a lattice system) so that momenta larger than Λ are not needed. For the physical system Λ would be the size of the Brillouin zone but after the first iteration this meaning is lost.

The Hamiltonian H is rewritten in terms of the Fourier components of ϕ , and the functional integral becomes a multiple integral over the ϕ_{ik} . Let

s be some real number greater than 1; the integrals over ϕ_{ik} are broken into those over ϕ_{ik} with $k > \Lambda/s$ and those with $k < \Lambda/s$. We shall in fact perform only those with $k > \Lambda/s$. Thus

$$Z = \int \prod_i \sum_{\text{all } k} d\phi_{ik} \exp(-H) = \int \prod_i \sum_{k < \Lambda/s} d\phi_{ik} \exp(-H'') \quad (30.8)$$

where H'' is defined by

$$\exp(-H'') \equiv \int \prod_{i, k > \Lambda/s} d\phi_{ik} \exp(-H)$$

We are calling this H'' because we still have several steps to go before we reach our bona fide “new Hamiltonian,” fully rescaled (which will be called H'). To evaluate H'' , first consider the kinetic energy part of H

$$\begin{aligned} \int d^d x (\nabla \phi)^2 &= \int d^d x \sum_{i, j} L^{-d} \sum_{k, k'} (-) e^{i(k+k')x} (k_j) \cdot (k'_j) \phi_{ik} \phi_{ik'} \\ &= \sum_k \sum_i k^2 |\phi_{ik}|^2 \equiv \sum_k k^2 N_k \end{aligned} \quad (30.9)$$

(with $N_k \equiv \sum_i |\phi_{ik}|^2$) where we have used the fact that $\phi_i(x)$ is real ($\Rightarrow \phi_{ik} = \phi_{i-k}^*$). The form (30.9) will serve our purposes well, since the modes with $k < \Lambda/s$ and $k > \Lambda/s$ separate. With U we shall not be so fortunate. Its argument is ϕ^2 . For convenience we break ϕ into two pieces based on its Fourier representation:

$$\begin{aligned} \phi_i^{(<)}(x) &= \sum_{k < \Lambda/s} \phi_{ik} \exp(ikx) \\ \phi_i^{(>)}(x) &= \sum_{k > \Lambda/s} \phi_{ik} \exp(ikx) \end{aligned} \quad (30.10)$$

Then

$$\phi(x)^2 = \sum_{i=1}^n [\phi_i^{(>)2} + \phi_i^{(<)2} + 2\phi_i^{(>)}\phi_i^{(<)}] \quad (30.11)$$

In the spherical model the limit $n \rightarrow \infty$ plays an important role, and in (30.11) we make our first use of this limit. The terms $\phi_i^{(<)2}$ and $\phi_i^{(>)2}$ are positive for each n [note from (30.10) that $\phi_i^{(>)}$ and $\phi_i^{(<)}$ are real] and in the sum over i they add. The cross terms, however, will in general vary in sign

and may be expected to cancel relative to the positive contributions. Next consider $\Sigma_i \phi_i^{(>)2}$:

$$\sum_j \phi_j^{(>)2} = L^{-d} \sum_j \sum_{k > \frac{\Lambda}{s}} \sum_{k' > \frac{\Lambda}{s}} \phi_{jk} \phi_{jk'}^* \exp(i(k-k')x) \quad (30.12)$$

For any particular j the contributions from various k and k' , but $k \neq k'$, need not have any particular phase and so the sum over j will again lead to cancellation of these terms. What does survive is the $k=k'$ terms for which there is coherent addition. Hence

$$\phi^2(x) = \sum_i \phi_i(x)^{(<)2} + L^{-d} \sum_{k > \Lambda/s} N_k \quad (30.13)$$

Notice that the integral over $\phi_{ik}, k > \Lambda/s$, now depends only on N_k . Consequently we change variables of integration. For any particular k

$$\begin{aligned} \prod_{i=1}^n d\phi_{ik} d\phi_{ik}^* &= \prod_i d\text{Re } \phi_{ik} d\text{Im } \phi_{ik} \\ &= (\text{const.}) N_k^{n-1} dN_k \end{aligned} \quad (30.14)$$

The constant is related to the surface area of an n -sphere, but it is of no interest in this calculation. In our functional integral we are not actually integrating over ϕ_{ik} and ϕ_{i-k} or $\text{Re } \phi_{ik}$ and $\text{Im } \phi_{ik}$, but rather our basic integral is over the real $\phi_i(x)$. Thus if we were to replace $\prod_{i,k} d\phi_{ik}$ [of (30.8) and which has not been precisely defined] by a product over k of the expressions in (30.14) we would be overcounting degrees of freedom for integration by a factor of 2. However, instead of carefully working out the correct integration variables we simply cut the dimension of the integration in (30.14) and replace n in the last term by $n/2$ (we also drop the “ -1 ” in the exponent, since $n \gg 1$). The integral to be evaluated is therefore

$$\begin{aligned} \exp(-H'') &= \int \prod_{k > \Lambda/s} N_k^{n/2} dN_k \exp \left(-a \sum_k k^2 N_k - \int d^d x U \left(\sum_i \phi_i^{(<)2}(x) \right. \right. \\ &\quad \left. \left. + L^{-d} \sum_{k > \Lambda/s} N_k \right) \right) \end{aligned} \quad (30.15)$$

Now we are ready to use the $n \rightarrow \infty$ limit in a way reminiscent of the famous Kac-Berlin handling of the spherical model, namely n is a large parameter for the Laplace method (steepest descents) of evaluating an

integral. The argument of the exponent in (30.15) causes the integrand to decrease strongly with N_k while the factor $N_k^{n/2} = \exp[\frac{1}{2} n \log N_k]$ grows rapidly. The principal contribution to H'' comes where this product is maximum, that is, where N_k (for $k > \Lambda/s$) satisfies

$$0 = \frac{\partial}{\partial N_k} \left[\frac{1}{2} n \log N_k - ak^2 N_k - \int d^d x U \left(\sum_i \phi_i^{(<)2}(x) + L^{-d} \sum_{k' > \Lambda/s} N_{k'} \right) \right] \quad (30.16)$$

For convenience we introduce some further notation:

$$\begin{aligned} t(y) &= \frac{\partial U(y)}{\partial y} \\ \rho &= L^{-d} \sum_{k > \Lambda/s} N_k \\ \phi^{(<)2} &= \sum_i [\phi_i^{(<)2}(x)]^2 \\ \bar{N}_k &= \text{solution of (30.16) for the indicated } k \end{aligned} \quad (30.17)$$

In this notation (30.16) yields

$$\begin{aligned} \bar{N}_k &= \frac{1}{2} n (ak^2 + t(\phi^{(<)2} + \bar{\rho}))^{-1} \\ \bar{\rho} &= L^{-d} \sum_{k > \Lambda/s} \bar{N}_k \end{aligned} \quad (30.18)$$

(Note that the integral $\int d^d x$ was canceled by a factor L^{-d} .) The quantities \bar{N}_k are functions of $\phi^{(<)2}(x)$. By the Laplace method (to lowest order, i.e., dropping $\log n$ terms) H'' is obtained simply by substituting \bar{N}_k for N_k in (30.15); thus

$$H'' = \sum_{k > \Lambda/s} \left\{ \frac{1}{2} n \log \bar{N}_k - ak^2 \bar{N}_k - \int d^d x U(\phi^{(<)2}(x) + \bar{\rho}) \right\} - \sum_{k < \Lambda/s} ak^2 N_k \quad (30.19)$$

with $\bar{\rho}$ given by (30.18).

The first part of the “renormalization group” transformation has been done—elimination of shorter wavelengths—and now we must rescale so that the new system (i.e., Hamiltonian, space, and fields) is of the same form as the old one with the only possible change being in the parameters

that enter the Hamiltonian. Now, rescaling can be tricky business; the same kind of confusion can enter as in discussion of “active” and “passive” coordinate transformations in other areas of physics. We shall try to be careful on this point so as to save the reader the pain experienced by the author in preparing this section. What we do below is merely rename coordinates and fields—just a change of variables. After these changes we notice that the Hamiltonian in terms of the renamed objects looks just like the Hamiltonian in terms of the original objects, except for the appearance of different parameters. At that stage we call the things with the new names coordinates and fields and say that we have rescaled to new parameter values. This puts the conceptual leap at the end—a disadvantage perhaps—but it makes the arithmetic easier along the way. At various stages we insert motivational asides, to explain possibly obscure definitions, but these in no way interfere with the calculation. Define a set of fields

$$\psi_i(k) = s^{-1+\eta/2} \phi_i\left(\frac{k}{s}\right) \quad 0 < k < \Lambda \quad (30.20)$$

Motivational aside: In the end ψ_i will be the new fields. Their arguments extend over the entire “Brillouin zone,” $0 < k < \Lambda$, but they are defined only in terms of the ϕ_i 's left *after* the elimination through integration described above. The size of the fields has also been rescaled by a factor $\alpha_s = s^{1-\eta/2}$. The reason for this particular dependence on s is that we want the product of two successive scale transformations by s and s' to be the same as a single transformation by ss' , hence we demand $\alpha_s \alpha_{s'} = \alpha_{ss'}$. This requires α_s to be a power of s and we have written the power as $1-\eta/2$ because η turns out to be related to a critical exponent by that name.

With this definition consider the kinetic energy part of H'' in (30.19). Specifically

$$\begin{aligned} a \sum_{k < \Lambda/s} k^2 N_k &= a \sum_{i, k < \Lambda/s} k^2 |\phi_i(k)|^2 \\ &= a \sum_{i, q < \Lambda} \left(\frac{q}{s}\right)^2 |\phi_i\left(\frac{q}{s}\right)|^2 \\ &= a \sum_{i, q < \Lambda} q^2 \frac{s^{2-\eta}}{s^2} |\psi_i(q)|^2 = as^{-\eta} \sum_{i, q < \Lambda} q^2 |\psi_i(q)|^2 \end{aligned} \quad (30.21)$$

Aside: Thus the kinetic energy keeps its form, apart from the factor $s^{-\eta}$.

In H'' we also have integrals over x and the function $\phi^{(<)}$ depending on x , as well as the implicit dependence of $\bar{\rho}$ on $\phi^{(<)}$, hence on x . Define

then

$$\psi_i(x) = s^{d/2} L^{-d/2} \sum_{k < \Lambda} \exp(ikx) \psi_i(k) \quad (30.22)$$

Aside: About the factor $s^{d/2}$... When ψ assumes its rightful role as a rescaled field, ψ_{ik} and $\psi_i(x)$ should be Fourier transforms of one another. The factor $L^{-d/2}$ that appears in the original definition (30.7) is proportional to the square root of the unit cell in k -space and is the factor needed in order that the limit $L \rightarrow \infty$ take one from the Fourier series to the Fourier integral. The momenta for the fields ψ have the same size Brillouin zone but the number of individual k 's is cut down by s in each direction. Hence the side of a unit cell is not $1/L$ but s/L , s times as large.

With a change of variables $p = k/s$ (but no Jacobian in going from the discrete sum over k to that over p), (30.22) yields

$$\psi_i(x) = s^{-1+\eta/2+d/2} \phi^{(<)}(sx) \quad (30.23)$$

The potential energy becomes therefore ($y = x/s$)

$$\int d^d x U(\phi^{(<)}(x) + \bar{\rho}) = \int d^d y s^d U(s^{2-\eta-d} \psi^2(y) + \bar{\rho}) \quad (30.24)$$

with $\bar{\rho}$ understood to depend on $s^{2-\eta-d} \psi^2(y)$.

Equations 30.21 and 30.24 give H'' in terms of ψ . Now we reinterpret. Let H' be a new Hamiltonian, numerically equal to H'' but having its dependence on the fields expressed by

$$H' = \int d^d x [a'(\Delta\phi')^2 + U'(\phi'^2)] \quad (30.25)$$

and we identify

$$\phi'(x) = \psi(x) \quad (30.26)$$

$$a' = as^{-\eta} \quad (30.27)$$

$$U'(\phi'^2(x)) = s^d U(s^{2-\eta-d} \psi^2(x) + \bar{\rho}) \quad (30.28)$$

It is more convenient to replace (30.28) by an equation for the function t , and by taking a derivative of (30.28) with respect to ϕ'^2 we get

$$t'(\phi'(x)^2) = s^{2-\eta} t(s^{2-\eta-d} \psi^2(x) + \bar{\rho}) \quad (30.29)$$

Equations 30.27 and 30.29 are the renormalization transformation [with the definition of terms ψ , \bar{N}_k , and $\bar{\rho}$ given by (30.26), (30.20), (30.16),

(30.17), and (30.18) implicit]. Thus the parameter a has become a' and from (30.29) and (30.3) there is some new set of coefficients t'_1, t'_2, \dots . As mentioned earlier this is the form derived by Ma.

The first problem to tackle is to find the fixed points of the transformation. As indicated above, values of the parameters for which there is a fixed point correspond to systems at their critical points where there is scale invariance. Thus the fixed points occur where the correlation length ξ is either zero or infinite.

By (30.27) a critical point for which $\eta \neq 0$ will have $a=0$ or $a=\infty$. This would correspond, if one were to discretize space, to particles completely uncoupled from one another or to those coupled by powerful harmonic forces to the exclusion of all else. For various reasons such fixed points are considered less interesting than those for which a survives as a finite number, and we therefore consider only $\eta=0$. This makes the actual value of a irrelevant and we henceforth take it to be 1.

The renormalization group transformation can therefore be summarized as

$$t'(\psi^2) = s^2 t(s^{2-d} \psi^2 + \bar{\rho}) \quad (30.30)$$

$$\bar{\rho} = L^{-d} \sum_{k > \Lambda/s} \frac{n/2}{k^2 + t(s^{2-d} \psi^2 + \bar{\rho})} \quad (30.31)$$

with (30.31) being an implicit equation for $\bar{\rho}$. So as to deal with nonintegral dimension it is also convenient to put [in (30.31)]

$$L^{-d} \sum_{k > \Lambda/s} = K_d \int_{\Lambda/s}^{\Lambda} k^{d-1} dk \quad \text{with } K_d = 2[2^d \pi^{d/2} \Gamma(d/2)]^{-1} \quad (30.32)$$

By (30.30) and (30.32), (30.29) becomes

$$\bar{\rho} = \frac{n}{2} K_d \int_{\Lambda/s}^{\Lambda} k^{d-1} dk \frac{1}{k^2 + t'(\psi^2)/s^2} \quad (30.33)$$

Suppose now that t has a fixed point (function), which we designate t^* , which is finite. Then allowing $s \rightarrow \infty$ in (30.33) yields

$$\bar{\rho}^* = \frac{n K_d \Lambda^{d-2}}{2(d-2)} \equiv c \quad (30.34)$$

a constant. Inserting this in (30.30) and taking $d > 2$ gives the requirement

$$t^*(c) = 0$$

Next define ζ by

$$s^{2-d}\phi^2 + \bar{\rho} = \left(1 + \frac{\zeta}{s^2}\right)c \quad (30.35)$$

with $\bar{\rho}$ depending on ϕ^2 as in (30.31) or (30.33). By a change variable $p=ks$, (30.35) can be brought to

$$\frac{\phi^2}{c} = 1 + \zeta s^{d-4} - (d-2)\Lambda^{2-d} \int_{\Lambda}^{\Lambda s} dp p^{d-1} \left[\frac{1}{t'+p^2} - \frac{1}{p^2} \right] \quad (30.36)$$

The next step in finding the fixed point is to expand t :

$$t' = s^2 t \left(c \left(1 + \frac{\zeta}{s^2} \right) \right) = u\zeta + O(s^{-2}) \quad (30.37)$$

with

$$u = c \frac{dt(y)}{dy} \Big|_{y=c} \quad (30.38)$$

In (30.36) we would like to drop the term ζs^{d-4} , and Ma argues from (30.37) that for $s \rightarrow \infty$ this term behaves like $\zeta s^{d-4} \sim (t^*/u^*) s^{d-4} \rightarrow 0$ for $d < 4$. This gives [from (30.36)] for $s \rightarrow \infty$ (and thus $t \rightarrow t^*$)

$$\frac{\phi^2}{c} = 1 - (d-2)\Lambda^{2-d} \int_{\Lambda}^{\infty} dp p^{d-1} \left[\frac{1}{t^*+p^2} - \frac{1}{p^2} \right] \quad (30.39)$$

Equation 30.39 provides ϕ^2 as a function of t^* , hence the inverse function. This allows easy numerical calculation of the fixed point function $t^*(\phi^2)$.

Similarly for $d > 4$ one can see that $t^* = 0$, and the fixed point Hamiltonian is the trivial

$$H^* = \int d^d x (\nabla \phi)^2 \quad (30.40)$$

For $d=4$ (30.40) still holds.

Consider some d between 2 and 4, say $d=3$, and the nontrivial fixed point ($\eta=0, t^* \neq 0$) that we found there. We now discuss the physical significance of $\eta=0$. Let $g(\mathbf{r})$ be the two-point correlation function, $\langle \phi(\mathbf{r})\phi(0) \rangle$, the expectation being taken with respect to the Hamiltonian H . In mean field theory, at the critical point this goes like $|r|^{-d+2}$, but in general need not behave that way. This leads to the definition of an exponent η' (usually designated η , but we temporarily call it η' to avoid

confusion with the quantity we have already defined), such that

$$g(r) \sim \frac{1}{r^{d-2+\eta'}} \quad (30.41)$$

The Fourier transform of g is $G(k)$, and satisfies

$$G(k) = \langle |\phi_{ik}|^2 \rangle_H \quad (30.42)$$

where the H subscript appears to emphasize the probability distribution being used and G is presumed independent of i . For g satisfying (30.41), we have

$$G(k) \sim k^{-2+\eta'} \quad (30.43)$$

by the properties of Fourier transformation. Now let us see what G looks like in terms of the scaled quantities. Following the point of view expressed above, we simply are changing variables [in (30.42)] to new quantities ψ . Thus for some k less than Λ/s (dropping the subscript i)

$$\begin{aligned} G(k) &= \frac{\int \prod_{k'} d\phi_{k'} e^{-H} |\phi(k')|^2}{\int \prod_{k'} d\phi_{k'} e^{-H}} \\ &= \frac{\int \prod_{k' < \Lambda/s} d\phi_{k'} e^{-H''} (s^{1-\eta/2})^2 |\psi(ks)|^2}{\int \prod_{k' < \Lambda/s} d\phi_{k'} e^{-H''}} \end{aligned} \quad (30.44)$$

where we have integrated out $k' > \Lambda/s$ (possible since the given k is less than Λ/s) and obtained H'' as described above. Also we have simply rewritten $\phi(k')$ as $s^{1-\eta/2}\psi(ks)$, following (30.20). Now we change integration variables to $d\psi(k')$ [or $d\phi'(k')$ which is the same thing, by the definition (30.26)] and any Jacobian needed cancels in the numerator and denominator. The result is

$$G(k) = s^{2-\eta} \frac{\int \prod_{k' < \Lambda/s} d\psi(k') e^{-H''} |\psi(ks)|^2}{\int \prod_{k' < \Lambda/s} d\psi(k') e^{-H''}} \quad (30.45)$$

Next we reinterpret. The ratio of integrals is just the expectation value using the new Hamiltonian H' . The dummy index k' runs over the entire Brillouin zone, but the argument of ψ in $|\psi(ks)|^2$ remains ks since that is not a dummy index. If $G'(k)$ is the two-point function with the new Hamiltonian (i.e., $G'(k) = \langle |\phi'(k)|^2 \rangle_{H'}$) then we have found that

$$G(k) = s^{2-\eta} G'(ks) \quad (30.46)$$

At the critical point, the scale transformation should not change the k space functional dependence of the pair correlation function at all so that combining (30.46) with (30.43), which also holds at the critical point, we have

$$1 = s^{\eta' - \eta}$$

implying that η' and η are the same. Consequently our finding of a critical fixed point for which $\eta=0$ shows that for that critical point the pair correlation function behaves as in mean field theory.

What we have just shown is a way in which renormalization theory fixes a critical exponent. Generally the derivation is more involved than that given above and the scale transformation needs to be linearized about its fixed point. Critical exponents can then be related to the eigenvalues of the linearized transformation. The reader is referred to Ma's article or to the extensive literature available for further information.

A good ideal of philosophizing can be done about phase transitions, universality, and the fundamental problems of systems with an infinite number of degrees of freedom and an infinity of scales on which nontrivial coupling of modes (whatever they are) takes place. Such systems are drawn not only from statistical mechanics but from quantum field theory and hydrodynamics as well. However, we forbear such philosophizing (since literature abounds) and ask, from the standpoint of the functional integral technician, what has been done. The place where a functional integral was actually carried out was (30.15) to (30.19) and once again the method of choice was Laplace's method (steepest descents), conveniently possible because of the large parameter, n (number of components of the field). In other renormalization group (scale) transformations the functional integral also appears, because it is a good way to separate short range and long range behavior, but its actual implementation on short wavelength modes is different. Laplace's method though, as we have seen repeatedly in this book, certainly can help the calculation.

NOTES

Early ideas on scale transformations and critical exponents appear in

L. P. Kadanoff, W. Gotze, D. Hamblen, R. Hecht, E. A. S. Lewis, V. V. Palciauskas, M. Rayl, J. Swift, D. Aspnes, and J. Kane, *Rev. Mod. Phys.* **39**, 395 (1967)

The implementation of scaling ideas through the “renormalization group” and the smearing of short wavelength modes by means of functional integration appear in

K. G. Wilson, *Phys. Rev. B* **4**, 3174, 3184 (1971)

From then on the field developed rapidly and within a short time review papers appeared. The worked example given in the foregoing section was taken from

S. Ma, *Rev. Mod. Phys.* **45**, 589 (1973)

and the related references

S. Ma, *Phys. Lett. A* **43**, 475 (1973)

S. Ma, *J. Math. Phys.* **15**, 1866 (1974)

My presentation follows the *Rev. Mod. Phys.* paper quite closely. The model itself, the infinite component field, was studied by Stanley:

H. E. Stanley, *Phys. Rev.* **176**, 718 (1968)

who showed that in this limit the model was equivalent to the spherical model of Berlin and Kac:

T. H. Berlin and M. Kac, *Phys. Rev.* **86**, 821 (1952)

Not all scale transformations use functional integration, and the methods that most resemble the original ideas of Kadanoff were first suggested by

T. Niemeijer and J. M. J. van Leeuwen, *Phys. Rev. Lett.* **31**, 1411 (1973)

T. Niemeijer and J. M. J. van Leeuwen, *Physica* **71**, 17 (1974)

who do “coordinate space” scaling. A general overview of scaling can be found in

C. Domb and M. S. Green, *Phase Transitions and Critical Phenomena*, Vol. 6, Academic Press, London, 1976.

See also

K. G. Wilson, *Scientific American*, August 1979

The method of “decimation” related to the coordinate space scaling mentioned above has in a way been used for the evaluation of functional integrals by

M. Nauenberg, F. Kuttner, and M. Furman, *Phys. Rev. A* **13**, 1185 (1976)

A book that relies heavily on functional integral techniques and goes deeply into many topics of contemporary interest in phase transitions is

D. J. Amit, *Field Theory, the Renormalization Group, and Critical Phenomena*, McGraw-Hill, New York, 1978

THIRTY-ONE

Phase Space Path Integral

Here is a difficult form of the path integral. In this book we have occasionally gone beyond the bounds of mathematical propriety, but always in the belief that the answer was coming out right and that if the details were hard to justify it only meant that there was an interesting mathematical challenge. Not so with the path integral in phase space; in this method formal tricks of great power (e.g., canonical transformations) can give just plain wrong answers, and to avoid them one must pay careful attention to error terms that are ordinarily negligible. However, the method does have its advantages (in particular, there are many systems for which the Hamiltonian formulation is better than the Lagrangian), and used with caution it can provide useful physical information.

The method starts from the Trotter product formula and differs from what was done in Section 1 only in leaving the “ p ” integrals undone. Thus

$$\begin{aligned} G(x, t; y) &= \left\langle x \left| \exp \left(- \frac{iHt}{\hbar} \right) \right| y \right\rangle \\ &= \lim_{N \rightarrow \infty} \left\langle x \left| \left[\exp \left(- \frac{ip^2 t}{2m\hbar N} \right) \exp \left(- \frac{iVt}{\hbar N} \right) \right]^N \right| y \right\rangle \quad (31.1) \end{aligned}$$

where $H = p^2/2m + V$. Equation 31.1, namely the statement that the limit on the right is the propagator, is just the Trotter product formula. Between each pair of operators in the product we insert a resolution of the identity, alternately using coordinate and momentum space eigenstates. This gives

$$\begin{aligned} G(x, t; y) &= \lim_{N \rightarrow \infty} \int dp_1 dx_1 dp_2 dx_2 \cdots dp_{N-1} dx_{N-1} dp_N \langle x | e^{p^2 \epsilon / 2m} | p_N \rangle \\ &\quad \times \langle p_N | e^{V \epsilon} | x_{N-1} \rangle \cdots \langle x_1 | e^{p^2 \epsilon / 2m} | p_1 \rangle \langle p_1 | e^{V \epsilon} | y \rangle \quad (31.2) \end{aligned}$$

where

$$\epsilon = -\frac{it}{\hbar N}$$

and we shall also employ the notation $x_0 = y, x_N = x$. As usual the profit from (31.2) arises from the fact that

$$\begin{aligned} \left\langle x_i \left| \exp\left(\frac{p^2 \epsilon}{2m}\right) \right| p_i \right\rangle &= \exp\left(\frac{p_i^2 \epsilon}{2m}\right) \langle x_i | p_i \rangle \\ \langle p_i | \exp(V\epsilon) | x_{i-1} \rangle &= \exp(V(x_{i-1})\epsilon) \langle p_i | x_{i-1} \rangle \end{aligned} \quad (31.3)$$

Taking

$$\langle x | p \rangle = (2\pi\hbar)^{-1/2} \exp\left(\frac{ipx}{\hbar}\right) \quad (31.4)$$

gives δ -function normalization

$$\int dx \langle p | x \rangle \langle x | p' \rangle = \delta(p - p')$$

Equations 31.3 and 31.4 are inserted in (31.2) to give

$$\begin{aligned} G(x, t; y) &= \lim_{N \rightarrow \infty} \int dp_1 dx_1 \cdots dp_N (2\pi\hbar)^{-N} \\ &\times \exp\left\{ \epsilon \sum_{j=1}^N \left(\frac{p_j^2}{2m} + V(x_{j-1}) \right) + \frac{i}{\hbar} \sum_{j=1}^N p_j (x_j - x_{j-1}) \right\} \end{aligned} \quad (31.5)$$

Recalling the definition of ϵ , (31.5) becomes

$$G(x, t; y) = \lim_{N \rightarrow \infty} \int dp_1 dx_1 \cdots dp_N (2\pi\hbar)^{-N} \exp\left[\frac{i}{\hbar} \int_0^t \left(p \frac{dx}{d\tau} - H \right) d\tau \right] \quad (31.6)$$

where the integral in the exponent means neither more nor less than

$$\int_0^t \left(p \frac{dx}{d\tau} - H \right) d\tau \equiv \frac{t}{N} \sum_{j=1}^N \left[p_j \frac{(x_j - x_{j-1})}{t/N} - \frac{p_j^2}{2m} - V(x_{j-1}) \right] \quad (31.7)$$

As in Section 1 we interpret the formula (31.6) as a sum over paths. The integration variables suggest that one is adding trajectories in phase space. Moreover, the argument of the exponent is just the action expressed in phase space coordinates. However, the appearance of $2N$ (rather than N) variables coupled with the fact that a single point in phase space determines the classical path causes some difficulty. Here are two interpretations. (1) The particle's trajectory goes from x_j to x_{j+1} in the time interval $jt/N \leq \tau \leq (j+1)t/N$, $j=0, 1, \dots, N-1$; during that time the momentum equals p_{j+1} . The quantity x is continuous from interval to interval, while p jumps. Recalling the discontinuous derivatives of the Feynman path integrals the jumps in p should not be surprising. However, by contrast, because the dynamical variables are overspecified, during the interval $\tau \in [jt/N, (j+1)t/N]$ the paths are *not* classical paths. (2) The time period $[0, t]$ is considered to be broken into $2N$ intervals (let $\Delta = t/2N$) and with each end point of each interval is alternately associated either an x value or a p value. Within the interval the particle travels the classical path defined by the boundary values, a single x and a single p , albeit specified at different times. Thus $x(0) = x_0 = y$; $p(\Delta) = p_1$, $x(2\Delta) = x_1$, $p(3\Delta) = p_2, \dots$, $x(2j\Delta) = x_j$, $p((2j-1)\Delta) = p_j, \dots$. To evaluate the action we assume that p and x can be approximated by either end point value. Thus for the interval $2j\Delta \leq \tau \leq (2j+1)\Delta$ the initial x is x_j and the final p is p_{j+1} . We approximate H by $p_{j+1}^2/2m + V(x_j)$. The term $p(dx/d\tau)d\tau$ is approximated by $p_{j+1}(x_j^{(+)} - x_j^-)$ where $x_j^{(+)}$ is the final and unknown value taken by x at time $(2j+1)\Delta$ as determined by the classical path. Similarly, for the next time interval $\tau \in [(2j+1)\Delta, (2j+2)\Delta]$ the action is approximated by

$$p_{j+1}(x_{j+1} - x_{j+1}^-) - \left(\frac{p_{j+1}^2}{2m} + V(x_{j+1}) \right) \Delta$$

where x_{j+1}^- is the (unknown) value of x at time $(2j+1)\Delta$ as determined by the classical path between p_{j+1} and x_{j+1} . The action obtained in this way agrees with the expressions in (31.5) to (31.7) only to the extent that the difference between $x_j^{(+)}$ and x_{j+1}^- can be neglected. This does not say that (31.5) to (31.7) are wrong. On the contrary, it is only a limitation on the phase space path picture that we are trying to attach to those formulas.

From (31.5) we could immediately recover the Lagrangian path integral by doing the p_j integrations; in fact this was just the derivation of Section 1.

Quite a few applications have begun from the phase space path integral and one can even use it for gaining insight into classical mechanics. First, a hand waving stationary phase argument (for $\hbar \rightarrow 0$) applied to

(31.6) gives Hamilton's equations for the classical limit. Alternatively one can Fourier transform (31.6) with respect to E , yielding

$$G(E) \sim \int_0^\infty e^{iEt/\hbar} G(t) dt \sim \int_{\substack{\text{paths of} \\ \text{energy} E}} dp_1 dx_1 \dots dp_N dx \exp\left(\frac{i}{\hbar} \int p dx\right) \quad (31.8)$$

The last step follows by breaking up the sum over paths in (31.6) into those for which H is fixed making the Fourier transform give essentially a δ -function. Then (31.8) implies (suggests?) that the main contribution in the semiclassical approximation is that path for which $\int p dx$ is stationary at fixed E . Of course this is one of the variational statements of classical mechanics.

My own favorite application of (31.6) is unfortunately false. However, since the conclusions of the forthcoming false theorem are sometimes true even when one would not have guessed them to be, the material will be presented anyway.

Suppose that $H=0$. Then each integral over p_j in (31.5) gives a δ -function in successive values of x . The x integrals are then no less trivial, and we have

$$G(x, t; y) = \delta(x - y) \quad (31.9)$$

This is not a surprising result, since the propagator in this case is just the identity operator.

The power of Hamiltonian mechanics arises from the ability to make changes of variables in phase space. New coordinates P and X are defined in terms of the old ones (p and x) by means of a function of certain pairs of them and of time. Such a canonical transformation can be defined for example by a function $F(x, P, t)$ by

$$p = \frac{\partial F}{\partial x} \quad (31.10a)$$

$$X = \frac{\partial F}{\partial P} \quad (31.10b)$$

For appropriately nonsingular F this defines a 1-1 transformation of phase space. The new equations of motion are also canonical [that is, $dX/dt = \partial H'/\partial P$, and so on] in terms of the new Hamiltonian

$$H' = H + \frac{\partial F}{\partial t} \quad (31.11)$$

A central, and for us very useful, feature of this transformation is what happens to the action. For a path $(x(\tau), p(\tau))$ and its image under the transformation

$$\int_0^t \left(p \frac{dx}{d\tau} - H \right) d\tau = \int_0^t \left(P \frac{dX}{d\tau} - H' \right) d\tau + F(x'', P'', t) - F(x', P', 0) \quad (31.12)$$

with twice and once primed quantities referring to final and initial values respectively. The integrals in (31.12) are Riemann integrals along smooth paths. Another property of canonical transformations is that the Jacobian of the transformation is unity

$$\frac{\partial(P, X)}{\partial(p, x)} = 1 \quad (31.13)$$

A special transformation is that for which the new Hamiltonian is zero. When (31.10a) is inserted in (31.11) one gets the Hamilton-Jacobi equation for F :

$$H\left(x, \frac{\partial F}{\partial x}, t\right) + \frac{\partial F}{\partial t} = 0 \quad (31.14)$$

Since $H' = 0$, the new coordinates X and P are constants of the motion.

Returning to the path integral we proceed formally. For some path $(x(\tau), p(\tau))$ contributing to the sum in (31.6) there is a corresponding path $(X(\tau), P(\tau))$ in the coordinate system defined by the Hamilton-Jacobi transformation. If F is the generator of this canonical transformation then by (31.12)

$$\int (p dx - H d\tau) = \int P dX + F(x'', P'', t) - F(x', P', 0) \quad (31.15)$$

Note that the special nature of this transformation has eliminated the Hamiltonian on the right hand side of (31.15). The only trajectory dependent term on the right is $\int P dX$, since F depends on the end points alone. It would thus appear that the general case has been reduced to the zero Hamiltonian case. There is still a Jacobian to evaluate, since the path integral in (31.6) has $dp_1 dx_1 \dots dp_N$ as integration variables and we would like to use P and X . However, (31.13) suggest that most of the integration variable changes will yield unity. Counting variables there are $(N-1)$ x integrations and N p integrations. If the same number of X and P integrations occur, there will be a single $p \rightarrow P$ transformation whose

Jacobian needs to be included. Just what that transformation is requires stating things more carefully than we have. In addition the functional integral over P and X will yield a function of X while the original propagator involves x so that $X \rightarrow x$ Jacobians are also needed at both ends of the integration. All these Jacobians involve terms of the form $\partial X / \partial x$ or $\partial P / \partial p$ which are various second derivatives of F . We shall not provide any details, but it should not be too implausible that these factors eventually become the Van Vleck determinant $[\det \partial^2 S / \partial x''_\alpha \partial x'_\beta]^{1/2}$ when S is identified with F .

Having provided some motivation for ignoring the Jacobian we have now brought the propagator to the form

$$G = \int dP(\tau) dX(\tau) \exp \left[\frac{i}{\hbar} \int P dX + F(x, P'', t) - F(y, P', 0) \right] \quad (31.16)$$

There is a good deal of freedom in selecting the generalized momenta and a useful choice is to let P be the initial position of the particle. That is, under the equations of motion P is a constant and as that constant—which is in effect a trajectory label—we take y , the initial position. With this choice, $F(x, P, t)$ is just $S(x, y, t)$ the action along the associated classical trajectory. This is a function of the end points only and can be removed from the functional integral. The action at time zero, $F(x, P, 0)$ is zero. The remaining functional integral is trivial and yields (as for systems with zero Hamiltonian) a δ -function in $X'' - X'$, the initial and final positions expressed in the new coordinates. This δ -function just enforces the condition that the system has traveled the classical path, since in the new coordinates X is also constant of the motion. The propagator is therefore

$$G = e^{iS(x, y, t)/\hbar} \quad (31.17)$$

with S the action along the classical path from y to x . In (31.17) the δ -function in $X'' - X'$ disappeared as a result of transforming back to x and y as variables and the van Vleck determinant arising from various Jacobians has been ignored. In case there are several paths between given end points, a sum of such contributions would presumably appear.

Equation 31.17 clearly corresponds to the semiclassical approximation derived earlier (Section 13) from the Lagrangian path integral.

Because the foregoing derivation proceeded so cavalierly it may be difficult to see where the truly erroneous steps took place and why the apparent panacea of the canonical transformation only works to lowest order in \hbar . In fact, there are a number of weak points in the “proof.” Since

it is possible that some of them are curable either in general or under special circumstances, we mention several.

First recall that the path in phase space that we defined above is discontinuous, in contrast to the usual coordinate space path integral trajectories where it is only the derivative that goes bad. To see where this enters, consider again (31.15). This was written as if a continuous path led from $(x(0), p(0))$ to $(x(t), p(t))$. Instead recall that the path can be thought of as going from x_i to x_{i+1} in time t with momentum p_{i+1} and then from x_{i+1} to x_{i+2} with (in general different) momentum p_{i+2} . For each of the short time paths we make the canonical transformation to get ($i = 1, 2$)

$$p_2(x_2 - x_1) = \int P dX + F(x_2, P_2^{(+)}, \Delta t) - F(x_1, P_2^{(-)}, 0)$$

$$p_3(x_3 - x_2) = \int P dX + F(x_3, P_3^{(+)}, 2\Delta t) - F(x_2, P_3^{(-)}, \Delta t) \quad (31.18)$$

The quantities $P_i^{(\pm)}$ are the end point values of the canonical momenta and $P_{i-1}^{(+)}$ need not equal $P_i^{(-)}$ because the trajectory is not necessarily the classical trajectory. Now what one would like to do is add expressions like those appearing in (31.18) for all i and have the intermediate terms in F cancel, for example $F(x_2, P_2^{(+)}, \Delta t)$ cancel against $F(x_2, P_3^{(-)}, \Delta t)$. Then all that would be left is (31.15) with F evaluated at the extreme end points. However, because of the jump in p the quantities $P_3^{(-)}$ and $P_2^{(+)}$ may not be equal and (31.15) needs an additional term

$$\sum_i \frac{\partial F}{\partial P}(P_{i-1}^{(+)} - P_i^{(-)})$$

For the classical path this term vanishes, and to the extent that the particle sticks to the classical trajectory this correction can be ignored. However, for higher corrections in \hbar there is no reason for the sum to vanish as the differences in the P 's are related to the jumps in the velocity for the Lagrangian, which are not small.

The discontinuous path in phase space is what is usually considered in studying the phase space path integral. However, there have been suggestions for continuous paths and so we go on to consider another problem in the “proof.”

The generating function F effects a change of coordinates on the phase space. But we have not considered whether there is a single F giving a unique globally defined transformation, even though our analysis presup-

posed such a transformation. In fact, there is every reason to believe that the (phase space) transformation will be multivalued. This is because when there is more than one path between given end points there is a different Hamilton-Jacobi function (hence generator of canonical transformation) for each of them (since $S = \int L dt$ along the path). Moreover, these functions *are* related to one another, since at focal (conjugate) points different paths merge and we expect branch points in the Hamilton-Jacobi transformation (see page 224). Trouble in this area should have been evident when (31.17) was generalized by fiat to include a sum of several terms.

It may seem a bit strange to go to so much trouble to present a "proof" of (31.17) only to follow that proof by an analysis of its errors. Our purposes have been both exploratory and didactic. Exploratory, in the sense that there are many situations where the sum over classical paths *is* exact (see the appendix to Section 6), and it is conceivable that the phase space approach may give insight into them. The didactic message of our presentation has over the years made its absence known through the appearance of works in which formal operations with the phase space path integral led to incorrect conclusions.

On the optimistic side, suppose that the difficulty with discontinuous paths were overcome. Then one might look for special systems that behave especially well insofar as defining global transformations. For example, systems of n degrees of freedom that have solitons are known to have n globally defined invariants. Would the quantum propagator for such systems be given in terms of classical paths only? Statements that seem related to this have been made, but I am not sure of their exact meaning, of just what were the authors' claims regarding their validity or of how they are tied, if at all, to the methods described here.

As an example of pitfalls of the phase space path integral, consider any function $f(p, x)$ of momentum and position. In the usual quantization schemes there is an ambiguity in what operator to associate with f . One might have thought that a unique ordering rule could be generated simply by putting $f(p, x)$ into the argument of the exponent in (31.6) and integrating. Not so. Subtleties in phase space path integration make ordering no less ambiguous than for Lagrangian path integrals (see Sections 8, 24, and 27). The ordering problem has recently been revived in connection with real valued functional integrals (rather than path integrals with $e^{iS/\hbar}$) for the Fokker-Planck equation.

The phase space path integral finds its most significant applications when two elements are present: (1) for some reason one is committed to a Hamiltonian rather than a Lagrangian formalism and (2) semiclassical results are adequate. Examples of such applications will be given in the notes below.

NOTES

The phase space path integral has been discovered and rediscovered repeatedly since 1951. Some references are

- R. P. Feynman, An Operator Calculus Having Applications in Quantum Electrodynamics, *Phys. Rev.* **84**, 108 (1951), Appendix B
- W. Tobocman, Transition Amplitudes as Sums over Histories, *Nuovo Cim.* **3**, 1213 (1956)
- H. Davies, Hamiltonian Approach to the Method of Summation over Feynman Histories, *Proc. Camb. Phil. Soc.* **59**, 147 (1963)
- C. Garrod, Hamiltonian Path-Integral Methods, *Rev. Mod. Phys.* **38**, 483 (1966)

More abstract forms of the phase space path integral have also been developed in which one does not rely on the path's breakup as given in (31.5). This might be the right way to avoid the mire of $x_i^{(\pm)}$ end points and other inelegant complications that are encountered in trying to work with (31.5) and (31.6). These forms of the path integral are given in

- C. DeWitt-Morette, A. Maheshwari, and B. Nelson, Path Integration in Phase Space, *Gen. Rel. and Grav.* **8**, 581 (1977)
- M. M. Mizrahi, Phase Space Path Integrals, without Limiting Procedure, *J. Math. Phys.* **19**, 298 (1978)

Other definitions and presentations, covering the whole spectrum of rigor and pedagogy, are

- K. Gawedzki, Construction of Quantum-Mechanical Dynamics by Means of Path Integrals in Phase Space, *Rep. Math. Phys.* **6**, 327 (1974)

which uses techniques analogous to those used by Ito in the usual path integral.

- A. L. Alimov, Hamiltonian Form of a Feynman Path Integral, *Theor. and Math. Phys.* **20**, 837 (1975) [trans. from *Teor. Mat. Fizika* **20**, 302 (1974)]
- P. Pearle, Finite-Dimensional Path Summation Formulation for Quantum Mechanics, *Phys. Rev. D* **8**, 2503 (1973)
- M. Clutton-Brock, Feynman's Kernel and the Classical Path, *Path. Camb. Phil. Soc.* **61**, 201 (1965)
- R. Fanelli, Canonical Transformations and Phase Space Path Integrals, *J. Math. Phys.* **17**, 490 (1976)

Fanelli gets his canonical transformations by beginning with unitary transformations in quantum mechanics, a generally safer procedure than the reverse. He also points out instances where operator ordering ambiguities complicate matters and discusses other limitations of the method.

The need to define p and q at different times in the phase space path

integral is discussed by

A. Katz, *Classical Mechanics, Quantum Mechanics, Field Theory*, Academic Press, New York, 1965

Ambiguities in previously suggested operator ordering schemes were emphasized by

L. Cohen, Hamiltonian Operators via Feynman Path Integrals, *J. Math. Phys.* **11**, 3296 (1970)

who showed that a variety of orderings could be obtained. A summary that is short and to the point is given by

J. S. Dowker, Path Integrals and Ordering Rules, *J. Math. Phys.* **17**, 1873 (1976)

I think that the main wisdom on this subject is to be found in the following paper:

L. van Hove, Sur le Probleme des Relations entre les Transformations Unitaires de la Mecanique Quantique et les Transformations Canoniques de la Mecanique Classique, *Mem. Acad. Roy. Belg.* **26**, 610 (1951)

Van Hove seeks to relate the unitary transformations of quantum mechanics and the canonical transformation of classical mechanics and finds a 1-1 correspondence between continuous transformations *only* in those cases where the generating function is at most quadratic. This makes it reasonable that the canonical transformation on the phase space path integral recover the semiclassical propagator, since the semiclassical propagator is exact for quadratic Lagrangians. It makes it just as reasonable that the transformation not succeed in more general cases.

Ordering problems and a functional integral representation of solutions of the Fokker-Planck equation has attracted recent interest. Both coordinate and phase space functional integrals have been used and attempts have been made to resolve ordering ambiguities by functional integral definitions without discretization. Some of the literature (in which can also be found earlier references) is

R. Graham, Lagrangian for Diffusion in Curved Phase Space, *Phys. Rev. Lett.* **38**, 51 (1977)

H. Dekker, Functional Integration and the Onsager-Machlup Lagrangian for Continuous Markov Processes in Riemannian Geometries, *Phys. Rev. A* **19**, 2102 (1979)

F. Langouche, D. Roekaerts, and E. Tirapegui, Perturbation Expansions through Functional Integrals for Nonlinear Systems, *Phys. Rev. D* **20**, 433 (1979)

F. Langouche, D. Roekaerts, and E. Tirapegui, Discretization Problems of Functional Integrals in Phase Space, *Phys. Rev. D* **20**, 419 (1979)

A. C. Hirshfeld, Canonical and Covariant Path Integrals, *Phys. Lett.* **67A**, 5 (1978)

See also contributions in the book

L. Garrido, R. Segular, and P. J. Shepherd, Eds., *Stochastic Processes in Nonequilibrium Systems*, Proceedings Sitges International School of Statistical Mechanics, Barcelona, Springer Verlag, Berlin 1978

Ordering problems in a quantum field theory and the persistence of ordering ambiguities when doing phase space, canonical, path integral quantization is discussed by

H. Leschke, A. C. Hirshfeld, and T. Suzuki, Dyson-Wick Expansion for Nonlinear Bose Systems, *Phys. Lett.* **67A**, 87 (1978)

Solitons and their properties are reviewed in

A. C. Scott, F. Y. F. Chu, and D. W. McLaughlin, The Soliton: A New Concept in Applied Science, *Proc. IEEE* **61**, 1443 (1973)

Complete integrability, global invariants, and such matters are discussed by

V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, Benjamin New York, 1968

Possible exactness of WKB semiclassical results for systems involving solitons was found by

R. F. Dashen, B. Hasslacher, and A. Neveu, Particle Spectrum in Model Field Theories from Semiclassical Functional Integral Techniques, *Phys. Rev. D* **11**, 3424 (1975)

These authors used methods that in the main had been developed by Gutzwiller (reference below, see also Section 18), and the suspicion of exactness arose because of comparison with field theory results due to S. Coleman. (It was noted, however, that agreement to all orders in perturbation theory is not the same as equality. For asymptotic series there remains the possibility of essentially singular terms, a possibility that is plausible in a WKB context where barrier penetration contributions have just that form). For the quantization of solitons functional integration has been one of the main tools and in fact this is the background to the interest shown by particle physicists in the instanton calculation of Section 29. Much of the work in this area is discussed in the following conference report

J. L. Gervais and A. Neveu, Extended Systems in Field Theory, *Phys. Rep.* **23**, 237 (1976) where functional integral calculations are reported by Gervais, Neveu, A. Jevicki, B. Sakita and others. Various authors there are concerned with the "zero mode" problem that arises because of symmetries of the system (this is the mode that arose from translational invariance in Section 29). More recent work on this same problem is

L. D. Fadeev and V. E. Korepin, *Phys. Lett.* **63B**, 435 (1976)

S. S. Chang, *Phys. Rev. D* **17**, 2595 (1978)

See also

L. D. Fadeev and V. E. Korepin, Quantum Theory of Solitons, *Phys. Rep.* **42**, 1 (1978)

The phase space path integral has been used effectively by a number of authors:

M. C. Gutzwiller, Phase-Integral Approximation in Momentum Space and the Bound States of an Atom, *J. Math. Phys.* **8**, 1979 (1967)

This is the first in a series of articles getting semiclassical approximations for spherically symmetric and nonspherically symmetric potentials. Gutzwiller explores some rather subtle properties of the classical mechanics in the course of these investigations. In this application the phase space integral is merely a step on the way to getting the momentum space propagator in which coordinate space has been Fourier transformed away.

A phase space formalism for spin has been given by

Peter Horvathy, *Feynman Integral for Spin*, CNRS preprint, April 1979

who uses Gawedski's form (mentioned above) of the phase space functional integral.

An application in which the use of the phase space integral is essential is

S. Levit, U. Smilansky, and D. Pelte, A New Semiclassical Theory for Multiple Coulomb Excitation, *Phys. Lett.* **53B**, 39 (1974)

S. Levit and U. Smilansky, The Hamiltonian Path Integrals and the Uniform Semiclassical Approximations for the Propagator, *Ann. Phys.* **108**, 165 (1977)

S. Levit, K. Möhring, U. Smilansky, and T. Dreyfus, Focal Points and the Phase of the Semiclassical Propagator, *Ann. Phys.* **114**, 223 (1978)

In the work of Levit et al. coordinate space is not the best way to define the boundary value problem, and so they are naturally led to the use of a phase space path integral (cf. Section 19). In fact, though, they are mainly interested in the semiclassical approximation and come away unscathed from their use of the phase space object. For their study of caustics (see Section 16) they must go beyond the lowest semiclassical approximation but this apparently does not cause any harm.

In some presentations of field theory methods it has been found convenient to use phase space path integrals (as well as the usual Lagrangian path integral). See:

E. S. Abers and B. W. Lee, Gauge Theories, *Phys. Rep.* **9**, 1 (1973)

This paper provides an excellent background on the use of path integral methods in the renormalization of gauge field theories. For some path integral users, for example for a hypothetical particle physicist lacking in aesthetic sense, this application may be the entire *raison d'être* for path integration. Unfortunately I have not been able to provide in this book the extensive treatment that renormalization via path integrals deserves. For further reading, besides the Abers and Lee article, I can recommend

D. J. Amit, *Field Theory, The Renormalization Group and Critical Phenomena*, McGraw-Hill, New York, 1978

See also the soliton quantization references given above.

As mentioned in Section 27, the phase space path integral and the path integral in the coherent state representation are closely related. See the notes to that section for the appropriate references.

THIRTY-TWO

Omissions, Miscellany, and Prejudices

This section is devoted to a quick overview of some of the many topics not covered in this book. The discussion is brief—sometimes to the point of being merely a list of references.

32.1 FIELD THEORY

A major application of path integration and functional integration is to field theory. This is distinguished from those integrals we have considered earlier in that the objects summed over depend on more than one continuous parameter. Thus in the path integral we added contributions from “all” functions $x(t)$, $0 \leq t \leq T$ with some boundary conditions at $t=0$ and T . In field theory one wishes to sum over “all” fields $\varphi(\mathbf{r}, t)$ with both \mathbf{r} and t varying. Similarly the statistical mechanics of one-dimensional systems can often be cast as a Wiener integral while in higher dimensions functional integration with several independent continuous parameters is needed.

We first lay out what one would like to have—formally speaking—in a path integral for fields, then indicate what some of the problems are, and finally mention some of the achievements of this approach.

Straightforward generalization from quantum mechanics suggests that if one has a classical field $\varphi(\mathbf{r})$ then the basic quantum object should be the wave function $\Psi(\varphi(\mathbf{r}), t)$, the probability amplitude that the field has assumed the value $\varphi(\mathbf{r})$ at time t . Suppose that the dynamics of the classical

field are gotten from a variational principle for an action S given by

$$S = \int dt \int d^3x \mathcal{L}(\varphi, \dot{\varphi}) \quad (32.1)$$

$$\mathcal{L} = \frac{1}{2} \left(\frac{\partial \varphi}{\partial t} \right)^2 - \frac{1}{2} (\nabla \varphi)^2 - \frac{1}{2} m^2 \varphi^2 - P(\varphi) \quad (32.2)$$

where m is a constant and P some function of φ , such as $\lambda \varphi^4$. Then the classical field $\varphi(\mathbf{r}, t)$ obeys the Euler-Lagrange equation

$$\frac{\partial^2 \varphi}{\partial t^2} - \nabla^2 \varphi + m^2 \varphi + \frac{\partial P}{\partial \varphi} = 0 \quad (32.3)$$

while the quantum state evolves with the propagator

$$\begin{aligned} G[\varphi''(\cdot), t''; \varphi'(\cdot), t'] \\ = \int \prod_{\mathbf{r}\tau} d\varphi(\mathbf{r}, \tau) \exp \left[\frac{i}{\hbar} \int_0^t d\tau \int d^3x \mathcal{L}(\varphi, \dot{\varphi}) \right] \end{aligned} \quad (32.4)$$

The symbol $\prod_{\mathbf{r}, \tau}$ indicates a sum over field configurations. The sum is over fields that at time t' have configuration $\varphi'(\mathbf{r})$ and at t'' configuration $\varphi''(\mathbf{r})$. G propagates $\Psi(\varphi, t)$ in such a way that it satisfies a linear differential equation. This equation, the analogue of the Schrödinger equation, involves functional derivatives $\delta^2/\delta\varphi(\mathbf{r}_1)\delta\varphi(\mathbf{r}_2)$. Note that relativistic invariance would not be hard to come by if we would talk about spacelike surfaces, but we shall not bother about that here.

Starting from G one can—formally—do everything: wave equations, commutation relations, vacuum expectation values; derive the perturbation expansion; even apply the method of steepest descents. We do not give a systematic presentation of this material (references will be given shortly), but just to show the usefulness and a weakness of working with the formal object we derive the Landau theory of phase transitions. The starting point is the partition function for a system that has order parameter $\varphi(\mathbf{r})$ and for a given configuration (i.e., spatial dependence of the order parameter) it has energy

$$H = \int_{\Omega} d^3x [(\nabla \varphi)^2 + \epsilon \varphi^2 + \varphi^4] \quad (32.5)$$

where Ω is the volume of the system. Then Landau theory says that the

free energy of the system is H evaluated at that φ which minimizes it. For $\epsilon < 0$ the system goes into states of nonzero φ and $\epsilon = 0$ is the critical point. On the other hand, the more accurate statistical mechanical treatment takes into account fluctuations in φ and gives as the partition function the functional integral

$$Z = \int d\varphi(\cdot) \exp(-\beta H) \quad (32.6)$$

The notation here is slightly different from that in (32.4) and the dot in φ is for the dummy three-vector label (\mathbf{r}). The quantity β is inverse temperature and the sum is over all field configurations. (This object was studied in Section 30.) The free energy per unit volume is

$$F = \left(-\frac{1}{\beta \Omega} \right) \log Z \quad (32.7)$$

in the limit $\Omega \rightarrow \infty$. Now for large Ω , H is large, and one might suppose that Laplace's method could be used for (32.6), that is, Z is mainly given by the contribution of that φ which minimizes H . But then $Z \sim \exp(-\beta H(\varphi_{\min}))$ so that $F \sim (1/\Omega) H(\varphi_{\min})$, which is just Landau theory. From Section 30 we know that Z is not reasonably approximated by Laplace's method near the critical point and that the whole drama of renormalization and scaling is played out for just this reason.

Let us consider why the functional integrals in (32.4) or (32.6) are fundamentally more troublesome than the path integral or Wiener integral. For the v dimensional case we have

$$G = \int d\varphi(\cdot) \exp \left\{ - \int d^v x \left[\frac{1}{2} \sum_{i=1}^v \left(\frac{\partial \varphi}{\partial x_i} \right)^2 + \frac{1}{2} m^2 \varphi^2 + P(\varphi) \right] \right\} \quad (32.8)$$

One can obtain (32.8) from (32.4) by "Euclideanizing," that is, by going to imaginary time. For field theory on physical space time $v=4$; for physical space, for the computation of the partition function $v=3$; for ordinary quantum mechanics or the Wiener integral, $v=1$. Our notation does not specify what sort of boundary conditions, if any, are to be used for G . Now Fourier transform, so that the functional integral is over the Fourier components of φ :

$$G = \int \prod_k d\tilde{\varphi}(k) \exp \left\{ - \int d^v k \left[\frac{1}{2} (k^2 + m^2) \tilde{\varphi}(k) \tilde{\varphi}(-k) \right] - Q(\tilde{\varphi}(k)) \right\} \quad (32.9)$$

where Q is a functional of $\tilde{\varphi}(k)$ given by

$$Q(\tilde{\varphi}(k)) = \int d^v x P \left(\int e^{ikx} \tilde{\varphi}(k) d^v k \right) \quad (32.10)$$

The k^2 in (32.9) is $\sum_{i=1}^v k_i^2$ and is of course just the kinetic energy (or in statistical mechanics the nearest neighbor coupling).

Using (32.8) and (32.9) let us review some facts about the Wiener integral, the case $v=1$. For the Wiener integral (almost all) the paths summed over are continuous but nowhere differentiable. The latter feature has been repeatedly emphasized and arose from the appearance in the path integral of terms

$$\exp \left[\frac{i}{\hbar} \frac{(x_{i+1} - x_i)^2}{t/N} \right] = \exp \left(\frac{i}{\hbar} \frac{(\Delta x)^2}{\Delta t} \right) \quad (32.11)$$

When doing the path integral for given Δt the important range of x was such that

$$\frac{(\Delta x)^2}{\Delta t} = O(1) \quad (32.12)$$

The paths contributing to the sum are thus characterized by

$$\text{derivative} \sim \frac{\Delta x}{\Delta t} = O\left(\frac{1}{\sqrt{\Delta t}}\right) \quad (32.13)$$

which becomes infinite as $\Delta t \rightarrow 0$. (Actually $(\Delta x)^2/\Delta t$ grows logarithmically as we saw in our earlier discussions of asymptotic approximations, but for power counting purposes we can ignore this.)

Let us use the same sort of reasoning on (32.9) to find the large k dependence of $\tilde{\varphi}(k)$ (of course, large k is where infinite derivative or discontinuity problems would be manifested). From (32.9), for each mode k we expect

$$|\tilde{\varphi}(k)| = O\left(\frac{1}{k}\right) \quad (32.14)$$

Now we know from other sources that for $v=1$ the paths are continuous, and it would be nice if hand waving arguments based on (32.14) could be used to establish this fact. Unfortunately (32.14) is not quite enough to establish continuity, but it is nearly good enough and will serve our purpose for the heuristic points we wish to make for $v > 1$.

Relating asymptotic properties of Fourier coefficients, for example, the relation (32.14), to properties of the function transformed is one of the great subsciences of Fourier analysis. One such theorem, for example, is

THEOREM: Suppose that $f(x)$ satisfies a Lipschitz condition of order α for some $\alpha, 0 < \alpha \leq 1$, that is, as $h \rightarrow 0$

$$f(x+h) - f(x) = O(|h|^\alpha)$$

uniformly with respect to x . Then

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx dx = O(n^{-\alpha})$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx dx = O(n^{-\alpha}) \quad \text{for } n \rightarrow \infty$$

Another such result is that for functions of bounded variation both a_n and b_n are $O(1/n)$.

The property that we wish to use is that if for some $\alpha > 0$ $\tilde{\varphi}(k) = O(k^{-1-\alpha})$, then $\varphi(x)$ is continuous. This is verified by computing

$$\varphi(x+\delta) - \varphi(x) = \int dk [\tilde{\varphi}(k) k^{1+\alpha}] \frac{1}{k^\alpha} \left[\frac{e^{ik\delta} - 1}{k} \right] e^{ikx} \quad (32.15)$$

It is easy to show that the difference in (32.15) can be made arbitrarily small with δ provided that $\tilde{\varphi}(k)$ is assumed well behaved near zero (e.g., by assuming $\int_0^{k'} k |\tilde{\varphi}(k)| dk$ to be finite for some k'). For the functions of (32.14) this means that a little bit of smearing would make them continuous (recall that multiplying by k corresponds to taking derivatives, dividing to integrating). Actually we know Wiener paths to be continuous but our object now is to see how much can be obtained from inspection of Fourier coefficients.

Now consider the functional integral for $\nu > 1$. The exponent in (32.9) still has $k^2 |\tilde{\varphi}^2(k)|^2$ so that

$$|\tilde{\varphi}(k)| = O\left(\frac{1}{(\sum k_i^2)^{1/2}}\right) \quad (32.16)$$

For two dimensions the fields then look like

$$\begin{aligned} \varphi(x_1, x_2) &= \int dk_1 dk_2 e^{i(k_1 x_1 + k_2 x_2)} \tilde{\varphi}(k_1, k_2) \\ &= \int dk_1 e^{i k_1 x_1} \psi(k_1, x_2) \end{aligned}$$

which defines $\psi(k_1, x_2)$. Now (32.16) tells us that $\psi(k_1, x_2)$ as a function of x_2 is close to being continuous, but having integrated once the relation (32.16) no longer helps us with $\varphi(x_1, x_2)$. If we had had $\tilde{\varphi}(k) = O(\Pi k^{-1})$ then we would (nearly) have continuity; but alas φ does not have this property. An integral of φ would, however, have in its Fourier transform an additional power of k in the denominator. Consequently the integral of φ is “nearly” continuous; stated differently if we smear (i.e., integrate or average) φ $1+\alpha$ times for any $\alpha > 0$ it will be continuous. In ν dimensions $\nu - 1 + \alpha$ integrations are necessary.

It follows that the fields entering the functional integral are very rough, much worse than Brownian motion paths, with short wavelength problems getting more severe with increasing dimension. This is actually a reflection of ultraviolet divergences in the usual formulations of field theory. There when one evaluates the free Green's function $(\sum_i \partial^2 / \partial x_i^2 - m^2)^{-1}$ the integral

$$\int \frac{d^\nu k}{k^2 + m^2} \quad (32.17)$$

appears and this diverges with increasing ferocity for $\nu \geq 2$.

The expression (32.17) indicates that for $\nu > 1$ there are too many modes for the $1/k^2$ dropoff to compensate. This also suggests that even if one were brave enough to work with these rough fields, Laplace's method, an approximation that makes the functional integral so useful, would be in trouble. In Section 14 we showed how Laplace's method could be extended to infinite dimensions—that is, to the Wiener integral—by explicitly estimating the error in the integration over each mode and showing that the sum goes to zero with \hbar . Now there are far more modes but the dropoff is no more rapid; hence the usefulness of Laplace's method seems doubtful.

On the other hand, while the roughness of the fields certainly suggests difficulties for $\nu \geq 2$ it does not necessarily mean that all is lost, and one might expect cures to be available as they are for ultraviolet divergences in the usual treatments. In this book though we have mostly stayed away from such problems. To close our discussion of field theory, however, we briefly indicate how in fact Feynman managed to use path integrals to great advantage for quantum electrodynamics despite all the horrors that we have been at great pains to describe.

Quantum electrodynamics is the theory of matter in interaction with the electromagnetic field. The usefulness of path integrals arises mainly from the fact that the electromagnetic field can be represented as an infinite collection of harmonic oscillators whose interaction with matter is

expressed as a linear forcing term for the oscillators. The path integral for this system is known in closed form (Section 6), and so one can integrate out the field coordinates completely. Exactly this procedure was used to eliminate the phonon field in the polaron problem.

Let the matter field be treated nonrelativistically. For each momentum the electromagnetic field has two degrees of freedom (transverse polarizations), and since the matter field is nonrelativistic anyway, it is convenient to work in a noncovariant gauge in which ϕ (4th component of the vector potential for the electromagnetic field) is explicitly solved for in terms of the matter and thereby eliminated as a dynamical variable. The path integral theory can be formulated when we have a classical action for a specific history of the field plus matter system. This action is

$$S = S_{\text{mat}} + S_{\text{rad}} + S_{\text{int}} \quad (\text{subscripts: "matter, radiation, interaction"}) \quad (32.18)$$

$$S_{\text{mat}} = \int dt \left[\sum_{i=1}^n \frac{1}{2} m_i (\dot{\mathbf{q}}_i)^2 + \sum_{i,j} \frac{e_i e_j}{|\mathbf{q}_i(t) - \mathbf{q}_j(t)|} \right] \quad (32.19)$$

$$S_{\text{rad}} = \frac{1}{2} \int dt \int \frac{d^3 k}{(2\pi)^3} [\dot{a}_{1k}^* \dot{a}_{1k} - k^2 a_{1k}^* a_{1k} + (1 \leftrightarrow 2)] \quad (32.20)$$

$$S_{\text{int}} = \sqrt{4\pi} \sum_j e_j \int dt \int \frac{d^3 k}{(2\pi)^3} (a_{1k} \dot{q}_{1j} + a_{2k} \dot{q}_{2j}) \exp(i \mathbf{k} \cdot \mathbf{q}_j) \quad (32.21)$$

where particles are labeled by i, j , have masses m_i and charges e_i . The variables a_{1k} and a_{2k} are Fourier components of the (3-) vector potential for the electromagnetic field and \mathbf{k} is the wave vector of the mode. We are using the gauge $\nabla \cdot \mathbf{A} = 0$, so that a_{lk} , $l=1,2$ are amplitudes for the two polarizations perpendicular to \mathbf{k} . The term $\sum e_i e_j / |\mathbf{q}_i - \mathbf{q}_j|$ arises from the solved fourth component and appears as an instantaneous Coulomb interaction. Considering S_{rad} by itself would lead to a description of the field as a collection of free oscillators, each excitation of such an oscillator a photon of the appropriate momentum and polarization. S_{int} is essentially $e \mathbf{v} \cdot \mathbf{A}$ —a term familiar from other work in which the electromagnetic field is considered to be a given external field and not dynamical. The difference in the present treatment is that we do a path integral over the variables a_{lk} too. The quantity \dot{q}_{lj} ($l=1,2$) is the velocity of particle j along the polarization direction of a_{lk} . From the action given in (32.18) to (32.21) by means of the usual variational principle there follow the classical equations for charged matter in interaction with the electromagnetic field.

In principle the object of interest is the propagator for going from one field-matter configuration to another, as written in (32.4). However, the

electromagnetic field is more commonly specified by indicating the number of photons present initially and finally, and the desired physical quantity is a matrix element of G . For example, the Lamb shift is a radiative correction to hydrogen atom levels so that one would look at zero photon to zero photon (all k , both polarizations) transitions and at the change in the specified electron energy level. Photon emission amplitudes are obtained by letting the initial state have no photon and (say) starting a hydrogen atom in its $2p$ state. The final state has one photon and a hydrogen ground state.

Let us consider an amplitude for the electromagnetic field to start and end in vacuum and the matter to go from some set of values q'_i to values $q''_i \quad i=1,\dots,n$:

$$\begin{aligned} G(q'', \text{vac}, t; q', \text{vac}) = & \int \prod_{l,\mathbf{k}} da''_{l\mathbf{k}} da'_{l\mathbf{k}} [\text{ground state harmonic} \\ & \text{oscillator wave functions for all } a'_{l\mathbf{k}} \text{ and } a''_{l\mathbf{k}}] \\ & \times \int \prod_{l,\mathbf{k}} da_{l\mathbf{k}}(\cdot) \prod_i dq_i(\cdot) \exp \left\{ \frac{i}{\hbar} [S_{\text{mat}} + S_{\text{rad}} + S_{\text{int}}] \right\} \end{aligned} \quad (32.22)$$

Looking at the forms (32.20) and (32.21) we note that both the path integral and the ordinary integral over each $a_{l\mathbf{k}}$ can be done separately and, suppressing the index $l\mathbf{k}$ on the a 's, each of these becomes

$$\begin{aligned} \exp \left(\frac{i}{\hbar} I_{l\mathbf{k}} \right) = & \int da' da'' \left(\frac{k}{4\pi\hbar} \right)^{1/2} \exp \left[-\frac{k}{2\hbar} (a''^2 + a'^2) \right] \int da(\cdot) \\ & \times \exp \left\{ \frac{i}{\hbar} \int \frac{1}{2} (\dot{a}^* \dot{a} - k^2 a^* a) dt \right. \\ & \left. + \frac{i}{\hbar} \sqrt{4\pi} \int dt a \sum_j \dot{q}_j(t) \exp(i\mathbf{k} \cdot \mathbf{q}_j(t)) \right\} \end{aligned} \quad (32.23)$$

When (32.23) is evaluated it will be inserted in the amplitude G and finally the path integral over the matter coordinates q_j will be done. Consequently in (32.23) each $\mathbf{q}_j(t)$ is a fixed function of time (later to be integrated over) and (32.23) is the path integral for the forced harmonic oscillator. This is the key trick of this approach and it is what makes the polaron solution work too (Section 21). For the solution to (32.23) we refer the reader to

Section 6. The result is that for each mode I_{lk} is essentially

$$I_{lk} = \frac{i}{2} \int \sum_{i,j} e_i e_j \dot{q}_i(t) \dot{q}_j(s) \frac{4\pi}{2k} \exp(-ik|t-s|) dt ds \quad (32.24)$$

The path integral over q_j is therefore

$$G = \int \prod_i dq_i(\cdot) \exp \left[\frac{i}{\hbar} (S_{\text{mat}} + I) \right] \quad (32.25)$$

with

$$I = \sum_{l=1,2} \sum_{\mathbf{k}} I_{lk} \quad (32.26)$$

At the risk of providing fuel for scoffers, I have to confess that the result (32.25) has always seemed to me to be deep and full of not well understood implications. Equation 32.25 says that the electromagnetic field is a fiction. Particles interact with an action given by $S_{\text{mat}} + I$, both terms of which depend only on matter coordinates. Mention need never be made of the electromagnetic field. Of course we could introduce auxiliary coordinates and rewrite the quantity I as a functional integral over them—as in (32.23)—but then one would be hard put to assign any fundamental significance to the auxiliary coordinates. All the physics can be obtained from (32.25) with never a mention of X-rays, gamma rays, or even glowing sunsets.

One can take these strange views a step further and ask about the classical limit of (32.25). This should give a theory of matter in which the matter exhibits a peculiar self interaction that precludes the need for any mention of a mediating electromagnetic field. In fact, there is such a theory due to Fokker, Schwarzschild, Tetrode, Wheeler, and Feynman, which follows from a variational principle based on the following action for the particles:

$$S = - \sum_{i=1}^n \int m_i \sqrt{dx_i^\mu dx_{i\mu}} - \sum_{i < j} e_i e_j \int \int \delta[(x_i(\tau_i) - x_j(\tau_j))^2] dx_i^\mu dx_{j\mu} \quad (32.27)$$

where x_i^μ are the (4-vector) coordinates of the i th particle, τ_i is a proper time parametrization of the particle trajectory, a summation convention for μ is adopted, $c=1$ and all is relativistic. Note that the interaction is nonlocal and in fact symmetric in time. From the action (32.27) one can

recover the usual field and matter equations by defining auxiliary quantities

$$A^\mu(x) = \sum_j e_j \int \delta((x - x_j(\tau))^2) dx_j^\mu \quad (32.28)$$

which behave exactly like the vector potential at x .

Not surprisingly, the effective action from (32.25) is essentially the same as (32.27). There is one difference, however, which I don't think anyone really understands—it is the real part of $(S_{\text{mat}} + I)$ that gives (32.27). The imaginary part is a bit of a puzzle, but cannot be just dropped, since it enters calculations in quantum electrodynamics.

Having put forth this strange view of the electromagnetic field, let me add a few comments to detract from its mystic significance. First, (32.25) as it stands does not say everything there is to say about transition amplitudes. It is a vacuum-vacuum transition amplitude and the quantity I would be changed slightly by the presence of excitations in the initial or final states. However, by lengthening the time interval and putting in extra particles to pick up outgoing waves, general amplitudes could in principle be reduced to vacuum-vacuum amplitudes. A more cogent criticism is that we could have done the path integral on the matter first, ended up with an amplitude between electromagnetic field configurations and declared all matter to be fictional. The mathematics—lacking an explicit solution analogous to (32.23) and (32.24)—would be more cumbersome, but the principle is the same.

Starting from (32.25), which is an ordinary path integral, one can use all the usual techniques to do the quantum mechanics of particles interacting through a quantized electromagnetic field. Perturbation theory consists of expanding $\exp(iI/\hbar)$ in a series, and now the diagrams of this series—as developed in Section 10—really are Feynman diagrams in the form originally propounded.

Notes

We do not present the way in which physical quantities of interest are calculated from (32.25), as all this is given in detail in the book

R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965

where will be found derivations of the equations we have used. See also

R. P. Feynman, Mathematical Formulation of the Quantum Theory of Electromagnetic Interactions, *Phys. Rev.* **80**, 440 (1950)

G. C. Dente, Classical Limit of Quantum Electrodynamics, *Phys. Rev. D* **13**, 1733 (1975)

The classical theory of action at a distance electrodynamics, (32.27), is given in

J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.* **17**, 159 (1945); **21**, 425 (1949)

A formal presentation of field theory through the use of the path integral appears in

G. Rosen, *Formulations of Classical and Quantum Dynamical Theory*, Academic Press, New York, 1969

In rigorous work on quantum field theory the functional integral—despite the roughness of the fields—has proved quite useful. A review of some of the developments in this area can be found in

J. Glimm and A. Jaffe, Functional Integral Methods in Quantum Field Theory; J. Glimm and A. Jaffe, A Tutorial Course in Constructive Field Theory, in M. Levy and P. Mitter, Eds., *New Developments in Quantum Mechanics and Statistical Mechanics; Cargese 1976*, Plenum, New York, 1977.

32.2 UNCOMPLETING THE SQUARE

Consider the identity

$$\exp(b^2) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx \exp(-x^2 + 2xb) \quad (32.29)$$

The proof is trivial and consists of writing

$$x^2 - 2xb = (x - b)^2 - b^2$$

For all its triviality, however, it is a powerful way of getting rid of squares (“ b^2 ”) in an exponent and replacing them with linear terms (“ $2xb$ ”) at the expense of introducing auxiliary variables. The relevance of this device to functional integration is that often one wishes to be rid of an infinity of square terms and the resulting integrals over the infinity of auxiliary variables takes the form of a functional integral.

As a quick aside and as a preview to later applications, consider the following two examples that do not involve functional integration.

The first example is the Curie-Weiss model of a ferromagnet. A collection of N spins, σ_i , $i = 1, \dots, N$, each spin taking the values ± 1 , interact with long range forces, so that the energy of a given spin configuration is

$$E = E(\sigma_1, \dots, \sigma_N) = -\frac{J}{N} \sum_{i=2}^N \sum_{j=1}^{i-1} \sigma_i \sigma_j - h \sum_{i=1}^N \sigma_i \quad (32.30)$$

where J is a coupling constant and h an external magnetic field. The partition function and free energy are

$$Z_N = \sum_{\{\sigma\}} \exp(-\beta E(\sigma_1, \dots, \sigma_N)); \quad F_N = -\frac{1}{\beta N} \log Z_N \quad (32.31)$$

the sum being over all configurations and β is inverse temperature. Using the fact that $\sigma_i^2 = 1$ and defining $M = \sum_i \sigma_i$ Z_N becomes

$$Z_N = \sum_{\{\sigma\}} \exp \left\{ -\frac{\beta J}{2} + \frac{\beta J}{2N} M^2 + \beta h M \right\}$$

Now apply the identity (32.29) and interchange order of summation and integration to yield

$$Z_N = e^{-\beta J/2} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx \exp(-x^2) \sum_{\{\sigma\}} \exp \left[M \left(2x \sqrt{\frac{\beta J}{2N}} + \beta h \right) \right] \quad (32.32)$$

Ordinarily the configuration sum is the most difficult part, but in (32.32) the fact that M appears only linearly allows the sum over each σ_k , $k = 1, \dots, N$ to be done separately (the spins act like independent particles in a collective external field $h + 2x \sqrt{J/2\beta N}$). The configuration sum yields

$$Z_N = e^{-\beta J/2} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx \exp(-x^2) \left[2 \cosh \left(2x \sqrt{\frac{\beta J}{2N}} + \beta h \right) \right]^N \quad (32.33)$$

Laplace's method for the integral (change variables to $\eta = x \sqrt{2/\beta J N}$) gives for $N \rightarrow \infty$

$$F = \lim_{N \rightarrow \infty} F_N = \min_{\eta} \left[\frac{1}{2} J \eta^2 - \frac{1}{\beta} \log \cosh(\beta J \eta + \beta h) \right] \quad (32.34)$$

The value of η at which the minimum is assumed satisfies

$$\eta = \tanh(\beta J \eta + \beta h) \quad (32.35)$$

which is the well known equation for the mean field magnetization. (This is

no coincidence since the magnetization is $\partial F/\partial h$. Writing $F = \min \psi(\eta, h)$ it is easy to see that $\partial F/\partial h = \partial \psi/\partial h$ and that $\partial \psi/\partial h$ is just η .) For $h=0$ there is a first-order phase transition for $\beta J > 1$, and a second-order transition at $\beta J = 1$.

A second example of the use of the identity (32.29) involves analytic continuation. Consider the integral

$$F(\lambda) = \int_{-\infty}^{\infty} dx \exp(-x^2 - \lambda x^4) \quad (32.36)$$

For $\operatorname{Re} \lambda > 0$ this is obviously analytic while for $\operatorname{Re} \lambda < 0$ the integral is just as obviously bad. Uncomplete the square for the term $\exp(-\lambda x^4)$ and interchange the order of integration (justified for $\operatorname{Re} \lambda > 0$) to yield

$$F(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dy \exp(-y^2) \int_{-\infty}^{\infty} dx \exp(-x^2 + 2i\sqrt{\lambda} x^2 y) \quad (32.37)$$

The x integral can be done explicitly, so that

$$F(\lambda) = \int_{-\infty}^{\infty} dy \frac{\exp(-y^2)}{\sqrt{1 - 2iy\sqrt{\lambda}}} \quad (32.38)$$

This integral is healthy everywhere except for λ real and negative and uncompleting the square has provided an easy analytic continuation. (See Sections 13 and 29 for other methods of studying the analytic continuation of this integral.)

For many-dimensional uncompletions of the square (which lead to auxiliary functional integrals) it is convenient to use the more general formula

$$\begin{aligned} \exp\left[\frac{1}{2} \sum_{i,j=1}^N x_i A_{ij} x_j\right] &= (2\pi)^{-N/2} [\det A]^{-1/2} \int d^N u \\ &\times \exp\left\{-\frac{1}{2} \sum_{i,j=1}^N u_i (A^{-1})_{ij} u_j + \sum_{i=1}^N u_i x_i\right\} \end{aligned} \quad (32.39)$$

[see (6.24) for proof]. Again consider a spin model, but now not so trivial as that which we solved earlier. Let there be N spins on a d dimensional lattice. Let the energy of a configuration $\{\sigma_i\}$ be given by

$$E = -\frac{1}{2} \sum_{i,j} \sigma_i J_{ij} \sigma_j - h \sum_i \sigma_i \quad (32.40)$$

where $i, j \in \mathbb{Z}^d$. The nearest neighbor Ising model corresponds to J_{ij} = (positive const.) if $\|i-j\|$ = [Euclidean distance between i and j] = 1 and zero otherwise, but we do not restrict ourselves to this case. Again the partition function is the configuration sum of $\exp(-\beta E)$, and if we uncomplete the square by means of (32.39) we get

$$\begin{aligned} Z_N = & (2\pi\beta)^{-N/2} (\det J)^{-1/2} \int d\phi_1 \cdots d\phi_N \exp \left[-\frac{1}{2\beta} \sum_{i,j} \phi_i (J^{-1})_{ij} \phi_j \right] \\ & \times \sum_{\{\sigma\}} \exp \left[\sum_i \sigma_i (\phi_i + \beta h) \right] \end{aligned} \quad (32.41)$$

As before the configuration sum is trivial and

$$\begin{aligned} Z_N = & \left(\frac{2}{\pi\beta} \right)^{N/2} (\det J)^{-1/2} \int d\phi_1 \cdots d\phi_N \\ & \times \exp \left\{ -\frac{1}{2\beta} \sum_{i,j} \phi_i (J^{-1})_{ij} \phi_j + \sum_i \log \cosh(\phi_i + \beta h) \right\} \end{aligned} \quad (32.42)$$

For some matrices J_{ij} this is just the familiar functional integral. In particular J^{-1} can in some approximations look like the Jacobi matrix and for appropriate limiting behavior as $N \rightarrow \infty$ the term $\sum \phi_i J^{-1} \phi$ becomes $(\text{const})(\nabla\phi)^2 + (\text{const})\phi^2$. In this way discrete spin models become functional integrals.

Exercise: Let J_{ij} be translationally invariant (a function of $i-j$ or $r_i - r_j$ only). Under what circumstances is the limiting behavior as described above and what are the constants in terms of J_{ij} ?

There is a special class of J_{ij} 's whose inverses are particularly simple. For convenience we take $d=1$ although getting a phase transition in one dimension is a bit of a balancing act. Let

$$J_{ij} = g\gamma v(\gamma|i-j|) \quad (32.43)$$

with

$$v(r) = e^{-r} \quad (32.44)$$

Then with various changes of variables

$$\begin{aligned} Z_N = & e^{-g\beta N \gamma/2} \sum_{\{\sigma\}} \int d\phi_1 \cdots d\phi_N \exp \left[\sqrt{\beta J \gamma} \sum_{k=1}^N \phi_k \sigma_k + h\beta \sum_{k=1}^N \sigma_k \right] \\ & \times W(\phi_1, \dots, \phi_N) \end{aligned} \quad (32.45)$$

with

$$W(\phi_1, \dots, \phi_N) = w(\phi_1) P_\gamma(\phi_1, \phi_2) P_\gamma(\phi_2, \phi_3) \cdots P_\gamma(\phi_{N-1}, \phi_N)$$

$$w(\phi) = (2\pi)^{-1/2} \exp\left(-\frac{\phi^2}{2}\right)$$

$$P_\gamma(\phi, \phi') = [2\pi(1 - e^{-2\gamma})]^{-1/2} \exp\left\{\frac{-(\phi' - \phi e^{-\gamma})^2}{2(1 - e^{-2\gamma})}\right\}$$

The quantity $W(\phi_1, \dots, \phi_N)$ is the term involving “ J^{-1} ” and what is special about the form (32.44) is that W factors into products involving ϕ 's taken two at a time. The configuration sum is as usual trivial and the factorization of W allows Z_N to become the $(N-1)$ th power of a certain linear operator.

Thus after some algebra we have

$$Z_N = e^{-\beta g N \gamma / 2} 2^N \int d\phi d\phi' h(\phi) K^{N-1}(\phi, \phi') h(\phi') \quad (32.47)$$

where

$$h(\phi) = \left[w(\phi) \cosh(\sqrt{\beta g \gamma} \phi + \beta h) \right]^{1/2} \quad (32.48)$$

and K^{N-1} is the $(N-1)$ th power of the kernel

$$K(\phi, \phi') =$$

$$\left[\cosh(\sqrt{\beta g \gamma} \phi + \beta h) \right]^{1/2} \sqrt{\frac{w(\phi)}{w(\phi')}} P_\gamma(\phi, \phi') \left[\cosh(\sqrt{\beta g \gamma} \phi' + \beta h) \right]^{1/2} \quad (32.49)$$

This immediately gives the free energy as the logarithm of the largest eigenvalue of this kernel. See the references below for more details.

A more interesting version of the foregoing development concerns one-dimensional systems ($d=1$) for which as $r=|i-j|\rightarrow\infty$

$$v(r) \sim \frac{1}{r^\alpha} \quad 1 < \alpha \leq 2 \quad (32.50)$$

Such a medium range interaction has a phase transition even in one dimension (questions still exist for $\alpha=2$).

On the face of it (32.50) would seem to be a different problem from the potential (32.44) but Kac and Thompson take

$$v(r) = \int_0^1 d\lambda \lambda^{\alpha-1} e^{-\lambda r} \quad (32.51)$$

which has the right asymptotic properties. Then v is approximated by a discrete sum of m exponentials

$$v(r) \approx \sum_{k=1}^m \frac{1}{m} \left(\frac{k}{m} \right)^{\alpha-1} e^{-kr/m} \quad (32.52)$$

For any finite m the inverse of J can be found and Z is obtained from the largest eigenvalue of the kernel of an operator on $L^2(\mathbf{R}^m)$ [instead of \mathbf{R}^1 , as in (32.49)].

It turns out that the eigenvalues of K can be approximated by solving a certain Schrödinger equation in m dimensions and that a good way to find the eigenvalue is by introducing yet another functional integral in order to use the Feynman-Kac formula. Then this eigenvalue can be analytically continued by instanton techniques on the functional integral, as was done in Section 29. Details will be found in the references cited below.

Uncompleting the square, together with Laplace's method is also a way of generating various mean field theories. Consider the Lagrangian

$$\mathcal{L}(\phi) = \frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{2} m_0 \phi^2 - \frac{1}{4} \lambda_0 \phi^4 \quad (32.53)$$

for the relativistic field ϕ having ϕ^4 self interaction. Amplitudes of all sorts for the quantum field theory are obtained from the functional integral

$$G = \int d\phi(\cdot) \exp \left(- \int d^d x \mathcal{L}(\phi) \right) \quad (32.54)$$

The ϕ^4 term is now "uncompleted" by introducing an auxiliary field ψ so that

$$G = \int d\psi(\cdot) d\phi(\cdot) \exp \left(- \int d^d x \mathcal{L}(\phi, \psi) \right) \quad (32.55)$$

with

$$\mathcal{L}(\phi, \psi) = \frac{1}{2} (\partial_\mu \phi)^2 - \frac{1}{2} m_0 \phi^2 - \frac{1}{2} \lambda_0 \psi \phi^2 + \frac{1}{4} \lambda_0 \psi^2 \quad (32.56)$$

Now ϕ can be integrated explicitly [as for the one-dimensional integral (32.37) above] and we are left with an intractable integral in ψ . When Laplace's method is applied to the ψ integral a mean field theory together with a perturbation theory around that mean field emerges.

Yet another uncompletion of the square has proved useful even when the Hamiltonian contains operators that do not commute with one another.

Suppose

$$H = H_0 + AB \quad (32.57)$$

where

$$[A, B] = 0 \quad \text{but} \quad [H_0, A] \neq 0, \quad [H_0, B] \neq 0$$

We wish to “linearize” the product AB and to this end make use of the formula

$$AB = \frac{1}{4}(A+B)^2 - \frac{1}{4}(A-B)^2$$

and introduce two real (or one complex) auxiliary integrations. Because of the noncommutativity of H_0 with A and B it is not enough to introduce just two auxiliary variables but in fact a time ordered product must be used and the final form can be written

$$\begin{aligned} \exp(-H) &= T \int dz(\cdot) \exp \left[\pi \int_0^1 d\tau |z(\tau)|^2 \right] \\ &\times \exp \left\{ - \int_0^1 d\tau \left[H_0 + \sqrt{\pi} Az(\tau) - \sqrt{\pi} Bz^*(\tau) \right] \right\} \end{aligned} \quad (32.58)$$

where $z(\tau)$ is a complex function for $0 \leq \tau \leq 1$. As indicated by the symbol “ T ” the integral is a time (i.e., τ) ordered product like the product integrals of Section 22.1. The form (32.58) has motivated a number of approximations in, for example, the Anderson model (see references), although I find some of the steepest descent approximations on the time ordered products a bit much for my mathematical sensitivities.

Notes

Some of the material presented here on the uncompleting of the square can be found in

M. Kac, Mathematical Mechanisms of Phase Transitions, in M. Chretien, E. P. Gross, and S. Deser, Eds., *Statistical Physics: Phase Transitions and Superfluidity*, Vol. 1, Proc. Brandeis Univ. Summer Institute, Gordon and Breach, New York, 1968

where further references will be found. Analytic continuation of the free energy for the $r^{-\alpha}$ potential using instanton techniques has been carried out by

R. J. McCraw, Metastability in a Long Range One-Dimensional Ising Model, *Phys. Lett.* **75A**, 379 (1980)

and in more detail in an Indiana University thesis. Mean field theory of the auxiliary variable is studied by

C. M. Bender, F. Cooper, and G. S. Guralnik, *Ann. Phys.* **109**, 165 (1977)

A brief survey of applications of uncompletions of the square for noncommuting operators is given by

B. Mühlriegel, Functional Integrals and Local Many-Body Problems: Localized Moments and Small Particles, in *Functional Integration and its Applications*, A. M. Arthurs, Ed. (1974 Conference), Clarendon Press, Oxford, 1975

where other references will be found.

32.3 RUBBER: FUNCTIONAL INTEGRAL FORMULATION OF A POLYMER AS A RANDOM WALK

For some purposes a polymer molecule can be considered to be a random walk of step size l where l is the length of the individual monomers. A further refinement is to require that the random walk be self avoiding but this non-Markoffian property is a considerable complication and we ignore it here. A random walk polymer model exhibits elastic forces because of entropy effects alone. To see this consider a polymer molecule with fixed ends confined to a plane. Imagine putting your finger at a point R in the plane. This allows the paths to be classified with respect to their winding number (see Section 23) around the point R . The probability $p_n(R)$ that a path has winding number n is a function of R and is such that the entropy of an ensemble of such molecules is maximized with respect to other possible winding number probability distributions. Now move your finger. The rubber band does not slip, so p_n is unchanged, but your finger is at some other point R' where the maximum entropy is obtained by some other distribution $p_n(R')$ of winding numbers. Moving your finger thus lowers the entropy. Lowering the entropy raises the free energy so that the change in $p_n(R)$ provides a restoring force. In this way the elastic constants can be expressed in terms of $p_n(R)$ and $\partial p_n(R)/\partial R$. A nice way to characterize this entropy decrease is to describe it as arising from a topological constraint.

It remains to calculate $p_n(R)$. This is basically a Wiener integral, being the measure of a certain subset of the path space. It is convenient to calculate $p_\theta(R)$, the measure of the set of those paths that go through an angle θ around R , where θ is *not* measured modulo 2π , but includes whole rotations. This can be written

$$p_\theta(R) = (\text{const.}) \int dr(\cdot) \delta \left[\theta - \int_0^L \left(\frac{x\dot{y} - y\dot{x}}{x^2 + y^2} \right) d\tau \right] \exp \left[-\frac{1}{l} \int_0^L \dot{r}^2 d\tau \right] \quad (32.59)$$

with x and y coordinates in the plane and the integral of $(xy - y\dot{x})/(x^2 + y^2)$ being just the total angle traversed. The δ -function can be written $\delta(\theta - K) = (2\pi)^{-1} \int \exp[i\lambda(\theta - K)] d\lambda$ and the quantity of interest becomes the Fourier transform of

$$(\text{const.}) \int dr(\cdot) \exp \left\{ -\frac{1}{l} \int_0^L (\dot{r}^2 + i\mathbf{A} \cdot \mathbf{r}) d\tau \right\} \quad (32.60)$$

with

$$\mathbf{A} = \frac{\mathbf{r} \times \hat{\mathbf{z}}}{r^2} \quad (32.61)$$

In this way the elastic constants of rubber are computable from the path integral (or some analytic continuation thereof) for a particle in a certain magnetic field. The path integral can be carried out and expressions obtained for the elastic constants.

Notes

This work and further developments can be found in

S. F. Edwards, Statistical Mechanics with Topological Constraints, I, *Proc. Phys. Soc.* **91**, 513 (1967); II, *J. Phys. A (Proc. Phys. Soc.) Ser. 2*, **1**, 15 (1968)

S. F. Edwards, Functional Problems in the Theory of Polymers, in *Functional Integration and its Applications*, A. M. Arthurs, Ed. (1974 Conference), Clarendon Press, Oxford, 1975

32.4 HARD SPHERE GAS SECOND VIRIAL COEFFICIENT

The forthcoming calculation, due to Lieb, is noteworthy for its technical aspects. It comes closer to doing the multidimensional integral for the discretized path sum (i.e., the Trotter product formula) than most other methods—aside from those that somehow reduce to the harmonic oscillator—and as far as I know there is no previous or subsequent work at all similar to it.

The physical problem concerns a hard sphere gas and by standard statistical mechanics results the calculation of the quantum mechanical second virial coefficient reduces to a problem in ordinary one-particle quantum mechanics. Specifically, if the particle-particle pair interaction is $v(r)$ (for the hard sphere gas with atoms of radius $a/2$, $v(r)=0$ for $r>a$, $v(r)=\infty$ for $r<a$) and the mass of a single atom is m , then the second

virial coefficient can be obtained from the Green's function:

$$\left[-\frac{\hbar^2}{m} \nabla^2 + v(r) + \frac{\partial}{\partial t} \right] G(\mathbf{r}, t; \mathbf{r}') = 0 \quad t > 0$$

$$G(\mathbf{r}, t; \mathbf{r}') \rightarrow \delta(\mathbf{r} - \mathbf{r}') \quad \text{for } t \downarrow 0 \quad (32.62)$$

The second virial coefficient B is made up of two terms, direct and exchange, given respectively by

$$B_d = \frac{1}{2} N \int d^3 r \left[1 - 2\sqrt{2} \Lambda^3 G(\mathbf{r}, \beta; \mathbf{r}) \right] \quad (32.63a)$$

$$B_e = \mp \sqrt{2} \Lambda^3 N (2s+1)^{-1} \int d^3 r G(\mathbf{r}, \beta; -\mathbf{r}) \quad (32.63b)$$

and

$$B = B_d + B_e$$

In (32.63), N is Avogadro's number, β inverse temperature, s the atomic spin, and

$$\Lambda^2 = \frac{2\pi\hbar^2\beta}{m}$$

The upper and lower signs in (32.63b) correspond to bosons and fermions respectively.

For the hard sphere case we have the problem of a particle that is free except for being unable to enter a sphere of radius a . We study in particular the high temperature behavior of B_e . The physical interest attached to this quantity is discussed by Lieb. At high temperatures, β small, B_e has an exponential dependence on β^{-1} ; specifically Lieb shows that

$$\log \left[\frac{B_e}{B_e^{(0)}} \right] = -\frac{\pi}{4} \left(\frac{a}{\hbar} \sqrt{\frac{m}{\beta}} \right)^2 + O \left[\left(\frac{a}{\hbar} \sqrt{\frac{m}{\beta}} \right)^{3/2} \right] \quad (32.64)$$

where $B_e^{(0)}$ is the exchange coefficient for free particles in the absence of hard cores. $B_e^{(0)}$ can be calculated trivially, since it involves only the free particle Green's function and is

$$B_e^{(0)} = \mp \frac{N\Lambda^3}{4\sqrt{2}} (2s+1)^{-1} \quad (32.65)$$

Equation 32.64 is derived by finding upper and lower bounds on the Green's function, bounds having the virtue that their leading terms [the first term in (32.64)] coincide. Since we are dealing with a Wiener integral [cf. (32.62)] a lower bound can be obtained by throwing away the contributions of certain paths. This is not quite so straightforward as it sounds, and what we actually do is discard all paths outside of some tube and obtain a lower bound on the sum over paths in that tube. Then the location of the tube itself will be varied to get the best possible lower bound. For the upper bound one can include paths that enter the hard sphere, but then these are summed with somewhat reduced weight so the upper bound has some teeth to it. We discuss details of the lower bound only.

As usual, the Green's function is given by

$$\begin{aligned} G(\mathbf{r}, t; \mathbf{r}') &= \lim_{n \rightarrow \infty} G_n(\mathbf{r}, t; \mathbf{r}') \\ G_n(\mathbf{r}, t; \mathbf{r}') &= \int_{|\mathbf{r}_j| > a} d^3 r_1 d^3 r_2 \cdots d^3 r_n \left(\frac{m}{4\pi\hbar^2 \epsilon} \right)^{3(n+1)/2} \\ &\quad \times \exp \left[\frac{-m}{4\hbar^2 \epsilon} \sum_{j=0}^n (\mathbf{r}_{j+1} - \mathbf{r}_j)^2 \right] \end{aligned} \quad (32.66)$$

with $\epsilon = t/(n+1)$ and $\mathbf{r}_0 = \mathbf{r}'$, $\mathbf{r}_{n+1} = \mathbf{r}$ [\hbar is squared because t has been defined in (32.62) with unusual dimensions]. The range of integration for each of the \mathbf{r}_j , namely $|\mathbf{r}_j| > a$, takes into account the hard sphere constraint, and the range restriction could be replaced by an infinite potential in the exponent

$$\exp \left(-\epsilon \sum_j v(\mathbf{r}_j) \right) \quad (32.67)$$

as was done in previous discussions of first passage times (Sections 9 and 28). What we have often done with expressions of the form (32.66) is to expand about a classical path connecting the end points. However, when the sphere falls on the straight line connecting \mathbf{r} and \mathbf{r}' (as it does in our case, since $\mathbf{r}' = -\mathbf{r}$), there is no such classical path. One can then talk about "creep waves" around the sphere, and we mention work of this sort in the notes. However, in the present lower bound calculation the path around which we expand is the following.

Let b be a number satisfying $a < b < \min(r, r')$. Let W_b be a sphere of radius b centered at the origin. Let $\mathbf{x}(\tau)$ be a path that begins at \mathbf{r}' , is straight until it meets W_b , meets W_b tangentially, runs along a great circle of W_b , leaves W_b tangentially, and from there on is a straight line

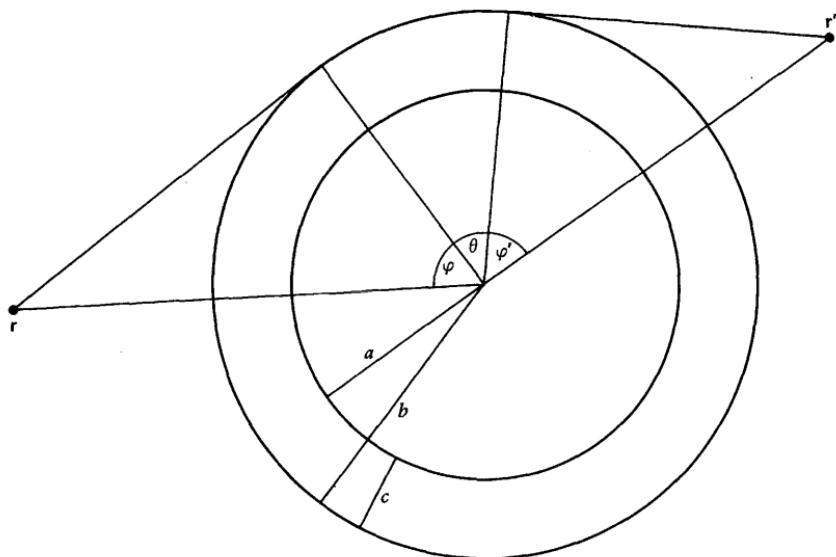


Fig. 32.1 Quantities involved in the hard sphere functional integral.

ultimately reaching \mathbf{r} . The path $\mathbf{x}(\tau)$ is taken to lie entirely in the plane containing \mathbf{r} , \mathbf{r}' and the origin (for $\mathbf{r}' = -\mathbf{r}$ the plane is not specified uniquely, which causes some difficulty related to caustics in the upper bound calculation, but can be ignored here). See Fig. 32.1. The length of the path is $L_b = r \sin \varphi + r' \sin \varphi' + b\theta$ in terms of the angles defined in the figure. To specify the path fully, its position on this curve at various times must be given: it covers equal distances in equal times, including the time spent on the surface of W_b .

The path integral is now expanded about $\mathbf{x}(\tau)$, but in contrast to an expansion about a classical path, a first-order term in the deviation remains. Let $\mathbf{y}(\tau) = \mathbf{r}(\tau) - \mathbf{x}(\tau)$ with $\mathbf{r}(\tau)$ the broken line path associated with the dummy variables in (32.66). Let the classical action be denoted by S . The action becomes

$$S[\mathbf{r}(\tau)] = S[\mathbf{x}(\tau)] + \delta S + \delta^2 S$$

Now $\delta^2 S$ is of the usual form, $\text{const.} \cdot \sum (\mathbf{y}_{j+1} - \mathbf{y}_j)^2$. The first-order term is

$$\frac{\text{const.}}{\epsilon} \sum \mathbf{y}_j \cdot (\mathbf{x}_{j+1} - 2\mathbf{x}_j + \mathbf{x}_{j-1}) \quad (32.68)$$

clearly arising from the curvature of $\mathbf{x}(\tau)$ on W_b and contributing only for those j for which $\mathbf{x}(\tau)$ lies on W_b .

Before analyzing (32.68) we note that the \mathbf{y} integrations will inherit a complicated constraint from the restriction on the \mathbf{r} integrations. However, if we limit \mathbf{y} to

$$|\mathbf{y}_j| < b - a \equiv c \quad \text{all } j \quad (32.69)$$

the \mathbf{r} constraints are satisfied. The price of this simplification is that paths on the straight stretches of $\mathbf{x}(\tau)$ cannot wander far nor can one get far from W_b . For the asymptotic order relations in $1/\beta$ that we seek these contributions disappear and so for that goal nothing has been lost. It is at this stage that paths are discarded in obtaining a lower bound. With the restriction (32.69) we can also get a bound on (32.68) which allows the remaining functional integral involving $\delta^2 S$ to be easily evaluated. The difference $\mathbf{x}_{j+1} - 2\mathbf{x}_j + \mathbf{x}_{j-1}$ is essentially $\epsilon^2 d^2 \mathbf{x}/d\tau^2$ and is constant along the circle. Hence using $|\mathbf{y} \cdot \ddot{\mathbf{x}}| \leq |\mathbf{y}| |\ddot{\mathbf{x}}|$ we have

$$\exp\left[-\frac{\left(\begin{array}{l} \text{positive} \\ \text{constant} \end{array}\right)}{\epsilon} \sum \mathbf{y}_j \cdot (\mathbf{x}_{j+1} - 2\mathbf{x}_j + \mathbf{x}_{j-1})\right] \geq \exp(-kc)$$

with k a positive constant of order unity (not ϵ). Lieb evaluates this constant and shows that the term in the functional integral of the form $\exp(\text{const.} \cdot \delta S)$ can be bounded from below by

$$\exp\left(-\frac{mL_b c \theta}{2\hbar^2 t}\right) \quad (32.70)$$

The term involving $S(\mathbf{x}(\tau))$ can easily be evaluated and contributes to the path integral a factor

$$\exp\left(-\frac{mL_b^2}{4\hbar^2 t}\right) \quad (32.71)$$

Finally, we confront the path integral involving the action $\delta^2 S$. The form of the action is just that involving a free particle, but the integration variables have the constraint $|\mathbf{y}_j| \leq c$. This constraint, however, is quite tractable, and we recognize our functional integral to be the Green's function for a particle constrained to remain *within* a sphere of radius c . Call this $G_c(\mathbf{r}, t; \mathbf{r}')$. Moreover, since $\mathbf{y}(\tau)$ by definition begins and ends at zero we want that Green's function with both its spatial arguments evaluated at the origin. Instead of doing the functional integral for G_c we appeal to its eigenfunction expansion and note that since $G_c(0, t; 0)$ is

needed only s -wave states contribute. Thus

$$G_c(0, t; 0) = \sum_{\substack{s \text{ wave} \\ \text{states } \alpha}} |\psi_\alpha(0)|^2 \exp(-tE_\alpha)$$

where E_α and ψ_α are energies and associated eigenstates. These quantities are well known, and for our bound we shall discard all but the lowest energy. Hence

$$G_c(0, t; 0) > \frac{\pi}{2c^3} \exp\left(-\frac{\hbar^2\pi^2}{mc^2}t\right) \quad (32.72)$$

Combining all the foregoing lower bounds we have

$$G(\mathbf{r}, t; \mathbf{r}') > \frac{\pi}{2c^3} \exp\left\{-\frac{L_b^2 + 2L_b c \theta}{\alpha t} - \left(\frac{\pi\sqrt{\alpha}}{2c}\right)t\right\} \quad (32.73)$$

with $\alpha = 4\hbar^2/m$ and $c = b - a$. At this point one should set $\mathbf{r}' = -\mathbf{r}$ and maximize the right hand side of (32.73) with respect to c (for the best possible bound). The complicated c dependence of L_b (through fixing the angles φ , φ' , and θ) makes this procedure difficult. However, we expect exchange effects to be most important when the particles are closest, that is, when $r \gtrsim a$ in the integral (32.63b). For this r the quantity c must be close to zero, that is $c \sim r - a$. For simplicity we now take $c = r - a$ for all r . Since for this choice it turns out that the upper and lower bounds meet, we know that to leading asymptotic order no other choice could have been better.

It remains to take the expression in (32.73) evaluated at $c = r - a$ and to integrate over r [as in (32.63b)] to obtain one of the inequalities needed for the order relation (32.64). We shall not study this integral and refer to Lieb's paper for details.

The upper bound on the path integral is obtained by expanding around the same path $\mathbf{x}(\tau)$ but with $b = a$. For the straight stretches of $\mathbf{x}(\tau)$ all constraints are dropped while on the sphere paths within the sphere are allowed but with reduced contributions. The final result is an upper bound providing the other inequality needed for (32.64).

Notes

The work just described is taken from

E. H. Lieb, Calculations of Exchange Second Virial Coefficient of a Hard-Sphere Gas by Path Integrals, *J. Math. Phys.* **8**, 43 (1967)

In that paper will be found a discussion of the physical significance of the exchange coefficient, further details of the lower bound calculation, and the upper bound calculation. Background on the virial expansion is available in any text on statistical mechanics.

The calculation of the exchange coefficient is essentially a diffraction problem and in addition to the references given by Lieb we mention

V. S. Buslaev, Continuum Integrals and the Asymptotic Behavior of the Solutions of Parabolic Equations as $t \rightarrow 0$. Applications to Diffraction, in *Topics in Math. Phys.*, Sh. Birman, Ed., English translation by Consultants Bureau, New York, 1967

This paper also uses a functional integral technique and gets an asymptotic formula for the Green's function $G(\mathbf{r}, t; \mathbf{r}')$ for $t \rightarrow 0$. Buslaev treats a general convex obstacle in many dimensions and in particular gets a creep wave (or an "enveloping ray" for the associated classical path on the surface of the convex body) contribution to the functional integral.

32.5 ADDING PATHS BY COMPUTER

On first encountering the path integral one may be put off by the immensity of the task of adding an infinity of paths. Then one learns the prescription for doing the sum that is implicit in the Trotter product formula as well as the plethora of analytical tricks for avoiding the doing of any sum at all, and the path integral seems less terrifying. But some brave souls, armed with the chief fruit of modern technology (next to lightweight bicycles) have undertaken to add the paths on a computer. Of course, the space and time coordinates must be discretized but even if an infinite dimensional integral is out of reach, the actual dimension that is attempted can be quite high. Now, the way to do a high dimensional integral is by Monte Carlo methods coupled with importance sampling—one tries to find the region in the multidimensional space that contributes most to the integral by means of a certain stochastic process taking values in that space. To apply the Monte Carlo method to the functional integral it is easiest to work in imaginary time so that all contributions are real and the most important regions provide maxima rather than stationary points of the integrand.

We know from the Feynman-Kac formula that a Green's function evolving in imaginary time is ultimately dominated by the ground state energy and wave function, and the sum over paths that we shall describe is proposed as a means for numerical evaluation of these quantities.

Consider then a spinless particle in one dimension with $H = \frac{1}{2} p^2 + V$ and $\hbar = 1$ for convenience. By summing over paths beginning and ending

at the same point x and taking time T we have

$$\int_{(x, 0)}^{(x, T)} dx(\cdot) \exp(-S) = G(x, -iT; x) = \sum_{\alpha} |\psi_{\alpha}(x)|^2 \exp(-E_{\alpha}T) \\ \sim |\psi_0(x)|^2 \exp(-E_0T) \quad \text{for } T \rightarrow \infty \quad (32.74)$$

where α labels eigenstates of H and 0 represents the ground state. Doing the functional integral on the left of (32.74) for large times thus provides E_0 and $\psi_0(x)$. (Although the Monte Carlo method can be used for quantum mechanics in any number of dimensions, in one dimension ψ_0 can always be taken real and positive and the square root of ψ_0^2 performed). These quantities may be extracted in a variety of ways. Letting the system evolve in T the total integral of G shrinks, this shrinkage being measurable and governed by E_0 . Then by renormalizing with $\exp(+TE_0)$ the wave function is extracted. The method used here gets $\psi_0(x)$ and then evaluates the expectation of H in this state. The problem is therefore to evaluate the left hand side of (32.74), that is, the infinite n limit of

$$G_n(x, -iT; x) = \int dx_1 \cdots dx_n \left(\frac{1}{2\pi\epsilon} \right)^{(n+1)/2} \\ \times \exp \left\{ -\epsilon \sum_{j=0}^n \left[\frac{1}{2} \left(\frac{x_{j+1} - x_j}{\epsilon} \right)^2 + V(x_j) \right] \right\} \quad (32.75)$$

with $x_0 = x_{n+1} = x$ and $\epsilon = T/(n+1)$. We wish to identify those sequences $\xi \equiv (x_0, x_1, \dots, x_n)$ (interpreted as paths) that are most important in giving the sum in (32.75). For this we start from an arbitrary sequence and move randomly toward the more important contributors in a particular way. Let

$$E(\xi; \epsilon) = E(x_0, x_1, \dots, x_n; \epsilon) \\ = \sum_{j=0}^n \left[\frac{1}{2} \left(\frac{x_{j+1} - x_j}{\epsilon} \right)^2 + V(x_j) \right] \quad (32.76)$$

(as usual $x_{n+1} = x_0$). Start with some initial sequence, designated $\xi^{(0)}$. Randomly select an integer j between 1 and n and real η between 0 and 1 and define ξ' to be the same as $\xi^{(0)}$ except for the j th component. Let the j th component of ξ' be

$$x'_j = x_j + \alpha(2\eta - 1) \quad (32.77)$$

where the number α , like ϵ and n , is a parameter characterizing the numerical integration and discretization. Now let $\Delta E = E(\xi'; \epsilon) - E(\xi^{(0)}; \epsilon)$.

If $\Delta E < 0$, the “energy” (E) is lowered and we take $\xi^{(1)} = \xi'$. If $\Delta E > 0$, we let $\xi^{(1)} = \xi'$ with probability $\exp(-\epsilon \Delta E)$ and let $\xi^{(1)} = \xi^{(0)}$ with probability $1 - \exp(-\epsilon \Delta E)$. This last selection is implemented by means of another random number η' (generated by the computer) between 0 and 1 such that if $\eta' < \exp(-\epsilon \Delta E)$ we let $\xi^{(1)} = \xi'$, and otherwise $\xi^{(1)} = \xi^{(0)}$. Next $\xi^{(2)}$ is gotten from $\xi^{(1)}$ in the same way. The sequence $\xi^{(k)}$ tends to relax toward a class of most important contributors to the sum, but because of the rule allowing $\Delta E > 0$ also, the sequence does not get stuck at local minima of E .

In principle, what can be done next is to evaluate $\exp(-S)$ for each $\xi^{(k)}$ in the sequence and add. Moreover, this calculation would have to be done for each x ($= x_0$) appearing as the argument of G . However, this method would waste information on the values of $\exp(-S)$ built into the sequence $\xi^{(k)}$ through its selection process and would ignore data on all positions, not just the nominal endpoints, that go into the construction of the sequence.

The efficient method actually used to evaluate $\psi_0(x)$ from $\{\xi^{(l)}\}$ is as follows. The real line is divided into a large number of bins $[m\Delta, (m+1)\Delta]$ $m=0, \pm 1, \dots$. Now consider some particular step of the stochastic process that generates $\{\xi^{(l)}\}$. At, say, the k th step some coordinate x_j is given an opportunity to assume a new value $x_j + \alpha(2\eta - 1)$. The j th component of $\xi^{(k+1)}$ is therefore either x_j or $x_j + \alpha(2\eta - 1)$. In either case the j th component of $\xi^{(k+1)}$ falls into some bin $[m\Delta, (m+1)\Delta]$. A point is scored for that bin. As the stochastic process continues each bin accumulates points with each successful entry or persistence in that bin. $\psi_0(x)^2$ is then proportional to the total number of points in the bin associated with x , and by normalizing the point score the square of the wave function is obtained.

Why does this method work? Basically because the bias in selecting $\xi^{(k)}$ is just the exponential we want to evaluate, and in the reference below this is the explanation given. One can, however, look at it in a number of other ways—either through statistical mechanics or through probability theory, the latter once again playing on the connection between probability and potential theory that was one of the themes of Section 9.

Let us first take a statistical mechanical approach. For fixed n in the approximation for G [cf. (32.75)] we can look upon $E(\xi; \epsilon) = E(x_0, x_1, \dots, x_n; \epsilon)$ as the energy of $n+1$ particles coupled by nearest neighbor harmonic forces (from $(x_{j+1} - x_j)^2$) and one particle forces (from $V(x_j)$) with periodic boundary conditions. It is easy to see that with the system undergoing the stochastic process described above $\xi^{(k)} \rightarrow \xi^{(k+1)}$ the probability of finding any particular ξ is proportional to $\exp(-\epsilon E(\xi, \epsilon))$, that is, the Gibbs distribution with inverse temperature ϵ . (This feature is used extensively in giving stochastic dynamics to the Ising model. The

simple proof of our assertion is given in the Landau and Alben reference cited below.)

Each of the $n+1$ "particles" has varying values of x , and we can ask for the density of particles at $x=a$; call this $\rho(a)$. Using the Gibbs distribution ρ is given by

$$\rho(a) = \frac{\sum_{\xi} \exp(-\epsilon E(\xi; \epsilon)) \delta(x_j - a)}{\sum_{\xi} \exp(-\epsilon E(\xi; \epsilon))} \quad (32.78)$$

where some particular j has been arbitrarily chosen and used in the argument of the δ -function by virtue of the translational invariance of the system (the gas of particles has periodic boundary conditions). Taking $j=0$, the object in the numerator of (32.78) is just the quantity G_n of (32.75) with $x=a$. The normalization is irrelevant and cancels. The denominator of (32.78) is the integral of the numerator over a and is therefore the trace of G . The value of T to be used in G_n is, following the definitions after (32.75), $T=(n+1)\epsilon$. It follows that

$$\rho(a) = \frac{G_n(a, -i(n+1)\epsilon; a)}{\int dx G_n(x, -i(n+1)\epsilon; x)} \quad (32.79)$$

But this—for sufficiently large $n\epsilon$ —is $\psi_0(a)^2$, by (32.74). Now the justification for the point scoring above is evident—it was simply a way of measuring the density of particles in x space.

Once the wave function ψ_0 is known, the ground state energy can be calculated from $E_0 = (\psi_0, H\psi_0)$.

There is yet another way to look at the stochastic process described above, namely as the motion of a single particle in an Ornstein-Uhlenbeck process. This takes place in a "potential" $U(x) \sim -\log \psi_0(x)$ and the density of particles satisfies a Fokker-Planck equation. This is not quite the same as the stochastic process discussed previously (Section 9) in connection with the expectation over Brownian motion paths of $\exp(-\int_0^T V(x(\tau)) d\tau)$. There the potential V was related to an absorption probability. The potential U differs and can be thought of as representing asymmetry in left-right step size or frequency for a discrete process. There are relations, however, between the two formulations. We shall not go into these matters here.

The numerical sum over paths does reasonably well on various one-dimensional systems, and comparison can be made to systems having known solutions. There is some difficulty where "metastability" is encountered and the ground state and first excited state are very close in energy. Further details can be found in the paper cited below.

Notes

The summing over paths by computer treated here appears in

S. V. Lawande, C. A. Jensen, and H. L. Sahlin, Monte Carlo Integration of the Feynman Propagator in Imaginary Time, *J. Comp. Phys.* **3**, 416 (1969)

The stochastic processes described above acts like a heat reservoir in bringing the system to the Gibbs distribution. A proof of this appears in an appendix to

D. P. Landau and R. Alben, Monte Carlo Calculations as an Aid in Teaching Statistical Mechanics, *Am. J. Phys.* **41**, 394 (1973)

The original use of Monte Carlo methods is the paper

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953)

Another computer evaluation of a Wiener integral, one that also uses Monte Carlo techniques and in fact preceded the work of Lawande et al. cited above is that of

L. D. Fosdick and H. F. Jordan, Path Integral Calculation of the Two Particle Slater Sum for He^4 , *Phys. Rev.* **143**, 58 (1966)

There is an extensive literature on functional integration and the Fokker-Planck equation. We cite a few recent papers:

R. Graham, Lagrangian for Diffusion in Curved Phase Space, *Phys. Rev. Lett.* **38**, 51 (1977)

L. Garrido, D. Lurie, and M. San Miguel, Stochastic Quantization and Path Integral Formulation of Fokker-Planck Equation, *Phys. Lett.* **67A**, 243 (1978)

F. Langouche, D. Roekaerts and E. Tirapegui, On the Path Integral Solution of the Master Equation, *Phys. Lett.* **68A**, 418 (1978)

H. Dekker, Path Integrals for Diffusion Processes in Riemannian Spaces, *Phys. Lett.* **69A**, 241 (1978)

The Ornstein-Uhlenbeck process, the associated Fokker-Planck equation, and the functional integral can in some cases be related to Brownian motion with absorption (cf. Section 9) by means of a simple transformation (related to a gauge transformation). See the paper

N. G. van Kampen, A Soluble Model for Diffusion in a Bistable Potential, *J. Stat. Phys.* **17**, 71 (1977)

32.6 A PERTURBATION EXPANSION USING THE PATH INTEGRAL

In Section 10 we found that the perturbation expansion for the path integral can be obtained simply by expanding an exponential

$$\begin{aligned} G(x, t; x') &= \int dx(\cdot) \exp \left\{ \frac{i}{\hbar} \int_0^t \left[\frac{1}{2} m \left(\frac{dx}{d\tau} \right)^2 - \lambda V(x) \right] d\tau \right\} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-i\lambda}{\hbar} \right)^n \int dx(\cdot) \left[\int_0^t V(x(\tau)) d\tau \right]^n \exp \left[\frac{i}{\hbar} \int_0^t \frac{1}{2} m \left(\frac{dx}{d\tau} \right)^2 d\tau \right] \end{aligned} \quad (32.80)$$

In the papers listed below the authors make use of the fact that if V is replaced by its Fourier transform a tremendous simplification results. Specifically, writing

$$V(x) = \int e^{ikx} f(k) dk$$

the x dependence becomes a linear term in an exponent and the path integrals in (32.80) all become Gaussian and yield simple closed form expressions. The price of course is the introduction of integrals over k , one integration for each order of perturbation theory. In some cases these integrals can be done and the series summed (to achieve this goal it was found useful to integrate over one of the position arguments of G).

Notes

This work is described in

M. J. Goovaerts and J. T. Devreese, Analytic Treatment of the Coulomb Potential in the Path Integral Formalism by Exact Summation of a Perturbation Expansion, *J. Math. Phys.* **13**, 1070 (1972)

M. J. Goovaerts, A. Babcenko, and J. T. Devreese, A New Expansion Method in the Feynman Path Integral Formalism: Application to a One-Dimensional Delta-Function Potential, *J. Math. Phys.* **14**, 554 (1973)

32.7 SOLVABLE PATH INTEGRAL WITH THE POTENTIAL $ax^2 + b/x^2$

Consider the Lagrangian

$$L = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} m \omega^2 x^2 - \frac{g}{x^2} \quad (32.81)$$

for a particle on the half line $x > 0$. The path integral for this Lagrangian can be done, even if ω has an explicit time dependence. This is not entirely surprising since a $1/x^2$ term is the same as the centrifugal contribution in polar coordinates and therefore a term of that sort already appears for a free particle. The path integral is actually done by iteration of the discretized expression for the propagator. The functions and functional identities that arise involve the modified Bessel functions and are much the same as those appearing in the path integral for a free particle in polar coordinates.

For the case of constant ω the propagator for the Lagrangian (32.81) is

$$G(b, t; a, 0) = \frac{m\omega\sqrt{ab}}{i\hbar \sin \omega t} \exp\left\{\frac{im\omega}{2\hbar}(b^2 + a^2)\cot \omega t\right\} I_\gamma\left(\frac{m\omega ab}{i\hbar \sin \omega t}\right) \quad (32.82)$$

with I_γ the modified Bessel function and

$$\gamma = \frac{1}{2}\left(1 + \frac{8mg}{\hbar^2}\right)^{1/2} \quad (32.83)$$

Note that for $\gamma \rightarrow \infty$ the modified Bessel function has the asymptotic behavior

$$I_\gamma(y) = \left(\frac{1}{2\pi y}\right)^{1/2} \exp\left[y - \frac{\frac{1}{2}\left(\gamma^2 - \frac{1}{4}\right)}{y} + O\left(\frac{1}{y^2}\right)\right]$$

so that for $g \rightarrow 0$ G reduces to the harmonic oscillator functional form. From (32.83) the requirement $g > -\hbar^2/8m$ is also evident, and in the corresponding polar coordinate problem is what is needed to prevent the particle from "falling to the center."

The Lagrangian (32.81) is also obtained from an appropriate separation of variables in a three-body problem (where the particles move in one dimension) with Lagrangian

$$L = \frac{1}{2}(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) - \frac{1}{4}\omega^2[(x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_1 - x_3)^2] - g(x_1 - x_2)^{-2} \quad (32.84)$$

(one can also include terms $-g(x_2 - x_3)^{-2} - g(x_1 - x_3)^{-2}$). Hence this work also represents a path integral solution for a three-body problem.

Notes

The foregoing work appears in

K. C. Khandekar and S. V. Lawande, Path Integral for a Three Body Problem, *J. Phys. A* **5**, 812 (1972)

K. C. Khandekar and S. V. Lawande, Path Integration of a Three Body Problem, *J. Phys. A* **5**, L57 (1972)

K. C. Khandekar and S. V. Lawande, Exact Propagator for a Time-Dependent Harmonic Oscillator with and without a Singular Perturbation, *J. Math. Phys.* **16**, 384 (1975)

M. J. Goovaerts, Path Integral Evaluation of a Nonstationary Calogero Model *J. Math. Phys.* **16**, 720 (1975)

The path integral in polar coordinates was done by

S. F. Edwards and Y. V. Gulyaev, Path Integrals in Polar Coordinates, *Proc. Roy. Soc. A* **279**, 229 (1964)

D. Peak and A. Inomata, Summation over Feynman Histories in Polar Coordinates, *J. Math. Phys.* **10**, 1422 (1969)

The three-body problem of (32.84) was considered by

F. Calogero, Solution of a Three-Body Problem in One Dimension, *J. Math. Phys.* **10**, 2191 (1969)

32.8 SUPERFLUIDITY

This is one of the early applications of path integration and an exposition can be found in the following sources:

R. P. Feynman, Atomic Theory of the λ Transition in Helium, *Phys. Rev.* **91**, 1291 (1953)

R. P. Feynman, *Statistical Mechanics: A Set of Lectures*, Benjamin, New York, 1972, Chapter 11

32.9 FERMIONS

Functional integrals for boson fields are easy. For fermions they are not. A recent contribution with references to some of the early literature is

D. E. Soper, Construction of the Functional-Integral Representation for Fermion Green's Functions, *Phys. Rev. D* **18**, 4590 (1978)

A different approach was taken in the following paper:

J. R. Klauder, The Action Option and a Feynman Quantization of Spinor Fields in Terms of Ordinary C-Numbers, *Ann. Phys.* **11**, 123 (1960)

32.10 BOOKS AND REVIEW PAPERS ON PATH INTEGRALS

Many of the following sources have been quoted at various places in the book, but for convenience we collect a number of general references.

J. Schwinger, Ed., *Selected Papers on Quantum Electrodynamics*, Dover, New York, 1958
 Feynman's original papers on path integration are reprinted here.

Mark Kac, *Probability and Related Topics in Physical Sciences*, Interscience, New York, 1959

W. T. Martin and I. Segal, Ed., *Analysis in Function Space*, MIT Press, Cambridge Mass., 1964 (conference proceedings)

R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965

A. Katz, *Classical Mechanics, Quantum Mechanics, Field Theory*, Academic Press, New York, 1965

A. M. Arthurs, Ed., *Functional Integration and Its Applications*, Clarendon Press, Oxford, 1975 (conference proceedings)

G. Rosen, *Formulations of Classical and Quantum Dynamical Theory*, Academic Press, New York, 1969

S. A. Albeverio and R. J. Hoegh-Krohn, *Mathematical Theory of Feynman Path Integrals*, Lecture Notes in Mathematics, Vol. 523, Springer-Verlag, Berlin, 1976

G. J. Papadopoulos and J. T. Devreese, Eds., *Path Integrals and their Applications in Quantum, Statistical and Solid State Physics* (Proc. NATO Adv. Study Ins., Antwerp 1977), Plenum, New York, 1978

S. Albeverio, P. Combe, R. Hoegh-Krohn, G. Rideau, M. Sirugue-Collin, M. Sirugue, and R. Stora, eds. *Feynman Path Integrals*, Proc. Int. Colloq., Springer, New York, 1979

B. Simon, *Functional Integration and Quantum Physics*, Academic Press, New York, 1979

D. J. Amit, *Field Theory, The Renormalization Group and Critical Phenomena*, McGraw-Hill, New York, 1978

Review Papers:

E. W. Montroll, Markoff Chains, Wiener Integrals and Quantum Theory, *Commun. Pure and Appl. Math.* **5**, 415 (1952)

I. M. Gel'fand and A. M. Yaglom, Integration in Function Space and its Application in Quantum Physics, *J. Math. Phys.* **1**, 48 (1960)

S. G. Brush, Functional Integrals and Statistical Physics, *Rev. Mod. Phys.* **33**, 79 (1961)

L. Streit, An Introduction to Theories of Integration over Function Spaces, *Acta Phys. Austriaca (Suppl.)* **2** (1965)

C. DeWitt-Morette, L'intégrale fonctionnelle de Feynman: Une Introduction *Ann. Inst. Henri Poincaré, Sec. A, Physique Théorique*, **11**, 153 (1969)

E. S. Abbers and B. W. Lee, Gauge Theories, *Phys. Rep. (Phys. Lett. C)* **9**, 1 (1973)

F. W. Wiegel, Path Integral Methods in Statistical Mechanics, *Phys. Rep. (Phys. Lett. C)* **16**, 57 (1975)

- J. B. Keller and D. W. McLaughlin, The Feynman Integral, *Am. Math. Monthly* **82**, 451 (1975)
- C. DeWitt-Morette, A. Maheshwari, and B. Nelson, Path Integration in Non-Relativistic Quantum Mechanics, *Phys. Rep. (Phys. Lett. C)* **50**, No. 5, 255 (1979)
- M. S. Marinov, Path Integrals in Quantum Theory: An Outlook of Basic Concepts, *Phys. Rep. (Phys. Lett. C)* **60**, 1 (1980)

Two general reviews of semiclassical and WKB methods are

- M. V. Berry and K. E. Mount, Semiclassical Approximations in Wave Mechanics, *Rep. Prog. Phys.* **35**, 315 (1972)
- A. Voros, Semi-Classical Approximations, *Ann. Inst. Henri Poincare, Sec. A*, **24**, 31 (1976)

Author Index

- Abers, E. S., 314, 347
Abramowitz, M., 78
Aharonov, Y., 15, 195, 196, 210
Alben, R., 342, 343
Albeverio, S., 63, 64, 117, 347
Alimov, A. L., 311
Amit, D. J., 189, 302, 314, 347
Anderson, P. W., 257
Anderson, J. L., 235
Arnold, V. I., 313
Arthurs, A. M., 128, 189, 211, 224, 332,
 333, 347
Aspnes, D., 302
Avez, A., 313

Babcenco, A., 48, 344
Baker, G. A., Jr., 289
Balian, R., 163
Bander, M., 157
Banks, T., 288
Bargmann, V., 245, 257
Bekenstein, J. D., 229, 235
Bender, C. M., 189, 288, 332
Berlin, T. H., 294, 302
Berry, M. V., 142, 348
Bezak, V., 270
Birman, Sh., 339
Bleistein, N., 169
Bliss, G. A., 91
Bloch, C., 163
Bohm, D., 195, 196, 210
Boon, M., 257
Bopp, F., 188, 223
Born, M., 146
Brezin, E., 288

Brush, S. G., 241, 347
Burton, W. K., 196
Buslaev, V. S., 41, 212, 339

Callan, C., 288
Calogero, F., 346
Cameron, R. H., 64
Campbell, W. B., 163
Casher, A., 253, 256
Chang, S. S., 313
Chester, C., 132, 141
Child, M. S., 142
Chretien, M., 331
Chu, F. Y. F., 313
Clutton-Brock, M., 41, 311
Cohen, L., 257, 312
Coleman, S., 106, 288
Collins, J. C., 288
Combe, P., 347
Connor, J. N. L., 141-142
Cooper, F., 332
Copson, E. T., 78
Courant, R., 80, 91
Coyne, S., 128

Dashen, R. F., 313
Davies, H., 311
DeBorde, A. H., 196
DeBruijn, N. G., 78
Dekker, H., 312, 343
Dente, G. C., 324
Deser, S., 331
DeWitt, B., 96, 106, 216, 217, 224
DeWitt-Morette, C., 63, 64, 96, 106, 130, 185
 188, 209, 210, 224, 236, 311, 347, 348

- Devreese, J. T., 48, 142, 188, 344, 347
 Dicke, R. H., 162
 Dirac, P. A. M., 14, 15
 Domb, C., 302
 Donsker, M., 96, 180, 266, 269
 Dowker, J. S., 41, 210, 223, 224, 258, 312
 Dreyfus, T., 106, 314
 Duistermaat, J. J., 142
 Dym, H., 106
 Dyson, F., 98, 106
- Edwards, R. E., 181
 Edwards, S. F., 224, 270, 333, 346
 Ellis, R. S., 129
 Erdelyi, A., 21, 78
 Esterling, D. M., 189
 Eve, M., 169
 Fadeev, L. D., 313
 Fanelli, R., 311
 Ferrell, R. A., 270
 Feynman, R. P., 3, 12, 15, 26, 180, 185,
 188, 214, 217, 226, 228, 234, 241, 311,
 323, 324, 325, 346, 347
 Fibich, M., 270
 Finkler, P., 163
 Fisher, M. E., 283, 289
 Fock, V. A., 214, 226
 Fokker, A. D., 323
 Fosdick, L. D., 343
 Frank, W. M., 107
 Friedman, B., 132, 141
 Fröhlich, 180, 270
 Furman, M., 302
- Garrido, L., 312, 343
 Garrod, C., 311
 Gawedzki, K., 311, 314
 Gel'fand, I. M., 39, 64, 347
 Gelman, D., 162
 George, T. F., 163
 Gervais, J. L., 313
 Gibbons, G. W., 236
 Glauber, R. J., 257
 Glimm, J., 47, 325
 Golubitsky, M., 130
 Goldstein, H., 88, 91, 205
 Goovaerts, M. J., 48, 344, 346
 Gotze, W., 302
 Graham, R., 312, 343
 Green, M. S., 302
- Griego, R. J., 63
 Gross, E. P., 181, 270, 331
 Gulyaev, V. V., 224, 270, 346
 Guralnik, G. S., 332
 Gutzwiller, M. C., 48, 154, 155, 157, 313,
 314
- Haag, R., 188, 223
 Hagedorn, G. A., 21, 116, 117
 Hamblen, D., 302
 Hamermesh, M., 202, 205
 Hamilton, J. F., Jr., 188
 Handelsman, R. A., 169
 Hardy, G. H., 196
 Hartle, J. B., 233, 234, 235
 Hassing, R. F., 189
 Hasslacher, B., 313
 Hawking, S. W., 233, 234, 235, 236
 Hecht, R., 302
 Helfand, E., 270
 Hellwarth, R. W., 180
 Hemmer, P. C., 283
 Hersh, R., 63
 Hibbs, A. R., 3, 12, 180, 188, 241, 324, 347
 Hilbert, D., 80, 91
 Hilton, P. J., 205
 Hirshfeld, A. C., 258, 312, 313
 Hirth, J. P., 288
 Hocking, J. G., 205
 Hoegh-Krohn, R. J., 63, 64, 117, 347
 Horvathy, P. A., 210, 314
 Huang, K., 289
- Iddings, C. K., 180
 Inomata, A., 346
 Israel, W., 236
 Ito, K., 28, 30, 311
 Itzykson, C., 157
- Jaffe, A., 47, 325
 Jeffreys, H., 78
 Jensen, C. A., 343
 Jevicki, A., 313
 Jones, C. E., 163
 Jones, R., 270
 Jordan, H. F., 343
- Kac, M., 42, 47, 58, 63, 241, 269, 283, 294,
 302, 329, 331, 347
 Kadanoff, L. P., 291, 302

- Kane, J., 302
Katz, A., 312, 347
Keller, J. B., 157, 163, 169, 348
Khandekar, D. C., 48, 346
Kim, D., 289
Kishore, R., 189
Klauder, J. R., 188, 256, 257, 346
Knoll, J., 163
Koeling, T., 163
Korepin, V. E., 313
Kostant, B., 210
Kramers, H. A., 288
Kruskal, M., 231
Kunz, H., 289
Kuratsuji, H., 257
Kuttner, F., 302

Laidlaw, M. G. G., 209, 210
Lam, C. S., 96, 106
Landau, L. D., 154, 342, 343
Landauer, R., 288
Langer, J. S., 256, 270, 284, 288
Langouche, F., 312, 343
Lawande, S. V., 48, 343, 346
Lebowitz, J. L., 289
Lee, B. W., 314, 347
Lee, L. L., 64
Leschke, H., 313
Levit, S., 95, 106, 130, 142, 163, 314
Lewis, E. A. S., 302
Lieb, E. H., 333, 337, 338, 339
Lifshitz, E. M., 154
Lipatov, L. N., 288
Lipkin, H. J., 257
Loeve, M., 181
Lothe, J., 288
Lowe, M. J., 289
Ludwig, D., 169
Lukes, T., 270
Lurie, D., 253, 256, 343
Lusternik, L. A., 212
Luttinger, J. M., 269

Ma, S. K., 241, 291, 298, 299, 301, 302
McCraw, R. J., 289
McLaughlin, D. W., 106, 128, 155, 157,
 163, 224, 313, 348
Maheshwari, A., 270, 311, 348
Malfliet, R. A., 163
Malgrange, B., 140

Marcus, R. A., 141
Marinov, M. S., 348
Martin, W. T., 347
Mather, J. N., 140
Matthews, P. T., 188
Mayes, I. W., 258
Messiah, A., 257
Metropolis, N., 343
Miller, W. H., 142, 163
Milnor, J., 90, 91
Misheloff, M. N., 163
Misner, C. W., 236
Miura, T., 188
Mizrahi, M. M., 258, 311
Möhring, K., 314
Montroll, E. W., 39, 347
Morawetz, C. S., 169
Morette, C. *see DeWitt-Morette*
Morse, M., 90, 91
Mount, K. E., 142, 348
Mühlschlegel, B., 188, 332

Nauenberg, M., 302
Nelson, B., 311, 348
Nelson, E., 12, 39, 64
Neveu, A., 313
Newman, C. M., 289
Newton, R. G., 162
Niemeijer, T., 291, 302
Nirenberg, L., 140
Norcliffe, A., 154, 157

Olver, F. W. J., 78

Pak, N. K., 211
Palciauskas, V. V., 302
Papadopoulos, G. J., 106, 142, 188, 270,
 347
Patrascioiu, A., 288, 289
Patterson, E. M., 205
Pauli, W., 96, 106, 223
Peak, D., 346
Pearle, P., 311
Pechukas, P., 142, 146, 162, 163
Pelte, D., 130, 142, 163, 314
Penner, A. P., 142, 162
Penrose, O., 289
Percival, I. C., 157
Perelomov, A. M., 257
Perry, M. J., 236

- Platzman, P. M., 180
 Poincaré, H., 73
 Postnikov, M. M., 91
 Rayl, M., 302
 Revzen, M., 253, 256, 258
 Rideau, G., 347
 Riesz, F., 212
 Roberts, M. J., 157
 Roekaerts, D., 312, 343
 Rosen, G., 325, 347
 Rosen, J. S., 129
 Rosen, N., 188, 223
 Rosenbluth, A. W., 343
 Rosenbluth, M. N., 343
 Sahlin, H. L., 343
 Sakita, B., 313
 Sakurai, J. J., 257
 Salam, A., 188
 San Miguel, M., 343
 Schaeffer, R., 163
 Schilder, M., 96, 106, 117, 129
 Schrieffler, J. R., 189
 Schulman, L. S., 41, 128, 188, 196, 210,
 211, 224, 235, 258, 289
 Schultz, T. D., 180
 Schwarzschild, K., 323
 Schweber, S. S., 256
 Schwinger, J., 15, 347
 Scott, A. C., 313
 Segal, I., 347
 Segular, R., 312
 Senitzky, I. R., 257
 Sewell, G. L., 284, 289
 Shepherd, P. J., 312
 Simon, B., 107, 270, 289
 Sirugue, M., 347
 Sirugue-Collin, M., 347
 Smilansky, U., 96, 106, 130, 142, 163, 314
 Sobolev, V. J., 212
 Souillard, B., 289
 Souriau, J. M., 210
 Soper, D. E., 288, 346
 Spanier, E. H., 205
 Spencer, T., 270
 Spruch, L., 162
 Stanley, H. E., 302
 Stora, R., 347
 Stegun, I. A., 78
 Stine, J. R., 141
 Stone, M., 288
 Stratovich, R. L., 30
 Streit, L., 347
 Sutcliffe, W., 257
 Suzuki, T., 257, 313
 Swanson, J. A., 288
 Swift, J., 302
 Sz-Nagy, B., 212
 Tarski, J., 188
 Teller, A. H., 343
 Teller, E., 343
 Testa, F. J., 257
 Tetrode, H., 323
 Thom, R., 140
 Thompson, C., 329
 Thorne, K. S., 236
 Tirapegui, E., 312, 343
 Tobocman, W., 311
 Truman, A., 63, 64
 Tschumi, P., 130
 Tze, H. C., 211
 Uhlenbeck, G. E., 283
 Ursell, F., 132, 141
 Van der Waals, J. D., 283
 Van Hove, L., 91, 312
 van Kampen, N. G., 343
 van Leeuwen, J. M. J., 291, 302
 Van Vleck, J. H., 106
 Varadhan, S.R.S., 96, 106, 117,
 180, 266, 269
 Vardi, M., 15
 Voros, A., 348
 Wallace, D. J., 289
 Wallace, R., 142, 162
 Wheeler, J. A., 217, 236, 323,
 325
 Wiegel, F. W., 241, 347
 Wiener, N., 28
 Wilson, K. G., 290, 291, 302
 Wittke, J. P., 162
 Wolf, E., 146
 Wright, E. M., 196
 Wu, T. T., 210, 212, 213, 288
 Yaglom, A. M., 39, 64, 347
 Yang, C. N., 210, 212, 213
 Young, G. S., 205
 Zak, J., 257
 Zinn-Justin, J., 288
 Zittarz, J., 270

Subject Index

- Action, 7
analytically continued, 155, 156
at a distance, 323, 325
at a distance electrodynamics, 174
forced harmonic oscillator, 38
harmonic oscillator, 38
instanton, 275
Aharonov-Bohm effect, 195, 196, 210
Airy function, 122, 123, 135, 139
generalized, 127
Algebraic topology, 197
Amplitude, probability, 13
Analytic continuation, 64, 99, 105, 162,
168, 230, 232, 272, 279, 280, 281, 289,
327, 331. *See also* Barrier penetration;
Tunneling
Anderson model, 270, 331
Anharmonic oscillator, 96-117, 271
Annihilation, particle-antiparticle, 227, 228
Anticommuting C-numbers, 188
Arcwise connected, 197
Asymptotic analysis, 73-78
Asymptotic approximation, 16-21, 106,
117, 126, 129
uniform, 131-142
Asymptotic expansion, 139, 169
Asymptotic series, 73-78, 97, 134
exercises, 76
Atomic scattering, 158
Auxiliary coordinates, 323

Bare mass, 170
Barrier penetration, 155, 157, 162, 163,
271, 272, 273. *See also* Analytic
continuation; Tunneling

Bead on a ring, 190
Bessel functions, 345
Black body temperature, 234
Black hole:
absorption probability, 234
emission, 229, 230, 232, 233, 234, 235
singularity, 231
temperature, 230, 233
thermodynamics, 236
Black holes, 229-236
Bloch waves, 195
Bohr-Sommerfeld quantization rules, 143,
154
Boosts (Lorentz transformations), 235
Borel resummation, 75
Born approximation, 287
Born-Oppenheimer approximation, 160
Bose-Einstein condensation, 259-270
Bosons, 209, 250
Boundary conditions, Dirichlet or Neumann,
212
Branch point, classical action, 224
Brillouin zone, 292, 297, 301
Broken line path, 7, 8, 29
Brownian motion, 28, 53-64, 342
exercise, 63
paths, 263, 264, 319, 320

Calculus of variations, 79-91
Calogero model, 346
Canonical form, 136, 139
Canonical morphology, 140
Canonical operators, 216
Canonical quantization,
215

- Canonical transformation, 20, 306, 307, 308, 309, 311
- Catastrophe:
dimensions of, 140
elementary, 140, 141
theory, 120-130, 139, 140, 142
- Caustic, fourth-order, 128
- Caustics, 17, 118-142, 168
- Caustic surfaces, 87
- Chapman-Kolmogorov equation, 42, 47
- Chemical potential, 250, 260
- Classical evolution, exercise, 19
- Classical limit, 16, 51. *See also* Correspondence limit; Semiclassical approximation, WKB approximation
- Classical path, 39-41, 159
coalescing, 125
- Coherent state representation, 242-256
- Col, 277
- Collective behavior, 290. *See also* Phase transitions
- Commutation relations, 49-52, 255
- Complete integrability, 313
- Complex paths, 162, 163, 168
- Complex potentials, 163
- Complex scaling, 99, 105
- Computer, 341
- Conjugate point, 84-90, 118, 122, 124. *See also* Focal point
- Continuity of paths in Wiener measure, 318, 319
- Continuous deformation, 197
- Continuous spectrum, 43
- Contractible, 198. *See also* Simply connected
- Convex function, 174
- Convex obstacle, 339
- Correlation function, two-point, 291, 299, 302
- Correlation length, 291, 298
- Correspondence:
limit, 16-21
principle, 154
identities, 157
- Covering projection, 202, 203, 206
- Covering space, 203
universal, 202, 203, 206
universal SO(3), 204
universal torus, 204
- Creep wave, 40, 339
- Critical droplet, 271, 278, 280, 288
explicit solution, 274
- Critical exponents, 290, 291, 296, 302
- Critical phenomena, 293, 295, 297, 299
- Critical point, 126, 136, 281, 290, 291, 298, 299, 301
coalescing, 125, 126, 132, 135, 136, 141
- Critical region, 121
- Curie point, 291
- Curie-Weiss model, 325
- Curvature, 217, 220
phase factor, 216, 217
- Curved space path integral, 41, 214-224
- Curved spacetimes, 229
- Cusp, 127, 128
- Cylinder sets, 57
- Cylindrical mirror, 125
- Decay rate, 273
- Decimation, 302
- Deficiency indices, 211, 212
- Density:
of paths, 88
of states, 268
- Deviations, large, 266
- Difference differential equations, 174
- Diffracting body, 212
- Diffraction, 168, 212, 339
geometrical theory of, 157, 163
- Diffusion, 54, 55, 60
- Dilute gas of excitations, 282
- Dirac equation, 185, 229
- Drift velocity, 55
- Effective field approximation, 98
- Effective mass, 170
- Ehrenfest theorem, 51
- EIGHT, 202
- Elastic forces, 332
- Electric field, uniform, 39
- Electrodynamics, 174
time symmetric, 323
- Electromagnetic field, 320-323
quantized, 324
- Electron in a solid, 267
- Emission by black holes, 235
- Emission probability from black hole, 234
- Energy levels, 44-48, 147-157. *See also* Feynman-Kac formula
- Energy, imaginary part of, 273

- Entropy, 332
 Enveloping ray, 339
 Equations of motion, complex, 250
 Equivalence class, 198, 200
 Error function, 76
 Essential singularity, 98, 107, 271, 283, 284, 289
 Euler-Lagrange equation, finite difference approximation, 109
 Event horizon, 229
 Exchange operator, 209
 Exchange second virial coefficient, 338
 Expectation value, complex, 49, 50
 Exponential potential, exercise, 89, 105
 Extensions of symmetric operators, 212
- Fermat's principle, 168
 Fermions, 209, 346
 Fermi statistics, 256
 Ferromagnet, 325
 Feynman diagrams, 65-69, 324
 Feynman-Kac formula, 43-48, 171, 173, 178, 272, 273, 330, 339
 Fields, rough, 320
 Field theory, 242, 251, 315, 316, 317, 320
 Fifth parameter, 214, 226, 230
 First passage time, 60-64, 265
 First-order phase transitions, *see* Phase transitions
 Fixed point, 291, 298, 299, 302
 Fluctuation, destabilizing, 277
 Focal point, 84-90. *See also* Conjugate point
 Focusing, 105, 118, 141
 Fokker-Planck equation, 310, 342, 343
 Force field, uniform, 32, 38
 Fourier coefficients, asymptotic properties of, 319
 Free energy, 272, 317, 326, 329
 Free particle, 31, 32, 45 spectrum, 45
 Fugacity, 260
 Functional calculus, 49-52
 Functional form, topological, 132, 133
 Functions of bounded variation, 319
 Fundamental group, 201, 202
 $\text{SO}(3)$, 221
 Fundamental homotopy group, 208, 210
 Fundamental theorem for theta functions, 223
- Gamma function, 75
 Gamma matrices, 229
 Gauge fields, 210
 Gauge invariance, exercise, 30
 Gauge transformation, 27-30, 210, 211, 213
 Gaussian functional integral, 269
 Generating function, 243
 Geodesic deviation, 87
 Geometrical optics, 164-169
 Gibbs distribution, 341, 342, 343
 Global invariants, 313
 Glory (in scattering), 163
 Goldstone boson, 84
 Gravitational field, quantization, 235
 Green's function, *see* Propagator
 Ground state energy, 342
 Group manifolds, 41
- Hamiltonian, time dependent, 8
 Hamilton-Jacobi equation, 15, 88, 101, 155, 224, 307
 Hard sphere, 334, 336 gas, 333, 338
 Harmonic oscillator, 32-39, 44 energy levels, 46 forced, 38, 172, 322 propagator, 247 wave function, 46
 Heat equation, 55, 60, 117
 Heavy ion scattering, 158, 163
 Hermite polynomials, 243, 244
 Higher-order terms, Laplace method, 115
 Hilbert space:
 of analytic functions, 245
 of paths, 63
 Homogeneous polynomials, 246
 Homogeneous spaces, 41
 Homotopy, 187
 class, 198
 equivalence relation, 197
 equivalence classes, 199, 206
 factor, 211
 theorem (for path integrals), 192, 208, 210
 theory, 197, 205, 209
 type, 199
 Hydrogen atom, 157
- Ideal bose gas, 252
 Ideal gas, 237, 240

- Images, method of, 40, 41, 61, 62, 156
 Implicit function theorem, 129, 133, 140
 Importance sampling, 339
 Impurities, 260, 262, 269
 Impurity potential, 267
 Index:
 of bilinear functional, 89
 of refraction, 164
 Indistinguishable particles, 209
 Infinitesimal generator, 9
 Infrared bounds, 270
 Instantaneous Coulomb interaction, 321
 Instanton, 122, 271-289, 331
 analytically continued, 330
 explicit solution, 274
 Integral:
 Gaussian, 6
 Gaussian, table, 20, 21
 Gaussian, multidimensional, 34
 Interacting field, 172
 Interaction:
 nonlocal, 269
 representation, 184
 Interference, 13, 18, 168
 Internal coordinate, 186, 187
 Intrinsic curvature, 187
 Invariant distance, 219
 Invariants, globally defined, 310
 Ising model, 328
 stochastic dynamics, 341
 Isotropy group, 186
 Ito formula, 28
 integral, 27-30, 215
 Jacobi equation, 37, 84, 95, 284
 field, 84, 85, 95
 matrix, 110, 328
 theta function, 156, 193, 195, 222, 223.
 See also Theta function
 Jensen's inequality, 174, 181
 Kinetic energy, in path integral, 51, 52
 Klein paradox, 228
 Klein-Gordon equation, 225, 226
 Kronecker's theorem, 195, 196
 Kruskal coordinates, 231, 234, 235
 Lamb shift, 322
 Landau-Ginzburg theory, 253, 254
 Laplace-Beltrami operator, 205, 215
 Laplace's method, 20, 76, 96, 114, 254,
 273, 276, 294, 295, 301, 317, 320, 326,
 330
 Laplacian, 205, 215
 Lebesgue measure, 56, 57
 Legendre condition, 80
 transform, 148
 Lie groups, 41
 Lifting (of Laplacian), 206
 Linked cluster expansion, 253
 Lippmann-Schwinger equation, 60, 286
 Lipschitz condition, 319
 Localization, 13
 Localized moments, 332
 Macroscopic occupation, 261, 262, 263,
 264, 266
 Magnetic field, 22-26
 uniform, 39
 varying, 185
 Magnetic moments, localized, 188
 Magnetic monopole, 212
 Magnetic resonance, 185
 Magnetization, 326
 Malgrange preparation theorem, 140
 Mass renormalization, 170
 Mass spectrum, 187
 Matrix product, infinite, 184
 Mean field theory, 254, 326, 330, 332
 Measure theory, 56
 Medium range interaction, 329
 Metastability, 278, 280, 289, 331, 343
 Metastable states, 271
 Method:
 of images, *see* Images, method of
 of stationary phase, *see* Stationary phase
 approximation
 of steepest descents, *see* Steepest descents,
 method of
 Midpoint evaluation in path integral, 23
 Minimum uncertainty wave packets, 244
 Mirror, cylindrical, 125
 Momentum operator, 51
 Momentum states, 6
 Monte Carlo methods, 339, 340, 343
 Morphology, 128
 Morse index theorem, 89-90
 Morse theory, 89, 91, 143-146
 Multiply connected spaces, 190,
 211

- Natural metric, induced, 187
 Negative energy state, 226, 228, 231
 Nonclassical effects, 124
 Nonintegrable phase factors, 210
 Normal form, 131-132, 136, 139, 140
 Numerical sum over paths, 343
- One loop approximation, 236
 Operator:
 annihilation (lowering), 243
 creation (raising), 243
 extensions of, 211
 ordering, 52, 214, 215, 310, 311, 313
 Optical modes, 170
 Optics, 124, 141, 164-169
 Orbit, 186
 Order parameter, 254
 complex, 253
 Ornstein-Uhlenbeck process, 342, 343
 O-symbol, 73
 Outgoing wave boundary conditions, 280
 Overcomplete basis, 245
- Pair production, 230
 Parabolic barrier, 163
 Parabolic cylinder function, 74, 78, 97
 Parastatistics, 209, 210
 Particle-antiparticle pair, 230
 Particles moving backward in time, 226
 Partition function, 237, 251, 252, 254, 262,
 275, 292, 316, 317, 326, 328
 grand canonical, 250, 260
 Path, nowhere differentiable continuous,
 28
 Path integral:
 circle, 208
 coherent state representation, 314
 complex paths, 155
 curved space, 223
 definition, 7
 Dirac equation, 188
 electromagnetic wave equation, 164
 fermion, 256
 Fourier transform of, 158
 free particle, 31, 32
 harmonic oscillator, 32-39
 phase space, 242, 249, 303-314
 relativistic spinning particles, 235
 Schwarzschild metric, 233
 spin, 182-189, 218, 223
- vector potential, 22-26, 29
 Percolation, 284, 289
 Periodic lattice, 190, 195, 196
 Perturbation expansion, 98,
 344
 high order terms, 288
 Perturbation theory, 65-69, 97, 100, 103,
 287
- Phase shift, 162
 Phase space path integral, 154, 303-314
 Phase transition, 259-270, 284, 289, 290-
 302, 328, 329
 first-order, 270, 283, 327
 Landau theory, 316
 in presence of impurities, 261
 second order, 290, 327
- Phonon, 170, 181
 field, 321
 spectrum, 90, 91
 Photon, 321, 322
 Planck length, 236
 Poisson summation formula, 194
 Polar coordinates, 345, 346
 Polar crystal, 170
 Polaron, 170-181, 321
 ground state energy, 180
 Polygonal paths, 63, 64
 Polymer, 332, 333
 Polynomial, homogeneous, 246
 Positive definiteness of $\delta^2 S$, 82
 Positive frequency functions, Hilbert space
 of, 228
 Potential theory, 53
 Potential:
 Coulomb, 48, 154, 157, 344
 delta-function, 48, 162, 344
 distribution function for, 267
 exponential, 89, 105
 harmonic, 32-39
 infinite square well, 46, 63, 156
 inverse square, 48
 nonlocal in time, 174
 random, 267, 270
 square well, 161
 Preparation theorems, 139
 Prequantising, 210
 Probabilities, conditional, 13
 Probability amplitude, 13
 Prodistribution (pseudomeasure), 63, 64
 Product integral, 182-185, 235, 331

- Propagation, *see* Path Integral
 Propagator, 7
 relativistic, 225-236
 on SU(2), 220
 Pseudomeasures (prodistribution), 63, 64
 Quadratic Lagrangian, 32-39, 92, 94, 110,
 144, 172
 Quantization, curved spaces, 214-224
 Quantum chromodynamics, 288
 Quantum effects, 93
 Quantum electrodynamics, 320, 324
 Quantum field theory, 325, 330
 Quantum gravity, 236
 Quantum statistics, 254
 Quantum tunneling, 229
 Radiative correction, 322
 Rainbow (in scattering), 163
 Random walk, 53-64, 332
 Ray density, 168
 Reflecting obstacle, 40, 41
 Relative homotopy, 198
 Relativistic propagators, 225-236
 Renewal equation, 177
 Renormalization, 293, 295, 297, 299
 group, 295, 301, 302
 theory, 302
 transformation, 297, 298
 Representation:
 of fundamental group, 208
 of SU(2), 221, 246
 Resolvent set, 11
 Riemann surface, 224
 Riemannian metric, 214
 Rotation group, 205
 group manifold, 202
 Rotation matrices, 187
 Rotator, fixed axis, 190
 Rubber, 332, 333
 Ruse's invariant, 223
 Saddle point, 277
 Scale:
 invariance, 298
 transformation, 291, 301, 302
 Scaling, 293, 295, 297, 299
 coordinate space, 302
 Scattering phase shift, 161
 Scattering, off mass shell, 228
 Scattering theory, 141, 158-163
 Schwarzschild metric, 230-231
 Second variation, 276
 Second virial coefficient, 333, 334
 Second-order phase transitions, *see* Phase
 transitions
 Self adjoint, 211
 Semiclassical approximation, 96-117, 142,
 143-146, 158, 159, 162, 308. *See also*
 WKB approximation
 Semigroup, contraction, 9
 Shadow region, 124
 Short distance cutoff, 292
 Simply connected, 187, 201. *See also* Con-
 tractible
 Singular perturbation, 346
 Singularity, unfolding of, 139-141
 Small oscillation integral, 275
 Small oscillations about the instanton, 284
 S-matrix, 158, 162
 classical, 163
 Solitons, 310, 313
 Spaces:
 homogeneous, 41
 multiply connected, 197
 Spectrum, 11
 Spherical model, 291, 294, 302
 Spin, 182-189
 continuous models for, 185
 half integral, 221, 222
 integral, 221
 model, 187, 327, 328
 model, continuous, 235
 model, relativistic path integral, 235
 Stationary phase approximation, 16-18, 76,
 97, 109, 113, 141, 147, 162
 Statistical mechanics, 237-241
 Steepest descents, method of, 125, 281
 Stirling's approximation, 54
 Stochastic process, 339, 341, 343
 String (magnetic monopole), 213
 Strong minimum, 90
 Structural stability, 128
 Sturm-Liouville problem, 81
 discrete form, 111
 Sum over classical paths, 223
 exactness, 39
 Sum over histories, 7
 Superfluidity, 250, 253, 259, 346
 Symmetric group on N objects (S_N), 209

- Thermal gravitons, 236
Thermodynamic limit, 264
Thermodynamics:
 black hole, 235
 second law of, 229
Theta function, 193, 195, 224
 fundamental identity, 194, 195. *See also*
 Jacobi theta function
Three-body problem, 345, 346
Time ordered product, 8, 182, 183, 331
Top:
 spherical, 187
 symmetrical, 187
Topological constraint, 332, 333
Topological equivalence, 140
Trajectories in phase space, 305
Transition element, 173
Translation invariance, 278, 280
Trotter product formula, 5, 9, 117, 211,
 303, 333
Tunneling, 160, 162. *See also* Barrier penetration
Turning point, 101, 149, 150
Two-fluid phenomenology, 254
Two-point function, 301
Two slit experiment, 13-15

Ultraviolet divergences, 320
Uncompleting the square, 325-332
Universality, 290, 301
Unstable states,
 271

Van Hove singularities, 90
Van Vleck determinant, 88, 95, 96, 106, 217
 308
Variational principle of Maupertuis, 167
Vector potential, 22-26, 212
Virial coefficient, 333, 334

Wave function, coherent state, 254
Wave packet, minimum uncertainty, 19
Wave packets, 160
Weierstrass preparation theorem, 139
Weierstrass E-function, 90, 91
Wiener integral, 53-64, 96, 114, 117, 129,
 262, 317, 318, 320, 332, 335, 343
Wiener measure, 53-64, 263, 272
 unconditional, 60
Wiener paths, *see* Browning motion paths
Wiener-Sausage, 264, 269
Wigner-Kirkwood expansion, 239, 240, 241
Winding number, 190, 191, 192, 193, 195,
 202, 208, 332
WKB approximation, 15, 16-21, 88, 92-107,
 118, 147-157, 158, 159, 162, 224, 273,
 276. *See also* Semiclassical approximation
WKB barrier penetration, *see* Barrier penetration
Wronskian, 285

Zero energy mode, 280, 282
Zero mode problem, 313
Zero point energy, 173

Supplements

Path integration now enjoys a central role in many areas of physics and chemistry. Its impact has been felt in other fields as well. For example, in seismology the path integral provides a way to deal with partial differential equations, while for finance it is a tool for the analysis of stochastic processes.

With the issuing of this Dover Edition, I am adding a few new topics. They are presented in Sec. I of this supplement. While some are elementary and hopefully of broad interest, I am not providing a comprehensive survey of developments, as might be suitable for a bona fide second edition (which this is not). The coverage ranges from self-contained to telegraphic. Sec. II lists recent texts where a fuller picture can be obtained. Finally, in Sec. III are errata and comments on particular parts of this book. Partly these are factors of two and such (which eluded me, despite the opportunity to correct them on each of Wiley's dozen or so reprintings), but mainly they are comments that were too lengthy to fit on the reprinted page.

References to "Sections" 1 through 32 are to sections of this book (a.k.a. chapters). Similarly, equation numbers with periods, (*n*.*m*), refer to equations in the book, while those consisting of a number alone refer to the supplement. Sections of the supplement have labels beginning with Roman numerals.

In preparing this supplement I have had the help of many individuals. Moreover, since the publication of the book, quite a few people have come to me with corrections, some of which were incorporated in previous reprintings and some of which appear here. Since these events stretch over 20 years, I hope those whose contributions have slipped my mind will forgive me. Those I recall—and thank, whether for counsel, corrections or both—are A. Auerbach, Y. Avron, C. DeWitt, P. Exner, P. Facci, B. Gaveau, H. Grabert, T. Jacobson, H. Jauslin, G. Junker, M. Kac, S. Kivelson, A. Mann, D. McLaughlin, E. Mihokova, D. Mozyrsky, D. Mugnai, S. Pascazio, P. Pechukas, A. Ranfagni, M. Revzen, G. Roepstorff, M. Roncadelli, A. Scardicchio, L. J. Schulman, D. Tolkunov, U. Weiss, and N. Yamada.

Contents

I. Topical supplements	363
A. Path integral in a magnetic field using the Trotter product formula	363
1. Splitting a sum	363
2. Splitting a product	365
3. A more accurate product formula	365
4. Precision and rough paths	367
Notes	368
B. Path decomposition expansion	369
1. The formula	369
2. Proof of the path decomposition expansion	370
3. More than one dimension	374
Notes	375
C. Checkerboard path integral	376
Notes	380
D. Exact solutions	381
1. δ -function path integral	381
2. Half-plane barrier	382
3. Coulomb potential and related solutions	387
Notes	387
E. Dissipation and other forced-oscillator applications	388
F. Mathematical developments	391
G. Electromagnetism and other PDE's	391
H. Homotopy	392
I. And more ...	393
1. Variational corrections to the classical partition function	393
2. A particular numerical approach	393
3. Semiclassical approximations for coherent states	394
4. Chaos and order	394
II. Literature	396
A. Texts and review articles	396
B. Conference proceedings	398
III. Page by page comments and errata	399
References and endnotes for the supplements	407

I. TOPICAL SUPPLEMENTS

A. Path integral in a magnetic field using the Trotter product formula

In Sec. 1 the path integral is derived using the Trotter product formula. This is basically the statement

$$\left| e^{-i\epsilon H/\hbar} - e^{-i\epsilon p^2/2m\hbar} e^{-i\epsilon V(\mathbf{x})/\hbar} \right| < O(\epsilon), \quad (1)$$

where standard notation, $H = p^2/2m + V$, etc., is used. Later, in Sec. 4, I turn to a more delicate derivation, the path integral in the presence of a magnetic field. This involves a term $\dot{\mathbf{x}} \cdot \mathbf{A}$ in the classical Lagrangian and, as explained in Sec. 5, to the need to evaluate $\mathbf{A}(\mathbf{x})$ at the *midpoints* along the broken line path within a path integral.

The derivation in Sec. 4 follows the original paper of Feynman. However, by a slight modification of the method of Sec. 1, it is possible to use the Trotter formula. The yet more delicate case of a curved-space metric can also be treated by operator methods.

1. Splitting a sum

As usual, we want the propagator, G , the kernel of the operator $\exp(-iHt/\hbar)$. With a magnetic field

$$H = \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A}(\mathbf{x}) \right)^2 + V(\mathbf{x}), \quad (2)$$

where \mathbf{A} is the vector potential. For $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, G satisfies the following sequence of identities

$$G(\mathbf{b}, t; \mathbf{a}) = \langle \mathbf{b} | \exp(-iHt/\hbar) | \mathbf{a} \rangle = \langle \mathbf{b} | [\exp(-iHt/N\hbar)]^N | \mathbf{a} \rangle \quad (3)$$

$$= \langle \mathbf{b} | e^{-iH\epsilon/\hbar} \int d^3x_1 | \mathbf{x}_1 \rangle \langle \mathbf{x}_1 | e^{-iH\epsilon/\hbar} \int d^3x_2 | \mathbf{x}_2 \rangle \langle \mathbf{x}_2 | e^{-iH\epsilon/\hbar} \\ \times \cdots \times \int d^3x_{N-1} | \mathbf{x}_{N-1} \rangle \langle \mathbf{x}_{N-1} | e^{-iH\epsilon/\hbar} | \mathbf{a} \rangle, \quad (4)$$

where after line (3) I use the definition $\epsilon \equiv t/N$. Eq. (3-4) can be written concisely as

$$G(\mathbf{b}, t; \mathbf{a}) = \int \prod_{k=1}^{N-1} d^3x_k \prod_{\ell=0}^{N-1} G(\mathbf{x}_{\ell+1}, \epsilon; \mathbf{x}_\ell), \quad (5)$$

with $\mathbf{x}_N = \mathbf{b}$, and $\mathbf{x}_0 = \mathbf{a}$. This is the starting point for the Sec. 1 derivation. The smallness of ϵ allowed H to be split into kinetic and potential energy terms. For convergence one must maintain $O(\epsilon)$ accuracy. That is, $G(\mathbf{x}_{\ell+1}, \epsilon; \mathbf{x}_\ell)$ can be replaced by other functions that differ only by terms going to zero faster than ϵ . For numbers this can be seen by recalling that $[1 + (x + a_N)/N]^N \rightarrow e^x$ provided $a_N \rightarrow 0$. For operators this is a more nuanced project and the reader is referred to Sec. 1 and its Notes.

The goal then is to approximate $G(\mathbf{x}, \epsilon; \mathbf{y})$ to first order in ϵ . It is helpful to phrase this in operator language. For operators A and B ,

$$\exp[\lambda(A + B)] = \exp(\lambda A) \exp(\lambda B) \exp\left(\frac{\lambda^2}{2}[B, A] + O(\lambda^3)\right). \quad (6)$$

The propagator is of this form, with $\lambda = \epsilon$, $A = -iK/\hbar$, $B = -iV/\hbar$, and K the kinetic energy (called “ T ” in Sec. 1). For present purposes it is sufficient to know that

$$\exp[\lambda(A + B)] = \exp(\lambda A) \exp(\lambda B) + O(\lambda^2). \quad (7)$$

In Sec. 1, without the vector potential, this quickly led to the path integral. To prepare for the vector potential case, here’s a recap.

Focus on one factor in the integrand of (5). Using Eq. (7) and the fact that V is diagonal in position we obtain

$$G(\mathbf{x}, \epsilon; \mathbf{y}) = \langle \mathbf{x} | \exp[-iK\epsilon/\hbar] | \mathbf{y} \rangle \exp[-iV(\mathbf{y})\epsilon/\hbar] + O(\epsilon^2). \quad (8)$$

Let $|\mathbf{p}'\rangle$ be an eigenvector of \mathbf{p} with eigenvalue \mathbf{p}' . Inserting $\mathbf{1} = \int d^3\mathbf{p}' |\mathbf{p}'\rangle\langle\mathbf{p}'|$ to the left of $|\mathbf{y}\rangle$ yields

$$\int d^3\mathbf{p}' \langle \mathbf{x} | e^{-i\mathbf{p}'^2\epsilon/\hbar 2m} | \mathbf{p}' \rangle \langle \mathbf{p}' | \mathbf{y} \rangle = \left[\frac{m}{2\pi i \hbar \epsilon} \right]^{3/2} \exp\left(\frac{i}{\hbar} \frac{m(\mathbf{x} - \mathbf{y})^2}{2\epsilon} \right). \quad (9)$$

[Eq. (16) gives $\langle \mathbf{p} | \mathbf{x} \rangle$ explicitly.] Combining Eqs. (8) and (9) leads to the expression, which when iterated, gives the path integral in three dimensions:

$$G(\mathbf{x}, \epsilon; \mathbf{y}) = \left(\frac{m}{2\pi i \hbar \epsilon} \right)^{3/2} \exp\left[\frac{i}{\hbar} \left(\frac{m(\mathbf{x} - \mathbf{y})^2}{2\epsilon} - \epsilon V(\mathbf{y}) \right) \right] + O(\epsilon^2) \quad (10)$$

[cf. Eq. (1.24)]. Inserting Eq. (10) in Eq. (5), we obtain the classical action in the exponent. Had we interchanged K and V in Eq. (7), the argument of V in Eq. (10) would be \mathbf{x} rather than \mathbf{y} . This changes the short-time propagator by less than $O(\epsilon)$ and therefore does not change the final result.

2. Splitting a product

When a magnetic field is present, $K = [\mathbf{p} - e\mathbf{A}(\mathbf{x})/c]^2/2m$. Inserting a momentum-state resolution of the identity, as in Eq. (9), is inadequate, because \mathbf{A} is a function of \mathbf{x} . The way around this is to look at the *square root* of K , which because it is a sum of \mathbf{p} and \mathbf{A} can be resolved by separately expressing \mathbf{p} and \mathbf{A} in momentum and position space bases. Taking the square root in the exponent is accomplished by introducing an additional integral: uncompleting the square, as in Sec. 32.5. We extend Eq. (32.29) by going to three dimensions and taking multiples of the variables used there, to obtain the identity

$$\exp\left(-i\epsilon \frac{\mathbf{b}^2}{2m\hbar}\right) = \left(\frac{1}{\sqrt{2\pi i}}\right)^3 \int d^3\mathbf{u} \exp\left(i\frac{\mathbf{u}^2}{2} - i\sqrt{\frac{\epsilon}{m\hbar}} \mathbf{b} \cdot \mathbf{u}\right). \quad (11)$$

For convenience, let $\mathbf{a}(\mathbf{x}) \equiv e\mathbf{A}(\mathbf{x})/c$ and $\bar{\epsilon} \equiv \sqrt{\epsilon/m\hbar}$. We want to evaluate Eq. (8) and concentrate on the kinetic energy part, designated G_K . Setting $\mathbf{b} = \mathbf{p} - \mathbf{a}$ in Eq. (11), this becomes

$$\begin{aligned} G_K(\mathbf{x}, \epsilon; \mathbf{y}) &\equiv \langle \mathbf{x} | e^{-iK\epsilon/\hbar} | \mathbf{y} \rangle = \langle \mathbf{x} | \exp\left(-i\epsilon \frac{(\mathbf{p} - \mathbf{a})^2}{2m\hbar}\right) | \mathbf{y} \rangle \\ &= (2\pi i)^{-3/2} \int d^3\mathbf{u} e^{i\mathbf{u}^2/2} \langle \mathbf{x} | \exp(-i\bar{\epsilon}(\mathbf{p} - \mathbf{a}) \cdot \mathbf{u}) | \mathbf{y} \rangle. \end{aligned} \quad (12)$$

In Eq. (12), focus on $\exp(-i\bar{\epsilon}(\mathbf{p} - \mathbf{a}) \cdot \mathbf{u})$. Because the noncommuting \mathbf{p} and $\mathbf{a}(\mathbf{x})$ no longer appear as a product one might expect that this would allow factoring of the exponential, as in the Trotter product formula. Unfortunately, applying Eq. (7) would *not* provide the needed accuracy, because it is $\bar{\epsilon}$ that appears in Eq. (12), not ϵ . The error is the *square* of $\bar{\epsilon}$, namely $\epsilon/m\hbar$, which cannot be neglected. This problem occurs whenever an expression mixes non-commuting variables, for example in a curved-space metric (or position-dependent kinetic energy). Generally speaking, this is why path integration does not fix operator ordering. The way to deal with this is to improve on Eq. (7).

3. A more accurate product formula

A variation of Eq. (7) provides $O(\lambda^3)$ accuracy:

$$\exp[\lambda(A + B)] = \exp(\lambda B/2) \exp(\lambda A) \exp(\lambda B/2) + O(\lambda^3). \quad (13)$$

This can be checked by direct expansion of the exponentials [1]. Apply Eq. (13) to $\langle \mathbf{x} | \exp(-i\bar{\epsilon}(\mathbf{p} - \mathbf{a}) \cdot \mathbf{u}) | \mathbf{y} \rangle$ to obtain

$$\begin{aligned}\langle \mathbf{x} | \exp(-i\bar{\epsilon}(\mathbf{p} - \mathbf{a}) \cdot \mathbf{u}) | \mathbf{y} \rangle &= \langle \mathbf{x} | e^{i\bar{\epsilon}\mathbf{a} \cdot \mathbf{u}/2} e^{-i\bar{\epsilon}\mathbf{p} \cdot \mathbf{u}} e^{i\bar{\epsilon}\mathbf{a} \cdot \mathbf{u}/2} | \mathbf{y} \rangle + O(\epsilon^{3/2}) \\ &= e^{i\bar{\epsilon}\mathbf{a}(\mathbf{x}) \cdot \mathbf{u}/2} \langle \mathbf{x} | e^{-i\bar{\epsilon}\mathbf{p} \cdot \mathbf{u}} | \mathbf{y} \rangle e^{i\bar{\epsilon}\mathbf{a}(\mathbf{y}) \cdot \mathbf{u}/2} + O(\epsilon^{3/2}).\end{aligned}\quad (14)$$

Errors are $O(\epsilon^{3/2})$, larger than ϵ^2 , but small enough. The operators \mathbf{a} , functions of the position operator, appear symmetrically, and are immediately evaluated on the last line as $\mathbf{a}(\mathbf{y})$ and $\mathbf{a}(\mathbf{x})$. *This is the step that gives the midpoint rule.* For the momentum-dependent portion, a momentum resolution of the identity is inserted:

$$\langle \mathbf{x} | \exp(-i\bar{\epsilon}\mathbf{p} \cdot \mathbf{u}) | \mathbf{y} \rangle = \int d^3 p' \langle \mathbf{x} | \exp(-i\bar{\epsilon}\mathbf{p} \cdot \mathbf{u}) | \mathbf{p}' \rangle \langle \mathbf{p}' | \mathbf{y} \rangle. \quad (15)$$

With

$$\langle \mathbf{x} | \mathbf{p}' \rangle = (2\pi\hbar)^{-3/2} \exp(i\mathbf{p}' \cdot \mathbf{x}/\hbar), \quad (16)$$

Eq. (15) becomes

$$\begin{aligned}\langle \mathbf{x} | \exp(-i\bar{\epsilon}\mathbf{p} \cdot \mathbf{u}) | \mathbf{y} \rangle &= \left(\frac{1}{2\pi\hbar} \right)^3 \int d^3 p' \exp[-i\bar{\epsilon}\mathbf{p}' \cdot \mathbf{u} + i\mathbf{p}' \cdot (\mathbf{x} - \mathbf{y})/\hbar] \\ &= \frac{1}{\hbar^3} \delta^3 \left(-\bar{\epsilon}\mathbf{u} + \frac{\mathbf{x} - \mathbf{y}}{\hbar} \right) = \frac{1}{(\bar{\epsilon}\hbar)^3} \delta^3 \left(-\mathbf{u} + \frac{\mathbf{x} - \mathbf{y}}{\hbar\bar{\epsilon}} \right).\end{aligned}\quad (17)$$

The integral over \mathbf{u} in Eq. (12) is now trivial, yielding

$$\begin{aligned}G_K(\mathbf{x}, \epsilon; \mathbf{y}) &= \left[\frac{1}{2\pi i} \frac{m}{\epsilon\hbar} \right]^{\frac{3}{2}} \int d^3 u e^{iu^2/2} \exp \left[i\bar{\epsilon}u \cdot \frac{\mathbf{a}(\mathbf{x}) + \mathbf{a}(\mathbf{y})}{2} \right] \delta^3 \left(\frac{\mathbf{x} - \mathbf{y}}{\bar{\epsilon}\hbar} - \mathbf{u} \right) \\ &= \left[\frac{m}{2\pi i\hbar\bar{\epsilon}} \right]^{\frac{3}{2}} \exp \left[\frac{i}{\hbar} \frac{m(\mathbf{x} - \mathbf{y})^2}{2\epsilon} \right] \exp \left[\frac{i}{\hbar} (\mathbf{x} - \mathbf{y}) \cdot \frac{\mathbf{a}(\mathbf{x}) + \mathbf{a}(\mathbf{y})}{2} \right],\end{aligned}\quad (18)$$

where in the last line the original ϵ has replaced $\bar{\epsilon}$. Recall that $\mathbf{a} \equiv e\mathbf{A}/c$. For the classical Lagrangian, a magnetic field contributes $\mathbf{v} \cdot e\mathbf{A}(\mathbf{x})/c$. If we multiply and divide by ϵ in the last expression in Eq. (18), we obtain

exactly the appropriate term. For completeness, we restore \mathbf{A} and V , yielding,

$$\begin{aligned} G(\mathbf{x}, \epsilon; \mathbf{y}) = & \left[\frac{m}{2\pi i \hbar \epsilon} \right]^{\frac{3}{2}} \exp \left\{ \frac{i}{\hbar} \epsilon \left[\frac{m}{2} \frac{(\mathbf{x} - \mathbf{y})^2}{\epsilon^2} \right. \right. \\ & + \frac{(\mathbf{x} - \mathbf{y})}{\epsilon} \cdot \frac{e}{c} \left(\frac{\mathbf{A}(\mathbf{x}) + \mathbf{A}(\mathbf{y})}{2} \right) - V(\mathbf{y}) \left. \right] \left. \right\} \\ & + O(\epsilon^{3/2}). \end{aligned} \quad (19)$$

The argument of the exponent is seen to be $i\epsilon/\hbar$ times the classical Lagrangian.

Note that what naturally arises is the average of the vector potential, \mathbf{A} , at the endpoints of the broken line path. The “midpoint” way of attaining the same level of precision uses $\mathbf{A}[(\mathbf{x} + \mathbf{y})/2]$. The difference between these is essentially a second derivative of \mathbf{A} times $(\mathbf{x} - \mathbf{y})^2$. The latter is of order ϵ and in turn multiplies an additional power of $(\mathbf{x} - \mathbf{y})$, so that the difference, of order $\epsilon^{3/2}$, can be neglected.

4. Precision and rough paths

The need for precision in “sensitive” path integrals has been emphasized in this book and in many other places. It was known to Feynman and—for the Wiener integral—is what lies behind Ito’s theorem, discussed in Sec. 5. This is not a quirk of the path integral, but is a central feature of quantum mechanics. It is evident in the fact that *velocity* cannot be defined as the limit of $\Delta \mathbf{x}/\Delta t$. Moreover, the same property plays a central role in applications of the Wiener integral, for example, in the derivation of the Black-Scholes formula.

Integration of terms in which the “kinetic energy” is space-dependent is even more sensitive than integration of $\int \mathbf{A} \cdot \dot{\mathbf{x}} dt$. The reason is easy to see. For the vector potential, \mathbf{A} multiplies $\Delta \mathbf{x}$, which is $O(\sqrt{\epsilon})$. On (e.g.) curved spaces, the metric function g multiplies $(\Delta \mathbf{x})^2/\Delta t$, which is $O(1)$, hence yet more stringent demands arise for the evaluation of g . Such dynamics can also be quantized operator methods.

The degree of roughness of the paths, reflected in the nonexistence of the derivatives, can be expressed in other ways. Like the paths that dominate Wiener measure, those that are important for the path integral can be assigned a fractal dimension (which is 2).

Exercise: *Another way to derive the midpoint rule.* Simplify notation by setting $a = eA/c$, $\hbar = 1$ and $m = 1$. You want to eval-

uate $\langle x|e^{-i\epsilon(p-a)^2/2}|y\rangle$. You can write this to order ϵ as $\langle x|e^{-i\epsilon p^2/2}[1 - i\epsilon(pa + ap)/2]e^{-i\epsilon a^2/2}|y\rangle$. In this expression it's obvious that one takes a 's argument symmetrically. What's not obvious is that you have to keep it that way, and why. Also in this formulation effort is needed to show that the a^2 disappears.

Notes

Feynman's derivation is in his Rev. Mod. Phys. paper [2]. Here I follow [3], which is in turn based on [4] and [5]. The Trotter formula variation of Eq. (13) has also been used in numerical evaluations of path integrals by Janke and Sauer [6]. The curved space path integral is derived in [4] by operator methods, although I expect that an extended Trotter expansion could be made to work in this case too.

Ref. [4] works with the path integral in a form not usually employed by physicists. The mathematician talking about the path integral will often call it the Feynman-Kac formula (not exactly the usage of Sec. 7) and write down something that looks little like our familiar sum over paths. In fact this form does appear in this book, in the " E_{xt} " of Eq. (9.17). In practice the expectation E_{xt} is essentially the expectation over Wiener measure. Usually Wiener measure depends on t alone, but E_{xt} can be recovered by inserting a δ -function to enforce the endpoint condition x . These remarks may be too telegraphic to make any sense. At a 1987 Trieste meeting/workshop I gave a path integral derivation in this language [7], although it may be a challenge to track down the notes.

If for some reason one prefers not to use the midpoint rule for a curved-space path integral, a correct propagator can still be obtained by adding a potential. This is discussed in [8].

The work of Nelson has already been mentioned in Sec. 1. The Trotter product formula continues to attract the attention of mathematicians; see for example Ichinose [9].

The Black-Scholes formula is a method for assigning a price to a financial option. The book of Hull [10] derives this formula using the Ito integral. But don't expect "Black-Scholes" plus your knowledge of path integration to make you rich, not immediately, anyway. Markets are *not* well described by Gaussian statistics and are closer to being Levy-distributed, for which the standard deviation does not exist. This has been stressed by Mandelbrot [11, 12]. A sampling of other sources is [13–15].

The fractal dimension of the paths that enter the nonrelativistic path

integral has been discussed extensively and it has even been proposed that the dimension could be experimentally measured. See articles by Abbott and Wise [16], Cannata and Ferrari [17], Kroger [18, 19] and Kroger, Lantagne, Moriarty and Plache [20].

B. Path decomposition expansion

1. The formula

It often happens that a coordinate space breaks up in a natural way, for example the inside and outside of a region surrounded by a high potential energy wall. One may know what the propagator looks like inside and outside the wall, and would like to put together a propagator for the entire space. The answer to an easier question has been known since the beginning of path integration: if the environment changes as a function of *time*, say there is one potential before a certain time t_0 and a different one afterward, then the “before” and “after” propagators combine easily, and the full propagator is given by

$$G(x, t; y) = \int dz G^{\text{after}}(x, t - t_0; z) G^{\text{before}}(z, t_0; y), \quad (20)$$

where x, y and z , are in the system’s coordinate space, whatever that happens to be.

The *path decomposition formula*, gives a way to connect different parts of the coordinate space, while summing over times. The simplest example is in one dimension, $x \in [a, b]$ (where $-\infty \leq a < b \leq \infty$). Let $a < c < b$, then

$$G(x, t; y) = \int_0^t ds G(x, t - s; c) \left[\frac{i\hbar}{2m} \frac{\partial}{\partial z} G^{(r)}(z, s; y) \right]_{z=c}, \quad (21)$$

where $G^{(r)}(z, s; y)$ is the propagator for a particle restricted to the region $y \in [a, c]$. There is no restriction on the potential except that it be time-independent.

Where does Eq. (21) come from? The idea is that the particle is restricted to $[a, c]$ up until time s , after which it can be anywhere. The integrand is thus the amplitude for this collection of events, that is, it is the sum of “ $e^{iS/\hbar}$ ” for paths that stay in $[a, c]$ up until time- s , after which they can be anywhere. The integral over s is therefore interpreted—as usual—as the sum over the possible times for which this restriction is obeyed. But the astute reader will have noticed that I’ve been glib on

two points: first, the propagator for times t greater than s is the *full* propagator, not the propagator for $[c, b]$. Second, the propagator for the time interval prior to s appears with a derivative.

The fact that the full propagator appears for times later than s can be offset by judicious choice of c . For tunneling applications you would like c to be such that on one side is the well, on the other side freedom. This works for the energy-dependent propagator, $\tilde{G}(E)$, if you take c on the outer slope of the potential, away from the well. For the time-dependent propagator this strategy may still work in an approximate way.

Regarding the second point, the derivative in Eq. (21) *had* to appear. This is because the restricted propagator $G^{(r)}(z, s; y)$ vanishes on the boundary, $z = c$. The path restriction forces Dirichlet boundary conditions on $G^{(r)}(z, s; y)$, just as for the hard wall case discussed in Sec. 6. The derivative in Eq. (21) represents a current across the boundary. If you take the restricted propagator for a particle confined to a half line [e.g., Eq. (6.43), page 40], apply $i\hbar\partial_x/2m$, and evaluate at the boundary, you will find the classical velocity times the free propagator.

The result, Eq. (21), has been derived in a number of ways. Most satisfying to me is time slicing [breakup into integrals at times $t_k = k(t_{\text{final}} - t_{\text{initial}})/N$], which I present below. But ordinary Green's function techniques can also be used, and there is another approach, using operators, that allows restrictions more general than those defined by the coordinate space.

2. Proof of the path decomposition expansion

For simplicity take the points a and b to be $\mp\infty$ and c to be 0. The time interval is $[0, t]$. Let $t_k = k\epsilon$ with $\epsilon = t/N$ for large N . Then for each path from y to x (with $y < 0 < x$) there will be a last time that that path was entirely to the left of the origin. Let that time lie between t_m and t_{m+1} . Then for all integrals $\int dx_k$, for $k \leq m$, the range of integration is restricted to $(-\infty, 0)$. On the other hand, x_{m+1} only varies from 0 to ∞ , since if it took negative values, t_m would not be the last time the particle was confined to $x < 0$ [21]. For all integrals *after* this one ($\geq m + 2$), the range of integration is over the entire real line.

The total propagator is the sum over all these possibilities, i.e.,

$$\sum_m \int_{-\infty}^{\infty} \prod_{k=m+2}^{N-1} dx_k \int_0^{\infty} dx_{m+1} \int_{-\infty}^0 \prod_{k=1}^m dx_k \mathcal{A} \exp(iS/\hbar), \quad (22)$$

where the integrand is only indicated schematically. If all the integrals from 1 to $m - 1$ are performed one gets the restricted propagator from

y to x_m in time t_m . Similarly the integrals beyond the $(m + 1)^{\text{th}}$ give the full propagator from x_{m+1} to the final point x in time $t - t_{m+1}$. As a result there are only two integrals to perform:

$$G(x, t; y) = \sum_m \int_0^\infty dx_{m+1} \int_{-\infty}^0 dx_m G(x, t - t_{m+1}; x_{m+1}) \\ \times \left[\frac{m}{2\pi i \hbar \epsilon} \right]^{\frac{1}{2}} \exp\left(\frac{i m (x_{m+1} - x_m)^2}{2\epsilon \hbar} \right) G^{(r)}(x_m, t_m; y). \quad (23)$$

[See [22] concerning the absence of $V(x_m)$ in Eq. (23).] We know that the $(x_{m+1} - x_m)^2/\epsilon$ in the exponent keeps these two variables together, which means they both must be close to zero. Before doing the integrals let me show how we “know” this.

Notation: define $u \equiv x_m$, $v \equiv x_{m+1}$, $\tau' \equiv t_{m+1}$, $\tau \equiv t_m$ and $D \equiv \hbar/m$. We also refer to the m^{th} term in the sum over m in Eq. (23) as Φ_m . Finally we define a shorthand for the free propagator, $g(x, t) \equiv \exp(ix^2/2Dt)/\sqrt{2\pi Dt}$.

Now rewrite the short time propagator

$$g(v - u, \epsilon) = \frac{\lambda}{2\pi} \int_{-\infty}^\infty dq e^{-iq^2} \exp(-i\lambda q v + i\lambda q u), \quad (24)$$

with $\lambda \equiv \sqrt{2/D\epsilon}$. With this split, the integral over u can be performed separately. Using the identity $\exp(i\lambda qu) = (1/iq\lambda)(d/du)\exp(i\lambda qu)$, we integrate by parts

$$\int_{-\infty}^0 du e^{i\lambda qu} G^{(r)}(u, \tau; y) = \left[\frac{e^{i\lambda qu}}{i\lambda q} G^{(r)}(u, \tau; y) \right]_{u=-\infty}^0 \\ - \frac{1}{i\lambda q} \int_{-\infty}^0 du e^{i\lambda qu} \frac{\partial}{\partial u} G^{(r)}(u, \tau; y) \quad (25)$$

The term in square brackets vanishes for the following reasons. At infinity we assume that there is a small imaginary part to regularize it, while at 0, a restricted propagator always vanishes at the endpoint of its restriction. Thus only the derivative of $G^{(r)}$ contributes. Furthermore, the integration by parts can be done again. The piece arising from the total derivative no longer vanishes, and *depends only on the value of the derivative of $G^{(r)}$ at zero*. The remaining integral is asymptotically smaller by a factor λ . (See Sec. ID 2.)

The same argument can be applied to the v integral. In this case the leading term is $G(x, t - \tau'; 0)$, which does not vanish. It will turn

out that the individual contributions, Φ_m , to the integral in Eq. (23) are $O(\epsilon)$ (so that \sum_m turns into $\int ds$). Contributions from the derivative of G or from the second derivative of $G^{(r)}$ carry an additional power of λ , for an overall $O(\epsilon^{3/2})$, which can be neglected.

The foregoing observations will be used at a later stage, and we return to the double integral, Eq. (23), undoing the transformation (24). The next step is establish integral relations satisfied by g , the free propagator. These relations are most easily derived using the Laplace transform, and in particular the Faltung theorem.

Faltung facts The Laplace transform of a function $f(t)$ is $\hat{f}(\omega) \equiv \int_0^\infty dt e^{-\omega t} f(t)$. The *Faltung* theorem says that if \hat{f} and \hat{g} are the Laplace transforms of f and g respectively, then the Laplace transform of $\int_0^t ds f(t-s)g(s)$ is $\hat{f}(\omega)\hat{g}(\omega)$. The integral of $f(t-s)g(s)$ is the convolution, or Faltung (folding). Bearing in mind the form of the path decomposition expansion it is not surprising that the Faltung theorem is relevant. Here are the several results that we will use.

Everything will be stated in real form, with the i 's inserted later in blind, optimistic analytic continuation. First an already familiar expression:

$$\int_0^\infty dt e^{-\omega t} \frac{1}{\sqrt{\pi t}} e^{-k^2/4t} = \frac{1}{\sqrt{\omega}} e^{-k\sqrt{\omega}}, \quad (26)$$

for $k > 0$. In Sec. 7 this formula took us from the propagator to its energy dependent form. Next, the trivial identity

$$\frac{1}{\sqrt{\omega}} e^{-k\sqrt{\omega}} \left(-\frac{\partial}{\partial \ell} \right) \frac{1}{\sqrt{\omega}} e^{-\ell\sqrt{\omega}} = \frac{1}{\sqrt{\omega}} e^{-(k+\ell)\sqrt{\omega}}, \quad (27)$$

and the Faltung theorem imply

$$\frac{e^{-(k+\ell)^2/4t}}{\sqrt{\pi t}} = \int_0^t ds \frac{e^{-k^2/4(t-s)}}{\sqrt{\pi(t-s)}} \left(-\frac{\partial}{\partial \ell} \right) \frac{e^{-\ell^2/4s}}{\sqrt{\pi s}}, \quad k, \ell > 0. \quad (28)$$

We will also need to evaluate the above integral *without* the derivative with respect to ℓ . Consider

$$F(t, k, \ell) \equiv \int_0^t ds \frac{e^{-k^2/4(t-s)}}{\sqrt{\pi(t-s)}} \frac{e^{-\ell^2/4s}}{\sqrt{\pi s}} = \int_0^\infty d\omega e^{-\omega t} \frac{e^{-k\sqrt{\omega}}}{\sqrt{\omega}} \frac{e^{-\ell\sqrt{\omega}}}{\sqrt{\omega}} \quad (29)$$

(by the Faltung theorem). To get rid of the extra $1/\sqrt{\omega}$ we use $\omega^{-1/2} = \int_0^\infty d\mu e^{-\mu\sqrt{\omega}}$. It follows that

$$F(t, k, \ell) = \int_0^\infty d\mu \frac{e^{-(k+\ell+\mu)^2/4t}}{\sqrt{\pi t}}. \quad (30)$$

This is essentially the Error function. We evaluate F in the limit of small t . Asymptotically

$$F(t, k, \ell) = \frac{e^{-(k+\ell)^2/4t}}{\sqrt{\pi t}} \int_0^\infty d\mu e^{-\mu(k+\ell)/2t} e^{-\mu^2/4t} \sim \frac{2t}{k+\ell} \frac{e^{-(k+\ell)^2/4t}}{\sqrt{\pi t}}. \quad (31)$$

The formulas will be applied with the following substitutions: $t \rightarrow i\epsilon D/2$, $s \rightarrow i\sigma D/2$, $k \rightarrow v$ and $\ell \rightarrow -u$.

Eq. (28) and the above substitutions give the following identity for the propagator (recalling that $u < 0$)

$$g(v-u, \epsilon) = iD \int_0^\epsilon d\sigma g(v, \epsilon-\sigma) \left(\frac{\partial}{\partial u} \right) g(u, \sigma). \quad (32)$$

From the relations (29) and (31) we have in turn

$$\int_0^\epsilon d\sigma g(v, \epsilon-\sigma) g(u, \sigma) \sim \frac{\epsilon}{v-u} g(v-u, \epsilon). \quad (33)$$

Now we are ready to calculate. Going back to Eq. (23) and recalling that the summands are denoted Φ_m , we have

$$\begin{aligned} \Phi_m &= \int_0^\infty dv \int_{-\infty}^0 du G(x, t-\tau'; v) g(v-u, \epsilon) G^{(r)}(u, \tau; y) \\ &= \int dv du G(x, t-\tau'; v) \int_0^\epsilon d\sigma g(v, \epsilon-\sigma) \left[\frac{iD\partial}{\partial u} g(u, \sigma) \right] G^{(r)}(u, \tau; y) \end{aligned}$$

Next do an integration by parts with respect to u . Since $G^{(r)}(0, \tau; y) = 0$ and there is regularity at infinity, this only amounts to a change of sign and a transferral of the $\partial/\partial u$ to $G^{(r)}$. The integral over σ now has the form of Eq. (33), which we use (replacing the asymptotic relation by equality):

$$\Phi_m = -iD \int dv du G(x, t-\tau'; v) \frac{\epsilon}{v-u} g(v-u, \epsilon) \frac{\partial}{\partial u} G^{(r)}(u, \tau; y). \quad (34)$$

We now set $v = 0$ and $u = 0$ in the arguments of G and $G^{(r)}$, respectively, but not in g and $1/(v-u)$. To integrate over u and v we go to sum and difference variables. Let $w = v-u$ and $\rho = (v+u)/2$ (with unit Jacobian). By inequalities or by pictures, you can check that the integration range [$v > 0$ and $u < 0$] corresponds to [$w > 0$ and

$-w/2 < \rho < w/2$. This leads to

$$\Phi_m = \int_0^\infty dw \int_{-w/2}^{w/2} d\rho G(x, t - \tau'; 0) \frac{-iD\epsilon}{w} g(w, \epsilon) \left[\frac{\partial}{\partial z} G^{(r)}(z, \tau; y) \right]_{z=0}. \quad (35)$$

The integral over ρ gives w , cancelling the w in the denominator. The integral over w is now a Gaussian and g 's normalization is such that it would yield unity if integrated over the entire line. Since the integration is only over positive w the result is $1/2$. It follows that

$$\Phi_m = \epsilon G(x, t - \tau; 0) \left[\frac{D}{2i} \frac{\partial}{\partial z} G^{(r)}(z, \tau; y) \right]_{z=0}, \quad (36)$$

where I have dropped the τ/τ' distinction. We next sum the terms Φ_m , and noting the ϵ in Eq. (36), recognize that this gives the integral in the path decomposition expansion.

3. More than one dimension

For the higher-dimensional generalization of Eq. (21) one considers a volume surrounded by a surface σ such that prior to time- s the particle was within the volume, at time- s it crossed the surface (for the first time), and thereafter is unrestricted. The formula becomes

$$G(b, t; a) = \int_0^t ds \int_\sigma G(b, t - s; c) \frac{i\hbar}{2m} \frac{\partial G^{(r)}(c, s; a)}{\partial n_c} \cdot d\sigma. \quad (37)$$

$G^{(r)}$ is the restricted propagator for the interior of the volume and $\partial/\partial n_c$ the normal derivative (going out of the volume).

The restriction can also be imposed for other time periods, not just the beginning. That is, the formula can easily be adapted to a particle that is anywhere up to some time s , after which it is confined. Multiple restrictions are also possible.

Given the intuitive nature of this formula it is natural that it was already known in probability theory. What is peculiar is that it took until the 1980's to discover it for quantum mechanics.

A surprising use of this formula arose in relativistic quantum mechanics. Halliwell and Ortiz were puzzled because the composition law for relativistic propagators is more complicated than its nonrelativistic counterpart [which is Eq. (20), with "before" and "after" the same]. In particular, it involves derivatives. What they found was that these

derivatives arise because the relativistic formulas can be phrased as applications of the path decomposition formula. One starts with a path integral with a fifth parameter (Sec. 25) in which particle paths can go forward and backward with respect to physical time. Then one uses Eq. (37) with the intervening surface a space-like surface separating initial and final events.

Notes

The path decomposition expansion was developed by Auerbach, Kivelson and Nicole [23, 24] and was part of Auerbach's Ph.D. thesis. They used it to deal with quantum tunneling in the articles just cited as well as in collaboration with van Baal [25].

In [26], Goodman evaluates the propagator for a particle bouncing off a hard wall using path integration—in contrast to my Sec. 9 justification, which is basically the method of images. See Sec. III, “*comment concerning page 40*”, for Goodman's argument. His way of keeping track of paths is reminiscent of the way Auerbach et al. do the path decomposition formula derivation. Amusingly, Auerbach turned things around and worked out the method of images, starting from the path decomposition expansion [27].

Auerbach et al. also used Green's function techniques to derive the path decomposition formula, with extensions by van Baal [28]. Another approach, by Halliwell [29], uses operators. Halliwell, together with Ortiz [30] studied relativistic propagators.

Neutralizers are used in some approaches to asymptotics where you want to focus on a particular portion of the range of integration and you multiply your integrand by a C^∞ function that is 1 in your range of interest, 0 far away. Together with Ziolkowski, I used the path decomposition formula to develop this for the path integral [31]. In the same article, we also applied that formula to tunneling-time issues. Yamada has studied quantum tunneling and tunneling time, using the path decomposition formula in his work [32, 33]. In particular he has looked quite deeply into the problem of (quantum) measurements when more than a single time is involved and has pointed out the dangers of careless interpretations of quantum tunneling. Besides [31], Sokolovski and Baskin [34] and Fertig [35] have also applied path integral methods to the tunneling problem.

C. Checkerboard path integral

The Dirac equation for a particle moving in one space dimension is

$$i\hbar \frac{\partial \psi}{\partial t} = mc^2 \sigma_x \psi - ic\hbar \sigma_z \frac{\partial \psi}{\partial x}, \quad (38)$$

where ψ is a two-component spinor and σ_x and σ_z are Pauli spin matrices. The propagator, G , for ψ is a 2×2 matrix function of space and time. That is, $\psi(x, t) = \int dy G(x, t; y) \psi(y, 0)$. Feynman found the following way to compute G : fix an N and consider all paths of the sort shown in Fig. 1. That is, we allow zig-zag paths that travel upwards at 45° or 135° (i.e., at velocity c) and which may switch direction at times kt/N , $k = 1, \dots, N$. For each such path, let R be the number of reversals (switches) it suffers. For say, the $(++)$ element of G , take all paths that start at $(y, 0)$ moving to the right and that arrive at (x, t) moving to the right, as illustrated, and use them to compute the following sum

$$\sum_{\substack{\text{zig-zag} \\ \text{paths}}} \left(i \frac{t}{N} \frac{mc^2}{\hbar} \right)^R \quad (39)$$

For $N \rightarrow \infty$ this becomes G_{++} . From the figure one also sees that these paths are the legal moves of a king in checkers or of a bishop in chess. For this reason the sum is often called the checkerboard or chessboard path integral, although unlike their game board counterparts, these particles only move in one vertical direction (but see the notes for generalizations).

What are we to make of this? It was invented by Feynman in the 1940's. Presumably, he did not consider it important since he only published it as an exercise in his 1965 book with Hibbs. His goal, he said (privately), had been to start with space only and get spin. Thus in one space dimension his formalism already demands a two-component object. But when he couldn't do the same thing for three space dimensions he dropped the whole business.

Let me tell you all the things *wrong* with this idea. First the two components in the one-space-dimension Dirac equation have nothing to do with spin—there is no spin in one dimension. They are related to parity and in this light their connection to the left- or right-going paths looks reasonable. In fact for three space dimensions spin only requires two-component spinors, and it is again parity that doubles the number, to give the usual 4-component Dirac spinor. The next thing I never liked about this approach was the absence of any classical action. Perhaps Feynman had found a clever way to compute G , but “clearly” it could

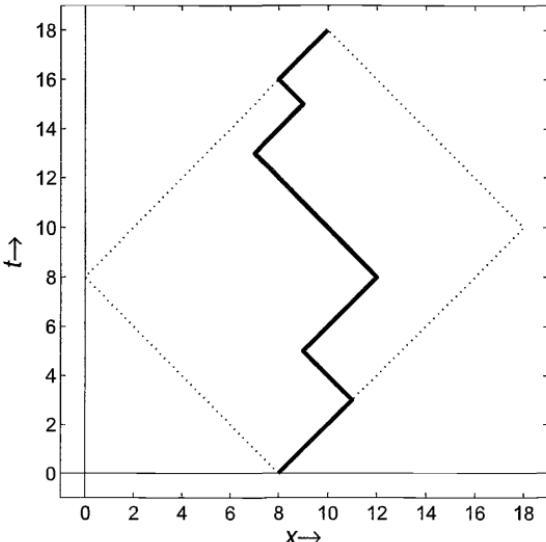


FIG. 1: A zig-zag path on a checkerboard discretization of space time. The dotted rectangle indicates the region of contributing paths. The velocity of light is taken to be 1.

have no relation to his more famous path integral if it had no action. Finally, and this I considered the most devastating observation, the scaling of Δx and Δt as $N \rightarrow \infty$ was wrong. By taking the paths to have light velocity you get $\Delta x = c\Delta t$ with $\Delta t = t/N$. Thus $\Delta x/\Delta t \sim \text{const}$ for $N \rightarrow \infty$. This is in dramatic contrast to the nonrelativistic path integral, for which Δt and the square of Δx are of the same size. A more detailed statement is

$$\frac{(\Delta x)^2}{\Delta t} \sim \frac{\hbar}{m}. \quad (40)$$

The quantity \hbar/m is the analogue of the diffusion coefficient. Thus in Eq. (9.6) the diffusion coefficient for Brownian motion emerges from the Δx and $\Delta t \rightarrow 0$ limits, while maintaining the ratio $D \equiv (\Delta x)^2/2\Delta t$ constant. This ultimately leads to a density, $\rho = \exp(-x^2/2Dt)/\sqrt{4\pi Dt}$. The “density” for the (nonrelativistic) path integral is $\exp(imx^2/2\hbar t)/\sqrt{2\pi i\hbar t/m}$, so that apart from factors of 2 and i , D and \hbar/m have the same role. That not-so-inconsequential i is the reason Eq. (40) has the symbol “~” rather than a proper equality.

Returning to Feynman’s formula, it is obvious that with so great a buildup I am about to tell you why I was wrong. The key to this problem

is to ask the right question. The right question is: for the important contributions to the sum in Eq. (39), how large is R ?

To answer this, group the terms in the sum according to the number of reversals. For given N , let $\phi^{(N)}(R)$ be the number of zig-zag paths with exactly R reversals; we have

$$G = \sum_R \phi^{(N)}(R) \left(i \frac{mc^2}{\hbar} \frac{t}{N} \right)^R, \quad (41)$$

where G and $\phi^{(N)}(R)$ have implicit subscripts, (\pm, \pm) , for path-direction labels. For that same N we want to determine for which R the quantity $\phi^{(N)}(R)(mc^2 t / \hbar N)^R$ is maximal. With a bit of combinatorics you can readily derive (see the exercise below) the fact that

$$R_{\max} = \frac{t}{\gamma} \frac{mc^2}{\hbar}, \quad (42)$$

where

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}, \quad v = \frac{b-a}{t}. \quad (43)$$

Remarkably, R_{\max} is independent of N . If you stare a bit at Eq. (42) you will see that the typical number of reversals is just the proper time in the particle's rest frame, measured in time units \hbar/mc^2 , which is the time for light to cross the particle's Compton wavelength.

Here is a way to look at this. The particle barrels along at the velocity of light. At any moment it can reverse direction just as when you have a large radioactive sample and at any moment one of the nuclei may decay. That is, there is a *rate* of reversal just as there is a *rate* of decay. For either random process, taking a finer mesh for time (increasing “ N ”) does not change the number of reversals or decays per unit physical time. The stochastic process associated with decay is the Poisson process and we have the familiar formula Prob(k decays in unit time for a process with decay rate r) = $r^k e^{-r}/k!$. Formally, Feynman's one-dimensional electron theory is such a process with one important—and familiar—difference. *Decay* in time dt has probability $r dt$. *Reversal* in time dt has probability *amplitude*, $i mc^2 dt / \hbar$, with an i . This parallels the real-imaginary correspondence discussed above following Eq. (40).

Why am I stressing this parallel? First to get across the idea that this checkerboard business has a certain richness to it, that it is not just plucked from the air. Second to pay tribute to one of the most important figures in the development of path integration, Mark Kac. It was Kac

who recognized around 1950 that Feynman's path integral was related by analytic continuation to the Wiener integral used in Brownian motion. As just shown (see also the notes and Footnote [36]) the checkerboard path integral is also an analytic continuation of a stochastic process, and what I wish to mention is that one of those involved in making this more recent connection was the same Mark Kac.

As a mathematical aside, the measure defined by the Poisson process with an imaginary parameter provides a bona fide measure on paths, unlike its nonrelativistic counterpart.

But there is still my main complaint. How can one get $(\Delta x)^2/\Delta t \sim \hbar/m$ from a theory that takes $\Delta x = c\Delta t$? The answer is a satisfying bit of asymptotics. For Brownian motion successive steps are completely uncorrelated. For the relativistic case, even though the electron is *allowed* to reverse at any moment it usually does not. In fact it has a correlation length $x \sim \hbar/mc$, the Compton wavelength. This means that if I tell you at some point which way it is going, it is likely to be going in the same direction for a distance $\Delta x \sim \hbar/mc$ no matter how finely the time interval is divided. On a short time scale the electron moves at the speed of light. But if you only check its position at widely separated intervals, many reversals will have taken place. Its (net) velocity is less than c and its direction in successive snapshots will be uncorrelated. This is Brownian motion.

Let us estimate the diffusion coefficient for that Brownian motion. The correlation length is \hbar/mc , so suppose that on shorter length scales it does not reverse, but that once it has traveled that far it can. For this motion $\Delta x \sim \hbar/mc$. Moreover, while it is travelling without reversing, its velocity is c , i.e., $\Delta t = \Delta x/c$. Therefore for this random walk

$$\frac{(\Delta x)^2}{\Delta t} = \frac{(\Delta x)^2}{(\Delta x)/c} = \left(\frac{\hbar}{mc} \right) c = \frac{\hbar}{m}$$

We have recovered the “diffusion coefficient” for the rough paths in the Feynman integral.

Poor time resolution therefore gives the nonrelativistic limit. You could have expected as much. For ordinary Brownian motion the infinite velocities arise from a mathematical idealization. Robert Brown's grains of pollen move at finite, if large, velocity between the closely spaced blows by water molecules. Similarly we now have insight into the infinite velocities predicted by nonrelativistic quantum mechanics. They are due to smearing over the motion at the smallest scales. As you improve time resolution, velocity increases, but only to its natural maximum, c . The checkerboard path integral shows how this transition is accomplished.

This picture of an electron's motion attributes its mass and lower-than- c velocity to the random process that makes it change direction; or stated more directly, (i times) the mass *is* the rate of flipping. Three dimensional versions of the checkerboard path integral (see below) preserve this property.

Exercise: Derive Eq. (42) by evaluating the $\phi^{(N)}(R)$ of Eq. (41). Don't forget the subscripts (\pm, \pm) on $\phi^{(N)}(R)$ and the associated constraints. Hint: it's a product of two combinatorial coefficients. See [37] for details.

Notes

Feynman first published the checkerboard path integral in his book with Hibbs [38], although Schweber, going through boxes of Feynman's notes, found his early calculations [39]. The checkerboard formulation was independently discovered by Riazanov [40].

The person who turned me around on the checkerboard path integral was Jacobson, who saw in it hints of a fundamental structure of space. Some of the material above is developed in [37], but Jacobson went further and set up a formalism for the full 3-space dimensional object in terms of spinors that pairwise reproduced space [41, 42]. When Kac learned of our results he recalled his own derivation of the telegrapher equation [36, 43] and together we worked out the relation between the Poisson process for the telegrapher equation and the prescription Feynman had developed for the one-dimensional Dirac equation [44].

The checkerboard path integral has attracted enthusiasts over the years, with applications to quantum tunneling, although most work has been devoted to extending the method itself, whether by going to higher space dimension, or by taking more general paths such as those going backward in time. Articles of this sort of which I am aware are by Ord, Jacobson, Ranfagni, DeWitt, Foong and their collaborators [45–54].

Gaveau and I also tried our hand at going to three space dimensions [55–57] and, as for other researchers, the simplicity of Feynman's one-dimensional method was not recovered. The problem lies in the way the gradient appears in the Dirac equation. Aside from the " σ " associated with parity, in one dimension the gradient is just d/dx , the generator of spatial translation. By pulling apart the two parity components, one can get this d/dx to produce paths in ordinary (one-dimensional) space. But in higher dimension one has $\sigma \cdot \nabla$, with the σ now the generator of rotations (and another set of Pauli matrices for parity). (The derivation in [55] allows this to be seen particularly clearly.) This object, $\sigma \cdot \nabla$, does *not* generate translations in 3-space. Instead one introduces other

objects, in our case Grassmann variables, that replace the simple “ dx ” of one dimension. For all these formulations the *square* of these “other objects” brings you back to ordinary space, which is in itself satisfying, but nevertheless lacks the appeal of the one-dimensional case. Perhaps I should be more positive: the “appeal” of a method is a matter of taste, and taste often needs to be educated. There may come a time when this is considered the natural way to think about the electron.

D. Exact solutions

In the Appendix to Sec. 6 I muse about the known exact solutions, circa 1980. What I mean by “exact” is an analytic expression for $\langle x | \exp(-iHt/\hbar) | y \rangle$, with t the physical time [58]. By this definition the list has not grown much, although the thesis that all such solutions correspond to semiclassical results has been knocked out. In the present section I will also mention another class of solutions that are significant contributions to the path integral literature and which are sometimes called “exact” under less stringent criteria.

1. δ -function path integral

Consider the Hamiltonian

$$H = \frac{p^2}{2m} + \lambda\delta(x), \quad x \in \mathbb{R}. \quad (44)$$

Given the ease with which one can solve for the eigenstates you might have thought the propagator would not present difficulties. Nevertheless, at the time this book was written, no complete form was known. As it turns out, the derivation can be skipped since the result—once you have it—can be verified directly.

Take units with $\hbar = 1$ and $m = 1$. As usual, the propagator $G(x, t; y) = \langle x | \exp(-iHt) | y \rangle$. Moreover, like the wave function itself, the propagator will have a discontinuity in its first (space) derivative at $x = 0$. Define first the free-particle propagator

$$G_0(x, t; y) = \sqrt{\frac{1}{2\pi it}} \exp\left(\frac{i(x-y)^2}{2t}\right). \quad (45)$$

Then the δ -function propagator is

$$G(x, t; y) = G_0(x, t; y) - \lambda \int_0^\infty du e^{-\lambda u} G_0(|x| + |y| + u, t; 0) \quad (46)$$

Since G_0 is a Gaussian integral, G itself is seen to be a Fresnel integral or Error function (if you do an analytic continuation). Verifying that this is indeed a solution involves just a slight subtlety in checking that certain quantities vanish, but I will not go into detail.

This propagator does *not* agree with what classical mechanics predicts. If you make a square barrier infinitely high but infinitely thin (preserving the product) a particle *never* gets through. If you make an attractive well infinitely deep and infinitely thin (again preserving the product), the particle passes through with zero contribution to the action (an easy calculation). This is at odds with Eq. (46), which gives the usual partial-transmission and partial-reflection.

Finally, it is possible using supersymmetry to generate entire classes of exact solutions, very much like that just presented (and including it as a special example). A reference is given below.

2. Half-plane barrier

Another exact quantum propagator, found since publication of this book, is for the half-plane barrier. In this case the particle is free except that the wave function must vanish on the half-plane defined by $y < 0$, all z . Since there is symmetry in the z direction we can drop to two dimensions and look only in the x - y plane, with the wave function vanishing on the negative y -axis. Using the angles defined in Fig. 2, the propagator, $G(\mathbf{b}, t; \mathbf{a})$, is built as follows: Let

$$\begin{aligned}\omega_1 &= (\theta_a + \theta_b)/2, & \omega_2 &= (\theta_a - \theta_b - \pi)/2, \\ \mu &= \sqrt{ab/\hbar t} \sin \omega_2, & \nu &= \sqrt{ab/\hbar t} \sin \omega_1,\end{aligned}$$

where $\mathbf{b} = b(\cos \theta_b, \sin \theta_b)$, etc. Define the function h as

$$h(u) \equiv \frac{1}{\sqrt{i\pi}} \int_{-\infty}^u \exp(iv^2) dv. \quad (47)$$

Take the particle's mass to be $1/2$. Then

$$G(\mathbf{b}, t; \mathbf{a}) = \frac{1}{4\pi i\hbar t} \exp \left[\frac{i(a+b)^2}{4\hbar t} \right] \quad (48)$$

$$\times \{ \exp(-i\mu^2)h(-\mu) \mp \exp(-i\nu^2)h(-\nu) \}, \quad (49)$$

where the upper sign corresponds to Dirichlet boundary conditions (the usual conditions for a barrier), but for completeness the lower sign is also given. The latter would be used for Neumann boundary conditions.

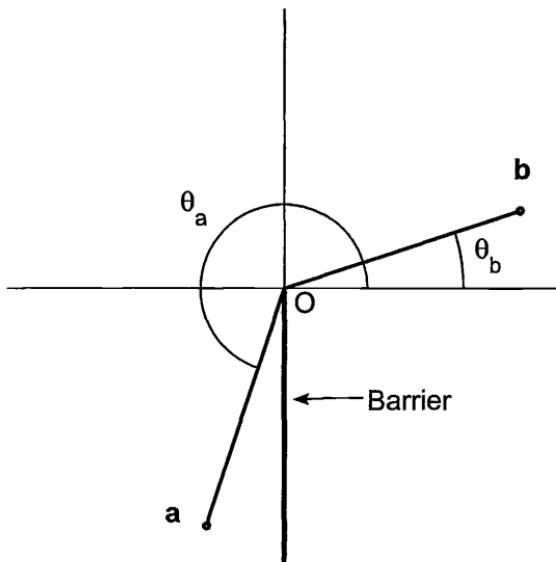


FIG. 2: Half plane barrier (two dimensional cut). The barrier lies on the negative y -axis. The points a and b are the spatial arguments of G , the propagator. The line $a-O-b$ is the path ξ given in Eq. (52).

Verification that this is indeed the propagator can be accomplished by taking a lot of derivatives. That's not the way G was found, but once you have the answer, it's the quickest. Note that through the function h the Error function or Fresnel integral again (as for the δ -function propagator) make an appearance.

Eq. (48-49) as it stands is not particularly transparent, but it simplifies if one uses the asymptotic form of h

$$h(u) \sim \Theta(u) - \frac{1}{2} \sqrt{\frac{i}{\pi}} \frac{\exp(iu^2)}{u} \sum_{n=0}^{\infty} a_n (iu^2)^{-n}, \quad |u| \text{ large}, \quad (50)$$

with

$$a_0 = 1, \quad \frac{a_n}{a_{n-1}} = n - \frac{1}{2}, \quad n > 0, \quad \Theta(u) = \begin{cases} 1 & \text{for } u > 0 \\ 0 & \text{for } u < 0 \end{cases}. \quad (51)$$

Geometrically, h will have its unit contribution from Θ when the points a and b are either mutually visible or when there is a path between them that reflects off the barrier. This allows the usual semiclassical formulas to be recovered. When there is no path, the propagator is smaller, and

in particular there is an additional factor $\sqrt{\hbar}$ multiplying it, as will be seen in detail below.

Nevertheless, for the case of no classical path, the term that *does* remain, cut down by $\sqrt{\hbar}$, has an interesting interpretation. Consider the path (with time variable s)

$$\xi(s) = \begin{cases} \mathbf{a}(t_0 - s)/t_0 & \text{for } s \leq t_0 \\ \mathbf{b}(s - t_0)/(t - t_0) & \text{for } s \geq t_0 \end{cases}. \quad (52)$$

This is the broken line $\mathbf{a}\text{-O-}\mathbf{b}$ (with O the origin). The time $t_0 \equiv at/(a+b)$ is chosen so there is no change in speed at the origin. In a sense it's the best you can do, given that you want to go from \mathbf{a} to \mathbf{b} and can't pass through the negative y -axis. How close is this to being a good classical path? By the Euler-Lagrange equations the first functional derivative of the action, $S = (1/4) \int \dot{x}^2 ds$, should be zero. Instead, for $\xi(s)$ it is

$$\frac{\delta S}{\delta \mathbf{x}(s)} \Bigg|_{\xi(\cdot)} = -\frac{(a+b)}{t} (\hat{\mathbf{a}} + \hat{\mathbf{b}}) \delta(s - t_0) \neq 0 \quad (53)$$

($\hat{\mathbf{a}} = \mathbf{a}/|\mathbf{a}|$, etc.). But $\xi(s)$ does have an extremal characterization: it minimizes S subject to the nonholonomic constraint forbidding transit through the barrier. It is easy to evaluate the action along ξ :

$$S[\xi(\cdot)] = \frac{1}{4} \left[t_0 \left(\frac{a}{t_0} \right)^2 + (t - t_0) \left(\frac{b}{t - t_0} \right)^2 \right] = \frac{1}{4} \frac{(a+b)^2}{t}. \quad (54)$$

Comparing this to the half-plane propagator, Eq. (48-49), one sees that the phase on line (48) is the action of a path from \mathbf{a} to the origin to \mathbf{b} . Moreover, the power of \hbar appearing on the left of line (48) is the correct power for two dimensions. But the point that I consider of greatest interest arises from the rest of G . For the case that there is neither a direct nor a reflected path, the last factor in G is asymptotically

$$\begin{aligned} \left[\begin{array}{l} \text{Shadow} \\ \text{correction} \end{array} \right] &\equiv \exp(-i\mu^2)h(-\mu) \mp \exp(-i\nu^2)h(-\nu) \\ &\sim \frac{1}{2} \sqrt{\frac{i\hbar t}{\pi ab}} \left[\sec\left(\frac{\theta_a - \theta_b}{2}\right) \mp \sec\left(\frac{\theta_a + \theta_b}{2}\right) \right]. \end{aligned}$$

This has the additional factor $\sqrt{\hbar}$, mentioned earlier, complete with a geometrical factor. At this point you might be tempted to say, "Great! I know how to deal with flawed paths: hit them with an extra $\sqrt{\hbar}$ and

work out a geometrical term that has information about how deeply shadowed the endpoints are from each other.” That would indeed be an excellent idea, and you would have rediscovered the beginnings of the “Geometric theory of diffraction,” due to Keller, which gives prescriptions for a variety of flawed paths. The context is not quantum mechanics, but electromagnetic wave propagation, and the power of the theory is not that you produce the rare exact solution, but rather that for geometrical configurations that approximate the exact layouts this prescription works asymptotically. Thus for a curved barrier with a sharp edge, you would calculate the field, or amplitude, in the shadow by using the geometrical diffraction theory prescription, except that you would have an asymptotic result. As long as the scale of the differences between the true barrier and the ideal one is larger than the asymptotic parameter (wavelength in Keller’s case), the method works.

Infinite dimensional perspective. The result (48-49) was first derived in a peculiar way: a third point, call it c , was defined in the plane, mutually visible from a and b . Next, the semiclassical propagators from a to c and from c to b were multiplied by one another and the result integrated over all appropriate c . (Reflection off the barrier was also included.) The result, after several asymptotic approximations, was Eq. (48-49), which is an *exact* solution. I won’t focus on the good fortune of unjustifiably landing on an exact solution, but instead will elaborate on the asymptotic features. First recall simple facts about asymptotics. Consider the integral

$$F(\lambda) = \int_0^\infty e^{i\lambda f(x)} g(x) dx, \quad x \in \mathbb{R}, \quad \lambda \text{ large.} \quad (55)$$

Usually (e.g., Sec. 11) one studies the “stationary phase approximation,” where for some $x_0 > 0$, $f'(x_0) = 0$. In that case

$$F(\lambda) \sim \sqrt{\frac{2i\pi}{\lambda f''(x_0)}} e^{i\lambda f(x_0)} g(x_0) \quad (56)$$

But if there is no stationary point, that is, $|f'(x)| > 0$, $\forall x \geq 0$, one can do an integration by parts:

$$F(\lambda) = \frac{i}{\lambda f'(0)} g(0) - \frac{1}{i\lambda} \int_0^\infty dx e^{i\lambda f} \frac{d}{dx} \left(\frac{g}{f'} \right) \quad (57)$$

The new integral on the right in Eq. (57) is like the old one, but carries an extra $1/\lambda$. Therefore, asymptotically

$$F(\lambda) \sim \frac{ig(0)}{\lambda f'(0)}. \quad (58)$$

Two features of this result should be emphasized: first, the asymptotic value of the integral is determined by the boundary of integration, in this case $x = 0$. Second, the overall scale is cut down by a factor $\sqrt{\lambda}$ relative to its value in the stationary phase case [Eq. (56)].

Returning to the path integral, we take the brave perspective that it is a sum over paths, an infinite dimensional integral,

$$G(\mathbf{b}, t; \mathbf{a}) = \int_{\Omega} \mathcal{D}\mathbf{x}(\cdot) e^{iS[\mathbf{x}(\cdot)]/\hbar}, \quad (59)$$

where Ω is the domain of functional integration. In the half-plane barrier case, Ω should not include paths crossing the negative y -axis (now down to 2 dim.). Therefore, depending on \mathbf{a} and \mathbf{b} , there may *not* be stationary “points” of the action, S in Ω . But Eq. (48-49) tells us what to do in that case: go to the best point on the boundary of Ω , in this case a path that just nicks the tip of the barrier, evaluate S for this path, and cut down the entire result by the square root of the asymptotic parameter—exactly what we did for one-dimensional asymptotics.

I have been tantalized by this connection for more than 20 years, and off-and-on have tried to put it all together. I will mention a bit more of what I believe must go into the connection, and hope that some enterprising reader can succeed. (Presumably one should first do this for the Wiener integral, where Eq. (59) has rigorous meaning, but I will continue to speak of the path integral and stationary phase approximations.) The n -dimensional generalization of Eq. (57) is

$$\begin{aligned} \int_{\Omega} d^n x g(x) e^{i\lambda f(x)} &= \frac{1}{i\lambda} \int_{\partial\Omega} d^{n-1} x \frac{g(x)\nabla f}{|\nabla f|^2} e^{i\lambda f} \\ &\quad - \frac{1}{i\lambda} \int_{\Omega} d^n x \nabla \left(\frac{g\nabla f}{|\nabla f|^2} \right) e^{i\lambda f} \end{aligned} \quad (60)$$

This formula is valid only if $\nabla f \neq 0$ throughout Ω . If ∇f had vanished in Ω there would be a $1/\sqrt{\lambda}$ multiplying the leading term, rather than $1/\lambda$. Moreover, as before, the first term on the right of Eq. (60) is larger by a factor λ than the second term. Thus the step from 1 dimension to n does not substantially change the asymptotics. The tough part is the case $n = \infty$ and establishing that the surface integral gives the desired term. One feature of Eq. (60) that only appears in dimension greater than one is that the “surface integral” (over $\partial\Omega$) is also subject to the stationary phase approximation, which is why in our half-plane example the flaw in the path $\xi(t)$ occurred only on the boundary. As a last remark I mention that the concept of neutralizer, so useful in asymptotics, can also be defined for the path integral [31].

3. Coulomb potential and related solutions

In classical mechanics there is a way to make the Kepler (or Coulomb) problem into a harmonic oscillator. It is known as the Kustaanheimo-Stiefel transformation and involves both a change of spatial coordinates and a change of time variable. The new time is well known in classical mechanics and is called the “eccentric anomaly” [59]. It is explicitly a function of a particular path: $s(t) \equiv \int_0^t d\tau / r(\tau)$, with r the radial coordinate of the particle. The change of coordinates is more complicated (in three dimensions) and one must ascend to a 4-dimensional coordinate space. Nevertheless, it brings considerable simplification to the problem and has been extensively used in classical contexts. Duru and Kleinert [60, 61] had the lovely idea that once you had a path integral you could take advantage of the Kustaanheimo-Stiefel transformation since, thanks to Feynman, there were paths, and functions $r(t)$, to work with *inside* the path integral. Thus for each path you go over to the new time. With the additional spatial variable, the dynamics becomes that of the harmonic oscillator, and can be done explicitly.

What these authors and others derived from this formalism is mainly information about the *energy*-dependent Green’s function. This is because the harmonic oscillator propagator that emerges is for paths in the “new” time, s . Since there is a different s for each path, they cannot all be gathered into a single *physical-time*-dependent propagator. For this reason the goal of an explicit function $G(\mathbf{r}'', t; \mathbf{r}) \equiv \langle \mathbf{r}'' | \exp(-iHt/\hbar) | \mathbf{r} \rangle$ has still not been attained with H the Coulomb Hamiltonian.

Notes

The propagator for the δ -function potential given in Sec. ID 1 was found by Gaveau and me [62] using two methods. The simpler was an expansion over eigenstates; the other was probabilistic and used the path decomposition expansion. Ref. [63] provides the verification (mentioned in the text) that the given propagator satisfies the time-dependent Schrödinger equation. Some years earlier Goovaerts, Babcenko and Devereese [64] had worked out a perturbation expansion for the path integral and applied it to the δ -function potential. They had a particular projection of the full propagator which nevertheless sufficed for the recovery eigenfunctions and scattering states.

The supersymmetry-based generation of exact propagators was found by Jauslin [65] and includes as a special case the δ -function.

In more than one dimension the δ -function is invisible; nevertheless,

people in nuclear physics long ago developed singular objects for idealized point scattering centers. Explicit propagators for such “potentials” were found by Scarlatti and Teta [66].

In 1982 there was a conference on “wave-particle duality” in honor of de Broglie’s ninetieth birthday. Keller’s geometrical diffraction theory [67] provides counterpoint to that theme: historically diffraction was a deciding factor in calling light a wave, but here was Keller with a way to calculate diffractive fields using rays, something you might be tempted to call particle paths. Because of the importance of paths in Keller’s story it seemed a good idea to derive his results from the path integral (cf. page 169). Ref. [68] uses a semiclassical approximation to derive the knife-edge (or half-plane) propagator, and noting factors $\sqrt{\hbar}$ and the dominance of the boundary term (of the range of functional integration) I conjectured that this was part of a larger asymptotic theory. The later realization that the result was exact led to an additional publication [69]. Apparently the exactness arises because the plane-with-barrier is the projection of a two-sheeted Riemann surface with the cut located along the barrier [70]. On this surface the motion looks free. This was exploited by Sommerfeld and for the heat kernel by Carslaw. For some of the history of this problem see [71].

As to getting exact results after lots of approximations, my only reaction is to recall the witticism (of Dyson?) that a good physicist is one who makes an even number of mistakes.

Following the original papers by Duru and Kleinert [60, 61] there was a great burst of activity, some offering alternative demonstrations, and many extending the technique to other systems. A small part of the literature is Refs. [72–77].

E. Dissipation and other forced-oscillator applications

You could make the case that the most powerful, effective, and famous applications of the path integral are based on the forced harmonic oscillator. The explicit path integral is given in Eqs. (6.41) and (6.42). As seen in Sec. 21 on the polaron, this comes into play when you have a complicated degree of freedom in contact with a bunch of oscillators, notably the electromagnetic or phonon field, and the coupling to those oscillators is linear in the oscillator coordinate. Call the “complicated” degree of freedom r and the oscillator coordinates Q_n , $n = 1, \dots$. Then the coupling in the Lagrangian will be $\sum f_n(r)Q_n$, with the form of f_n depending on the specific problem. Because $\int D\tau(\cdot)$ is performed last, as far as Q_n is concerned $f_n(r(\cdot))$ is just an external force.

At the beginning of the 1980's a new application of this idea sprung to life and has been going strong ever since. The physical system first studied was the Josephson junction, for which the state of a large number of electrons is subsumed into a single degree of freedom, the trapped flux in a superconducting ring interrupted by a resistor. At a phenomenological level this flux variable, ϕ , satisfies an equation of motion

$$M\ddot{\phi} + \eta\dot{\phi} = -dV/d\phi + F_{\text{ext}}(t), \quad (61)$$

where the various parameters, the potential and the external force characterize the system, and in particular η is a measure of dissipation. The case of particular interest is the "SQUID" (superconducting quantum interference device) for which the potential V can be made to be a double well. For the usual situation, having the system variable ϕ be in one or the other well means that the supercurrent is flowing in one direction or the other around the loop that constitutes the SQUID. At the quantum level an amazing possibility opened up. This variable could tunnel, by *quantum* tunneling, between the wells. This would mean that a collective variable for perhaps 10^{11} electrons would behave in a purely quantum fashion.

The theoretical framework for studying this tunneling was to model the source of the dissipation, namely to put in the coordinates of the phonons that provided the resistance in the junction. Thus one would have a quantum treatment of the whole system: energy would go from the principal degree of freedom to the resistor, but that degree of freedom could itself go from being in one state (localized on one well) to another (the other well).

The coupling of the phonons to the system degree of freedom could indeed be set up in the desired form (linear forcing) and the tunneling calculated. The Lagrangian used by Caldeira and Leggett [78] was

$$L = \frac{1}{2}M\dot{q}^2 + V(q) + \frac{1}{2}\sum_{\alpha}m_{\alpha}\dot{x}_{\alpha}^2 + \frac{1}{2}\sum_{\alpha}m_{\alpha}\omega_{\alpha}^2x_{\alpha}^2 + q\sum_{\alpha}c_{\alpha}x_{\alpha},$$

with q the dynamical variable (ultimately identified as ϕ), $\{x_{\alpha}\}$ the phonon coordinates, and $\{c_{\alpha}\}$ characterizing the coupling. One of the important physical properties that played a role in the results was the phonon spectrum, embodied in $\{c_{\alpha}\}$ and implicit in \sum_{α} . Simplifications were also introduced, for example replacing q by a two-state variable, giving the entire system the name "spin-boson model." As for the polaron, one first does the path integral for the $\{x_{\alpha}\}$, then sandwiches the propagator between states of the oscillators, typically the ground states. After this the functional integral involves the original Lagrangian in q

plus the result of all the completed integrations, which can now depend *only* on q . This gives an action contribution (the imaginary part of the logarithm of the result of all those integrals), which is an *effective* self-interaction of q . This self-interaction is non-local in time (as for the polaron) and represents the back-reaction of the field of oscillators on the motion of the variable of interest, q .

By now there are books and review papers that not only describe the SQUID application but also the many other situations where a system interacts with a heat bath, and nevertheless retains quantum properties. Typically, but not always, energy concentrated in the principal degree of freedom leaks into the oscillator modes, i.e., there is dissipation. I will not document the history of this application, but only cite sources where the reader can get into the subject [79–83]. More on this (and other) literature appears in Sec. II.

Meanwhile, other applications of Feynman's forced-oscillator integration continue to flourish. The polaron has historically had a special role: even those content to go through life without the pleasures of path integration will concede that for the polaron you can compute things using the path integral that are difficult to get by other methods. A review by Devreese [84] summarizes the polaron situation as of 1996. His group in Antwerp continues to work in this area [85, 86] and to use the path integral for other physical systems as well [87].

I recently used the forced oscillator technique for a condensed matter application, the quantization of *discrete breathers*. These are lattice excitations involving strong nonlinearity. Classically such localized oscillations are a kind of stationary soliton and can live forever [88]. By massaging the system it can be brought to a form suitable for the forced oscillator machinery, and the quantum stability of the system can be studied [89]. Let me pass on one lesson I learned from this: trust no one (including me). Even if you take Eqs. (6.41) and (6.42) as correct (which I believe they are), you still have several Gaussian integrals to do. Then there's usually an analytic continuation and a double integral variously written $\int_0^T dt \int_0^T ds$ or $\int_0^T dt \int_0^t ds$, which brings in another factor two. For the serious people, I trust them to have gotten all these signs and factors right in their research papers—but not in their lectures! In my work, [89], I wasn't sure of all these factors and signs until I found a way in which the effective action could be used to recover the spectrum of the original system, and when it agreed with that spectrum calculated by other means, I stopped worrying. By the way, symbolic manipulation programs, at least in the hands of my collaborators, were of no help. (See also the “comment on page 173” in Sec. III.)

F. Mathematical developments

Making proper mathematics out of the path integral has been a dream since Feynman's first paper on the subject. Like Newton and Dirac, mathematical innovators in quest of physical knowledge, Feynman knew how to get things right even if not all the mathematical i's were dotted. And perhaps, as for the inventions of his predecessors, a full mathematical theory will follow. In Sec. 9 I mentioned work of DeWitt, Albeverio, Hoegh-Krohn and Truman, all of them having this goal in mind. Cameron (cited in the Sec. 9 notes) had found that the path integral could not provide a bona fide measure. But if you weaken the demands, redefine "bona fide," new possibilities open. For example, Tepper and Zachary [90] have relaxed the need for countable additivity in the measure and find a useful object nevertheless. Another tack has been taken by Fujiwara and Kumano-go [91, 92], who focus on the properties of the time-sliced expression for the propagator, establishing classes of functions for which that is meaningful, and developing asymptotics in that context as well. Yet another approach is by Nakamura [93], who has been using non-standard analysis [94], taking advantage of that framework's ability to deal with explicit infinities and infinitesimals.

One avenue that appeals to me physically (although it has received only slight attention [95]) would start from the checkerboard path integral. For this there is a well defined measure. Then find its non-relativistic limit, which as discussed in Sec. IC, is the usual path integral.

G. Electromagnetism and other PDE's

In Sec. 20 I showed how the path integral could be applied to geometrical optics, basically to the wave equation, where the big step is to get rid of the $\partial^2/\partial dt^2$ by looking at a fixed frequency. Then, by hand, I inserted an artificial time variable, so as to make it look like the Schrödinger equation. My goal was aesthetic, to show the common features of Fermat's principle of optics and the semiclassical approximation of quantum mechanics.

It turns out that there's a practical side to this resemblance as well. People really do use the path integral for the solution of partial differential equations other than the Schrödinger and diffusion equations. Furthermore, in some of these applications the "time" variable has a clear physical meaning as the spatial variable along the direction of propagation (analogous to the eikonal approximation). Quite a few papers on

sound transmission in channels have appeared and Refs. [96–98] should help you find your way into this field. More recently, Schlottmann [99] pioneered the use path integral methods in geology. This is a difficult wave transmission problem, since not only is the medium nonuniform, but it's that uniformity that you are interested in. Impressive results of in this area have been obtained by Fishman [100, 101], whose path integral methods for the Helmholtz equation can provide a detailed picture of geological structure, especially useful for oil exploration.

For electromagnetism path integral methods have been used to find the low order modes of complicated waveguides [102, 103]. This is not the time-dependent propagator, but an application of the method of Sec. 7, taking advantage of the dominance of the lowest mode for large imaginary times. As a practical method this was first proposed by Donsker and Kac in the 1950's, and I believe they even did numerical calculations (punch cards and all). There have also been significant extensions of ray tracing techniques [104].

H. Homotopy

For diverse reasons, there are physical systems where the homotopy of the coordinate space is not trivial. As described in Sec. 23, the path integral provides an intuitive approach to multiple connectedness, since homotopy theory, like the path integral, focuses on paths, with its starting point being an equivalence relation among them.

The phase that can be associated with winding number, discussed in Sec. 23.1, has as its most common application the Aharonov-Bohm effect. The path integral/homotopy perspective goes back, as far as I know, to my thesis and its extension (for the Aharonov-Bohm effect), Ref. [105]. A physical arena where this frequently arises is the mesoscopic superconducting ring, a recent paper on this being [106]. See also [107, 108]. A closely related way to think about the richness permitted by the multiple connectedness is in terms of the Berry phase.

These phenomena also are present in non-Abelian gauge theories and have been studied by Sundrum and Tassie [109].

A somewhat different application came up in connection with the concept of *anyons*. As Laidlaw and DeWitt realized (Sec. 23.3), in two dimensions you can get additional kinds of quantum statistics, not just fermions and bosons. Such excitations were proposed in the attempts to understand high temperature superconductivity [110, 111]. (See also Canright's lectures in [112].) Ultimately, in the realm of superconductivity these ideas were put aside, but other applications have been proposed.

I. And more ...

Sec. 32 of this book is entitled “Omissions, Miscellany and Prejudices,” and represented my effort to get just a bit more material under the wire of a publication deadline. In this respect, little has changed in 20+ years. Here too I present just the briefest guide to a number of developments that have caught my attention.

1. *Variational corrections to the classical partition function*

An approximation for the partition function was independently developed by Feynman and Kleinert [113] and by Giachetti and Tognetti [114] in the mid 1980’s. They sought quantum corrections to the semi-classical partition function, making use of quadratic approximations to the potential energy function. These approximations were not merely the lowest quadratic fit, but were themselves found through a variational principle. In both cases, the initial publication was followed by several others of which a small sample is [115, 116].

2. *A particular numerical approach*

There has been an enormous amount of numerical work using the path integral (usually its real-valued analytic continuation), including cases where fermion statistics create a severe sign problem. I will not review this, although lectures by Pollock can be found in [112] and other literature is given in Sec. II. Here I want to mention one scheme that has not been pursued on an industrial scale and which at first sounds as if it shouldn’t work.

Since the path integral is the sum $G = \int \mathcal{D}x(\cdot) \exp(iS/\hbar)$, you could imagine a change of variable to a one dimensional integral $G = \int ds \Omega(S) \exp(iS/\hbar)$, where $\Omega(S)$ is the number of paths with action S . It’s an ambitious change of variables and requires real belief in the path-sum concept. By this method all the trouble is transferred to the evaluation of $\Omega(S)$. Creswick [117] proposed this idea and worked out a number of examples with numerical and analytic estimates for $\Omega(S)$. My first impression was that the roughness of the significant paths mitigated against having any reasonable limit for $\Omega(S)$ as the time-slicing mesh went to zero. However, although the limit may not exist, for any particular time-slicing (say N “slices”) Ω can be evaluated, and as a function of N it is possible to define sensible properties. In a way this

reflects the way I believe one should deal with path integral subtleties: go back to the multiple integral, $\int \prod_{n=1}^N dx_n$. Or as Kac is reputed to have said: when in doubt, discretize.

3. Semiclassical approximations for coherent states

In molecular systems there occasionally develop large magnetic moments. Such a moment may have more than one preferred orientation and there will be a quantum tunneling amplitude for transiting between classically degenerate states. Given the largeness of the moments it would be reasonable to expect semiclassical tunneling calculations to be helpful. One would also expect that the coherent state formalism, which has often been used to represent such degrees of freedom, would be a good vehicle for such a calculation, especially since its phase space interpretation brings it close to a semiclassical language. Nevertheless, considerable difficulties were encountered. In recent work more reliable techniques have been developed and I refer the reader to [118] as an entry into this literature.

4. Chaos and order

The use of the path integral in studying quantum chaos was established with Gutzwiller's trace formula [119]. In many places in this book [e.g., Eq. (13.25)] the propagator is approximated by a sum over classical paths, which means that for given endpoints in $G(b, t; a)$ there may be several solutions of the classical equations of motion. The propagator is then given by their sum, with interesting possibilities for interference and other quantum phenomena.

For a classically chaotic system the number of solutions to the two-time boundary value problem associated with $G(b, t; a)$ generally grows exponentially with t . If t is long enough you'd therefore expect a breakdown in the semiclassical approximation. This is not what happens. Even for long times a great deal of useful information can be extracted from the propagator.

By taking a trace of the propagator and applying a stationary phase approximation, Gutzwiller focused on the *periodic* orbits. This is because the trace forces initial and final positions to be the same while the stationary phase approximation takes care of matching the momentum. These periodic orbits can be quite long; nevertheless, their contribution to the propagator reliably gives information about the quantum system,

for example concerning the density of states. The literature on this subject extensive. A convenient source, with many references, is [120] by Cvitanovic et al., which is directly available on the web (the Gutzwiller trace formula is in Chapter 30).

Better than expected accuracy in the path sum for a classically ergodic system was also seen by Heller and Tomsovic [121]. They studied a “stadium billiard” (a free particle in a kind of elongated oval) and found that even when there were 30000 paths to add, a great deal of quantum information survived intact. An important technical point is that they were not actually looking at the propagator as a function of its position endpoints; rather they sandwiched it between coherent states, thereby smoothing the singularities that would otherwise arise from caustics of the motion. One feature that helped me understand this discovery was that although in the two-dimensional coordinate space the paths crossed and re-crossed each other, in *phase space* there is lot more room [122]. Another path-integral oriented approach to the smoothing of caustic singularities, especially in chaotic situations, is by Takatsuka [123].

II. LITERATURE

The library of path integral books has grown substantially since the present volume was first published. There are textbooks, conference proceedings and review papers. Books focussed on other subjects, for example quantum field theory, will often have introductions to the path integral. Below is a partial list of publications since this book appeared. I have not tried to be comprehensive.

A. Texts and review articles

Roepstorff, *Path Integral Approach to Quantum Physics: An Introduction* [124]. Mathematically careful, but not at the expense of good instruction.

Kleinert, *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics, and Financial Markets* [15]. Covers an impressively wide range of topics, including the Duru-Kleinert Green's function solution discussed in Sec. ID 3.

Kac, *Integration in Function Spaces and Some of Its Applications* [125]. Notes for lectures Kac gave in Italy. Marked by his usual lively style, although they may be hard to locate.

Wiegel, *Introduction to Path-Integral Methods in Physics and Polymer Science* [71]. Emphasizes polymer applications and is a readable and informative volume.

Gutzwiller, *Chaos in Classical and Quantum Mechanics* [119]. Extensive coverage of quantum chaos, with the path integral serving as a (very effective) tool.

Glimm and Jaffe, *Quantum Physics: a Functional Integral Point of View* [126]. The work of these authors played a part in the development of “constructive quantum field theory,” a subject for which the functional integral is used in a mathematically rigorous way.

Weiss, *Quantum Dissipative Systems* [127]. An introduction with many applications to the use of functional integration in handling dissipation in quantum systems.

Dittrich, Hägggi, Ingold, Kramer, Schön and Zwerger, *Quantum transport and dissipation* [83]. A collection of articles offering introductions to the fields indicated by the title.

Cerdeira, Lundqvist, Mugnai, Ranfagni, Sa-yakanit and Schulman, *Lectures on Path Integration: Trieste 1991* [112]. A collection of diverse lectures at a pedagogical level, given at a workshop at the ICTP, Trieste. As a co-organizer I wanted the topical coverage to be an update of sorts for this book.

Roncadelli and Defendi, *I Camini di Feynman* [128]. A pedagogical work in Italian, also reflecting contributions of the authors to the field.

Choquet-Bruhat and DeWitt-Morette, *Analysis, Manifolds and Physics* [129]. Introduction to many mathematical methods in physics, with in-depth treatment of the path integral.

Ranfagni, Mugnai, Moretti and Cetica, *Trajectories and Rays: The Path-Summation in Quantum Mechanics and Optics* [130]. A path oriented approach to problems in quantum mechanics and optics.

Freidlin and Wentzell, *Random Perturbations of Dynamical Systems* [131]. Functional integrals as well as other probabilistic methods, with emphasis on asymptotics.

Steiner and Grosche, *A Table of Feynman Path Integrals* [132]. A useful compendium.

Books on quantum field theory. As the reader may have guessed, quantum field theory is not my favorite arena for path integration. But for lots of other people it is (“*Let a hundred flowers bloom, ...*”). A few of the more well-known texts that make extensive use of path integration are Popov [133], Ramond [134], Itzykson and Zuber [135], Faddeev and Slavnov [136], and Rivers [137].

Review articles

Leggett, Chakravarty, Dorsey, Fisher, Garg and Zwerger, *Dynamics of the dissipative two-state system* [81]. Justification and use of the spin-boson model.

Grabert, Schramm and Ingold, *Quantum Brownian Motion: The Functional Integral Approach* [82]. Pedagogical presentation at an advanced level.

Ceperley, *Path integrals in the theory of condensed helium* [138]. Reviews an extensive literature in which Feynman’s original idea of how to deal with Helium (as a classical model of interacting “polymers”) is implemented with intensive numerical computation.

Khandekar and Lawande, *Feynman Path Integrals: Some Exact Results and Applications* [139]. Quite a few explicit path integrals including some that are nonlocal in time.

Gaspard, Alonso and Burghardt, *New Ways of Understanding Semiclassical Quantization* [140]. Extensive background on semiclassical methods, with the path integral playing a prominent role. Application to quantum chemistry, including reactions and chaotic scattering. Higher order (in \hbar) corrections to the Gutzwiller trace formula.

Littlejohn, *The Semiclassical Evolution of Wave Packets* [141]. Not really path integration, but a trove of information and perspectives on semiclassical approaches.

Shankar, *Renormalization-group approach to interacting fermions* [142]. Extensive use of the path integral for *fermions* (something not covered in this book) to study interaction between them. The fermion path integral uses Grassmann variables and these too are introduced in this article.

B. Conference proceedings

In 1983 the first of the “Path integrals from meV to MeV” conferences took place, organized by Akira Inomata, in Albany, New York. Since then, every three or four years there has been another such meeting, although the name of the conference has evolved. The original title was selected to emphasize physical phenomena, in contrast to several mathematically oriented meetings in the then recent past. Today we also have mathematicians coming and the limits meV and MeV have variously been replaced by quarks, peV, quantum information, TeV, galaxies and cosmology. The proceedings give an overview of the range of applications, although the mode of publication has varied from meeting to meeting and some of the proceedings may be hard to track down. The meetings: Albany 1983 [143], Bielefeld 1986 [144], Bangkok 1989 [145], Tutzing 1992 [146], Dubna 1995 [147], Florence 1998 [148], Antwerp 2002 [149], Prague 2005 [150].

Besides the meV/MeV series, there have been conferences and workshops where the path integral played a significant role. These can be a useful way to survey activity in the field, although these days people are more inclined to search the web than to get off their chairs and go to the library. A non-exhaustive list of conferences is [151–154].

III. PAGE BY PAGE COMMENTS AND ERRATA

Conventions: references to “Sections” 1 through 32 are to sections of this book (a.k.a. chapters). Equation numbers with periods, $(n.m)$, refer to equations in the book; those consisting of a number alone refer to the supplement. Sections of the supplement have labels beginning with Roman numerals.

- *Sec. 1, page 4, Eq. (1.11); elaboration of the associated footnote.*
Show that

$$f(\lambda) \equiv e^{\lambda A} e^{-\lambda(A+B)} e^{\lambda B} = 1 - \frac{1}{2}\lambda^2[A, B] + \dots$$

Evaluate $df/d\lambda = Af - e^{\lambda A}(A+B)e^{-\lambda A}f + fB$. Clearly $f'(0) = 0$. The second derivative is $f'' = Af' - e^{\lambda A}A(A+B)e^{-\lambda A}f + e^{\lambda A}(A+B)Ae^{-\lambda A}f + f'B$. Since $f'(0) = 0$, the only term surviving is $f''(0) = [B, A]$.

- *Sec. 6, page 39, solution to part (b), exercise at top of page.*
The correct form of the propagator is

$$G(\mathbf{x}_2, t; \mathbf{x}_1, 0) = \left[\frac{m}{2\pi it} \right]^{3/2} \exp \left\{ \frac{im}{2} \left[\frac{\omega}{2} \cot \frac{\omega t}{2} (\hat{\mathbf{B}} \times \mathbf{r})^2 - \omega \hat{\mathbf{B}} \cdot \mathbf{x}_2 \times \mathbf{x}_1 + \frac{1}{t} (\hat{\mathbf{B}} \cdot \mathbf{r})^2 \right] \right\}$$

For more on this propagator, see Glasser [155]. When there are constant electric and magnetic fields on a lattice, although the fields share the underlying periodicity, the potentials giving rise to those fields do not. See [156] for a path integral treatment of this situation.

- *Sec. 6, page 40, discussion following Eq. (6.43).*
Eq. (6.43) asserts

$$G(b, t; a) = \sqrt{\frac{1}{2\pi it}} \left[e^{i(b-a)^2/2t} - e^{i(b+a)^2/2t} \right], \quad \text{Sec. 6, Eq. (6.43)}$$

where G is the propagator for a particle confined to $x > 0$, $x \in \mathbb{R}$ but otherwise free ($V = 0$). The justification (in Sec. 9) uses what is essentially the method of images. I complain on page 40 that it is an “embarrassment to the purist” that I did not present a sum-over-paths derivation. Shortly after the book was published, Goodman [26] gave the following purely path integral argument.

As usual write

$$G(b, t; a) = \sum \exp(iS/\hbar),$$

with the sum taken over $x(\cdot)$ such that $x(0) = a$, $x(t) = b$ and $x(s) > 0$, for $0 \leq s \leq t$. We also define the free propagator, denoted $G_0(b, t; a)$, which is the sum over paths from a to b in time t , but *without* the restriction. The difference between G and G_0 is the sum over paths that have at least once entered the region $x < 0$, but returned to b nevertheless. Call these “forbidden” paths. Then $G(b, t; a) = G_0(b, t; a) - \sum_{\text{forbidden paths}} \exp(iS/\hbar)$.

Consider a forbidden path, $x(\cdot)$. There will be a last time, t_ℓ , for which $x(t_\ell) < 0$. Define the path $\bar{x}(\cdot)$ as follows

$$\bar{x}(s) = \begin{cases} x(s) & \text{for } s < t_\ell \\ -x(s) & \text{for } s > t_\ell \end{cases}.$$

The final position of $\bar{x}(\cdot)$ is thus $-b$ rather than b . The difference between $S(x(\cdot))$ and $S(\bar{x}(\cdot))$ goes to zero as the mesh for time-slicing goes to zero, because the flip in direction at t_ℓ has negligible impact on S . On the other hand, the set of paths $\bar{x}(\cdot)$ includes *all* paths from a to $-b$ in time t . As a result the sum, $\sum_{\text{forbidden paths}} \exp(iS/\hbar)$, is in fact equal to the free propagator from a to $-b$. Therefore $G(b, t; a) = G_0(b, t; a) - G_0(-b, t; a)$.

- Sec. 7, page 47, Eq. (7.22).

An even easier way to do the integral. You want to evaluate

$$f(c) = \frac{1}{\sqrt{b}} \int_0^\infty dy \exp\left(-\frac{c^2}{y^2} - y^2\right) \quad \text{Sec. 7, Eq. (7.22)}$$

Rewrite the integrand slightly,

$$f(c) = \frac{1}{\sqrt{b}} \int_0^\infty dy \exp\left[-\left(y - \frac{c}{y}\right)^2 - 2c\right] = \frac{1}{\sqrt{b}} e^{-2c} \psi(c), \quad (62)$$

where $\psi(c) \equiv (1/\sqrt{b}) \int_0^\infty dy \exp\left[-(y - c/y)^2\right]$. Now compute

$$\frac{\partial \psi}{\partial c} = \int dy 2 \left[1 - c/y^2\right] e^{-(y - c/y)^2}. \quad (63)$$

But this derivative is zero, since the individual terms in square bracket give the same integral, as can be seen from the substitution $x = c/y$. To finish, set $c = 0$ to find $\psi(c) = \psi(0) = \sqrt{\pi}$.

- Sec. 9, page 54, prior to Eq. (9.4).

The calculation of $u(j, N)$ of Eq. (9.4) uses the assumptions $\alpha^2/N = O(1)$ and $j^2/N = O(1)$. Note too that for finite N Eq. (9.4) does not hold in the extreme tail of the distribution.

- Sec. 9, page 63, new exercises.

Exercise: You enter a casino with \$100, intent upon walking out with either nothing or \$1000. Assuming you place a series of \$1 bets using a fair coin (tell me where to find this casino . . .), use the methods of this section to compute the probability of each final state.

Exercise: Evaluate $E\left(\frac{[x(t+\delta)-x(t)]^2}{\delta}\right)$.

- Sec. 12, page 81, Eq. (12.9).

For the specific case $L = m\dot{x}^2/2 - V$, Eq. (12.9) takes the simple form $\ddot{\phi} + V''\phi + \lambda\phi = 0$. See Eq. (12.18).

- Sec. 12, page 82, Eqs. (12.17) and (12.18).

The approximation Eq. (12.17) is appropriate for the intended demonstration. But it could be sharpened if for example one wanted to better estimate the interval during which the path is a true minimum. Eq. (12.18) still holds (with $x(t)$ an Euler-Lagrange equation solution), but Eq. (12.19) need not.

Justification of the approximation: If the magnitude of the force, $F(x) \equiv -dV/dx$, is bounded (say $|F| \leq F_{\max}$) then it is easy to show that Eq. (12.17) is good to $O(T^2)$. The equation to solve is $\ddot{x} = -dV/dx$ with $x(0) = a$, $x(T) = b$. Let the true solution be $x(t) = a + (b-a)t/T + u(t)$, so that $u(0) = u(T) = 0$ and $\ddot{u} = F(x)$. We want to bound u . First integrate to $t \leq T$. $\dot{u}(t) = \dot{u}(0) + \int_0^t F(x(s))ds$ (lower limits of integration are 0 throughout). This implies $|\dot{u}(t)| \leq |\dot{u}(0)| + tF_{\max} \leq |\dot{u}(0)| + TF_{\max}$. Since $u(t) = \int_0^t \dot{u}$, $|u(t)| \leq (|\dot{u}(0)| + TF_{\max})T$. It remains to bound $\dot{u}(0)$. Integrate the equation for u twice: $u(t) = t\dot{u}(0) + \int_0^t dt' \int_0^{t'} dt'' F(x)$. Set $t = T$ and use $u(T) = 0$. This implies $|\dot{u}(0)| \leq (1/T) \int_0^T dt' \int_0^{t'} dt'' F_{\max} = F_{\max}T/2$, which shows that $u = O(T^2)$.

- Sec. 12, page 90, “THEOREM.”

In quoting the theorem I neglected to define “strong minimum.” The actual terminology used in Bliss [157] is “strong relative minimum.” The distinction between strong and weak refers to how broad is the class of paths relative to which this is a minimum. Earlier on page 90 I allude to the pitfalls of lapses from full rigor in the subject of variational principles, so that thus warned I won’t try to convey all the subtleties of Bliss’s presentation in these pages. His book, [157],

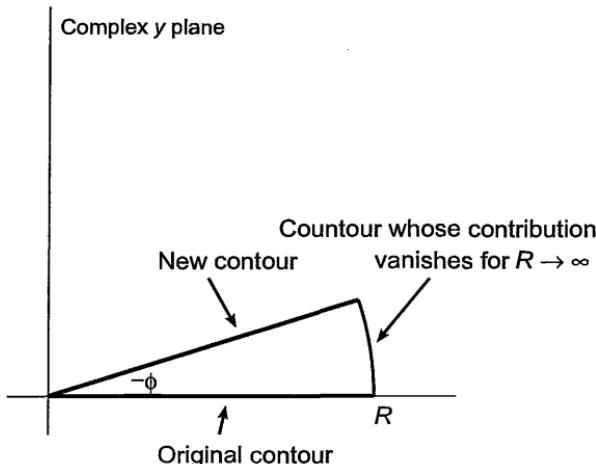


FIG. 3: For the “comment on Sec. 13, page 99.”

is part of a series, the Carus Mathematical Monographs, published by the Mathematical Association of America, which offers readable and enjoyable introductions to a variety of topics.

- *Sec. 13, page 94, material following Eq. (13.8).*

If Secs. 11 and 12 have not been read at this point, it is sufficient to begin at the paragraph preceding Eq. (12.29) and continue through the paragraph containing Eq. (12.45).

- *Sec. 13, page 99, discussion following Eq. (13.20).*

Fig. 3 may be useful for following the argument on page 99. The integrand in (13.20), with $\alpha = e^{i\phi}$, can be made to look like that in (13.12), but with the path of integration along the ray 0 to $\infty \cdot e^{-i\phi}$ (the “New contour”).

- *Sec. 13, page 104, paragraph following Eq. (13.42).*

Concerning the slope of the curves in Fig. 13.1, for small E and for the lowest curve, $T(E)$ is given by $(2\pi/\omega)[1 - (3/4)(\lambda E/\omega^4) + O(E^2)]$. The T versus E curve can be much richer than is illustrated in Fig. 13.1, and can bend around, connecting paths with different numbers of turning points. Doing the Exercise on page 105 will illustrate this point. Examples are given in [158].

- Sec. 14, page 110, Eq. (14.12).

The σ defined here and that defined in Sec. 6 differ by a factor $\epsilon/i\hbar$, i.e.,

$$\sigma_{\text{Sec. } 6} = \frac{\epsilon}{i\hbar} \sigma_{\text{Sec. } 14}.$$

- Sec. 18, page 155, Eqs. (18.27), (18.28a) and (18.28b).

A trivial example with which to see the interplay of path absence and imaginary time is the free particle. Eq. (7.18) has the same form as Eq. (18.27). An exact integration yields [Eq. (7.20)] $\tilde{G}(E) = \sqrt{m/2E} \exp(i|x|\sqrt{2mE/\hbar})$, but one could as well arrive at this result by a stationary phase approximation, with the stationary point in time [for Eq. (18.28)] given by $t_{\text{stationary}} = -x\sqrt{m/2E}$. Now consider $E < 0$, for which the free particle does not have classical paths. This gives us what we want: pure imaginary “stationary” time and exponential dropoff in space.

- Sec. 21, page 173.

Integrals of the form $\int_0^T \int_0^T$ and $\int_0^T \int_0^t$ make their appearance on this and nearby pages. Moreover they come with factors $\pm 2^{\pm n}$. A confession: I cannot promise you that I got it all right. I recently did a calculation of this sort and found that not only could I not rely on my own presentation, but as far as I could tell Feynman gave different versions in different places, and that other authors generally are not explicit about these terms. In my own recent calculation I could only be sure when I found a different way of getting the system’s spectrum and it agreed with whatever signs and factors I’d finally convinced myself to use (it also helped that Denis Tolkunov worked with me on this).

- Sec. 21, page 181

The reference given for the work of Donsker and Varadhan may not be easily accessible. A web search for their names plus “large deviations” will turn up many sources. The fourth in their series of articles is in the journal Comm. Pure Appl. Math. vol. 36 (1983), 183–212. I believe a significant stimulus to the large deviation quest was work of Friedberg and Luttinger [159] on disordered systems, which, at a heuristic level, arrived at results similar to those of Donsker and Varadhan.

- Sec. 22, page 188, just after the first paragraph.

Feynman suggested the use of time-ordered products for the Dirac equation in [160]. But he apparently did not consider them practical since according to [161] they would demand delving into “the geometrical mysteries involved in the superposition of hypercomplex numbers.”

- Sec. 23, page 195, paragraph beginning, “The degree of pathology...” In this paragraph, “entertaining” properties of the Jacobi θ -function are exhibited, based on its zeros. But you need to know that there is at least one zero per quasi-periodic cell:

Exercise: Show that $\theta_3(z_0, t) = 0$, with $z_0 = \pi(1 + t)/2$.

The fundamental theorem for θ -functions is Eq. (23.13). It is worth going through its derivation:

Exercise: The Poisson summation formula (for appropriate f) is

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{k=-\infty}^{\infty} \exp(2\pi i kx) \int_{-\infty}^{\infty} dy f(y) \exp(-2\pi i ky).$$

Find a function, f , such that this formula leads to Eq. (23.13).

For the solution to this exercise as well as a great deal more entertainment, I recommend the book of Bellman [162], where the field of elliptic functions is heralded not only as entertaining, but enchanting and wondrous as well. (See also the end of Sec. III.)

- Sec. 23, page 197, after the theorem.

In the proof I show one of the properties of an equivalence relation, transitivity. The other properties are reflexivity, $f \sim f$, and symmetry, $f \sim g \Rightarrow g \sim f$. If this is your first exposure to homotopy theory, it would be worth writing down the steps to prove these other properties.

- Sec. 23, page 203, **Definition** and THEOREM.

For an illustration of a covering projection see Fig. 23.1, page 193.

- Sec. 23, page 211, second paragraph of Sec. 23.4.

I needed to be more careful in my statement. The condition that the function must vanish at the endpoints of the interval (plus certain regularity conditions) *does* define H sufficiently to make it self adjoint. See [163], §119. But when you do not demand that the function vanish, there emerges the additional richness discussed in this section, in particular the absence of essential self adjointness and the possibility of defining the operator in many ways.

- Sec. 24, page 220, following Eq. (24.25)

For the record, the Laplacian on $SU(2)$ is

$$\Delta = \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \left(\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \psi^2} - 2 \cos \theta \frac{\partial^2}{\partial \phi \partial \psi} \right).$$

- Sec. 25, page 234, after line 5 (end of continuing paragraph).

Continuing the discussion of this paragraph, it's worth being more explicit. For Kruskal coordinates, $u \sim \cosh(t/4M)$ and $v \sim \sinh(t/4M)$. Consider $\exp(t/4M)$ with $t \rightarrow t - i\pi/\kappa$. Then $\exp(t/4M) \rightarrow -\exp(t/4M)$, since $\kappa = 1/4M$. Thus $u \rightarrow -u$ and $v \rightarrow -v$. The point $(-u, -v)$ is B' in Fig. 25.4.

- Sec. 27, page 244, Eq. (27.9).

Here is a proof of Eq. (27.9): Let $[[A, B], A] = [[A, B], B] = 0$. Then

$$e^{A+B} = e^A e^B e^{-[A,B]/2} = e^B e^A e^{+[B,A]/2} \quad \text{Sec. 27, Eq. (27.9)}$$

Lemma 1. $[A, B^n] = nCB^{n-1}$, where $C \equiv [A, B]$.

Proof by induction: $[A, B^{n+1}] = B[A, B^n] + [A, B]B^n$, which by the inductive hypothesis is $BnCB^n + [A, B]B^n$, and which by our assumptions and definitions is $(n+1)CB^n$. The case $n = 1$ is obviously true.

Note that $[B, A^n] = -nCA^{n-1}$, by interchange of letters.

Lemma 2. $[A, e^{\lambda B}] = \lambda C e^{\lambda B}$

To prove this, expand the exponential, use Lemma 1, and re-sum.

Now define

$$g(x) \equiv e^{-xA} e^{x(A+B)} e^{-xB}.$$

Take dg/dx . By straightforward manipulations and the use of Lemma 2 this becomes $dg/dx = e^{-xA}[B, e^{x(A+B)}]e^{-xB}$, which by Lemma 2 (with an interchange of letters) leads to $\frac{dg}{dx} = -xCg$ (also using the fact that C commutes with everything). Now integrate the differential equation from 0 to 1 to get $g(1) = \exp(-C/2)$. This is the desired result (proof attributed to Glauber).

I think of Eq. (27.9) in terms of the Baker-Campbell-Hausdorff series. This is an expression for $\exp(\lambda(A+B))$ in the case where $[A, B]$ need not commute with A and B . In this series, terms with powers higher than λ^2 all involve commutators of A and B with $[A, B]$. So Eq. (27.9) is simply a truncation of the Baker-Campbell-Hausdorff series. For information on (including a derivation of) this series see [164–166].

- Sec. 27, page 245, Exercise.

Solution to the exercise: One possibility is $v(z) = 1 - |z|^2/2$. By Eq. (27.12) it is only necessary to show that $\int \exp(-|z|^2/2) z^n v(z) d^2z = 0$, for all n . Polar coordinates will help.

- Sec. 27, page 247, discussion following Eq. (27.21).

Cross reference: the propagator “obtained earlier in this volume” is Eq. (6.38). Note that the “very easy” way to do this path integral only works for the harmonic oscillator.

- Sec. 27, page 248, just after Eq. (27.28).

Eq. (27.28) is often written without explicit “time-slicing” as

$$G(z_f, t; z_i) = \int_{z_i, 0}^{z_f, t} d^2 z(\cdot) \exp \left\{ i \int_0^t d\tau \left[\frac{i}{2} \left(z^* \frac{dz}{d\tau} - \frac{dz^*}{d\tau} z \right) - H(z^*, z) \right] \right\}$$

- Sec. 28, Eq. (28.3).

For completeness I give an indication of how Eq. (28.3), the expected number of particles in each state, is derived. For a fixed number, N , of particles, the *canonical* partition function is $Z_N = \sum_{\alpha} \exp(-\beta E_{\alpha})$, where α labels states and I have dropped the superscript (C) specifying the sphere-configuration that appears in Eq. (28.3). (Recall that the canonical partition function pertains to a system in contact with a heat reservoir at temperature $1/\beta$.) Since this is an ideal gas (no interaction between particles), its states can be labeled $\alpha = (n_1, \dots, n_k, \dots)$, where n_k is the number of atoms in the one-particle state k . The total energy is $E_{\alpha} = \sum_k n_k \epsilon_k$, with ϵ_k the k^{th} one-particle energy. The constraint $N = \sum_k n_k$ makes the partition function sum difficult so one goes to the grand canonical ensemble. For the latter ensemble one has a reservoir of particles with chemical potential μ and partition function $Z^{GC} = \sum_N Z_N e^{\beta \mu N}$. The probability of finding a particular state and a particular number of particles is thus $e^{-\beta E_{\alpha} + \beta \mu N} / Z^{GC}$. The chemical potential is then fixed so that the weaker requirement, $N = \sum_k \langle n_k \rangle$, is satisfied, with the expectation taken with respect to the probability just indicated. One can now calculate the expected occupation of a particular *one-particle* state:

$$\langle n_k \rangle = \sum_{N, \alpha} n_k \Pr(\text{state-}k \text{ occupied by } n_k \text{ particles}) = \frac{-1}{\beta} \frac{\partial}{\partial \epsilon_k} \log Z^{GC}.$$

All sums are geometric and one gets

$$\langle n_k \rangle = \frac{1}{\beta} \frac{\partial}{\partial \epsilon_k} \log \left(1 - e^{-\beta(\epsilon_k - \mu)} \right) = \frac{1}{e^{\beta(\epsilon_k - \mu)} - 1}.$$

- Sec. 23, p. 195 (out of order).

Solution to the new exercise for page 195: $f(x) = \exp(-ix^2/\pi t)$.

References and endnotes for the supplements

- [1] Those with a mathematician's aesthetic will recoil at the idea of going through a tedious expansion to derive Eq. (13), after having done a similar calculation to derive Eq. (6). The product on the right-hand-side of Eq. (13) can be written $\exp(\lambda B/2) \exp(\lambda A/2) \times \exp(\lambda A/2) \exp(\lambda B/2)$, with the first and second pairs giving $\exp(\lambda(A+B)/2)$ by Eq. (6), and with corrections of order λ^2 . Because the order of A and B is reversed in the two pairs, these corrections will be equal but of opposite sign. Putting the corrections next to each other, which can be done to this level of precision, allows the $O(\lambda^2)$ corrections to cancel, leaving an error that is $O(\lambda^3)$.
- [2] R. P. Feynman, *Space-Time Approach to Non-Relativistic Quantum Mechanics*, Rev. Mod. Phys. **20**, 367-387 (1948).
- [3] B. Gaveau, E. Mihokova, M. Roncadelli, and L. S. Schulman, *Path integral in a magnetic field using the Trotter product formula*, Am. J. Phys. **72**, 385-388 (2004).
- [4] B. Gaveau and L. S. Schulman, *Sensitive terms in the path integral: ordering and stochastic options*, J. Math. Phys. **30**, 2019-2022 (1989).
- [5] L. S. Schulman, *Selected Topics in Path Integration* (1993), in [112].
- [6] W. Janke and T. Sauer, *Properties of higher-order Trotter formulas*, Phys. Lett. A **165**, 199-205 (1992).
- [7] L. S. Schulman, *Introduction to the Path Integral* (1988), in [167].
- [8] D. W. McLaughlin and L. S. Schulman, *Path Integrals in Curved Spaces*, J. Math. Phys. **12**, 2520-2524 (1971).
- [9] T. Ichinose and H. Tamura, *On the Norm Convergence of the Trotter-Kato Product Formula with Error Bound*, Operator Theory: Adv. Appl. **126**, 149-154 (2001).
- [10] J. Hull, *Options, Futures, and Other Derivative Securities* (Prentice Hall, Englewood Cliffs, New Jersey, 1993), 2nd ed., chapter 10.
- [11] B. B. Mandelbrot, *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk* (Springer, New York, 1997).
- [12] B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman and Company, New York, 1983).
- [13] J.-P. Bouchaud and M. Potters, *Theory of Financial Risks* (Cambridge Univ. Press, Cambridge, 2000).
- [14] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics* (Cambridge Univ. Press, Cambridge, 1999).
- [15] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics & Polymer Physics, & Financial Markets* (World Sci., Singapore, 2004), 3rd ed.
- [16] L. F. Abbott and M. B. Wise, *The Dimension of a Quantum-Mechanical Path*, Am. J. Phys. **49**, 37-39 (1981).
- [17] F. Cannata and L. Ferrari, *Dimensions of relativistic quantum mechanical paths*, Am. J. Phys. **56**, 721-725 (1988).

- [18] H. Kroger, *Proposal for an experiment to measure the Hausdorff dimension of quantum-mechanical trajectories*, Phys. Rev. A **55**, 951-966 (1997).
- [19] H. Kroger, *Generalized Aharonov-Bohm effect, homotopy classes and Hausdorff dimension*, Phys. Lett. A **226**, 127-134 (1997).
- [20] H. Kroger, S. Lantagne, K. J. M. Moriarty, and B. Plache, *Measuring the Hausdorff dimension of quantum mechanical paths*, Phys. Lett. A **199**, 299-303 (1995).
- [21] It may be overkill, but let me explain this point in detail. Imagine that for fixed N (the time-slicing parameter) we also discretize space and write the path integral as an enormous sum. For small enough mesh (and large enough spatial cutoff) this sum approaches the integral, because a finite slicing of the path integral is an ordinary Riemann integral. Each contribution to the sum is labeled by N positions and one time. We want to define sets of labels such that the last time the particle is to the left of 0 is t_m . Then this set must have its x_{m+1} values to the right of 0. This breakup partitions the entire set of labels into disjoint sets whose union is the entire set of labels..
- [22] In Eq. (23), I omit the factor $\exp(iV\epsilon/\hbar)$. A single $O(\epsilon)$ contribution can be discarded without affecting the result.
- [23] A. Auerbach, S. Kivelson, and D. Nicole, *Path Decomposition for Multidimensional Tunneling*, Phys. Rev. Lett. **53**, 411-414 (1984).
- [24] A. Auerbach and S. Kivelson, *The Path Decomposition Expansion and Multidimensional Tunneling*, Nucl. Phys. B **257**, 799-858 (1985).
- [25] P. van Baal and A. Auerbach, *An Analysis of Transverse Fluctuations in Multidimensional Tunneling*, Nucl. Phys. B **275**, 93-120 (1986).
- [26] M. Goodman, *Path integral solution to the infinite square well*, Am. J. Phys. **49**, 843-847 (1981).
- [27] A. Auerbach and L. S. Schulman, *A path decomposition expansion (PDX) proof for the method of images*, J. Phys. A **30**, 5993-5995 (1997).
- [28] P. van Baal, *Tunneling and the Path Decomposition Expansion*, in [112].
- [29] J. J. Halliwell, *An Operator Derivation of the Path Decomposition Expansion*, Phys. Lett. A **207**, 237-242 (1995).
- [30] J. J. Halliwell and M. E. Ortiz, *Sum-Over-Histories Origin of the Composition Laws of Relativistic Quantum Mechanics*, Phys. Rev. D **48**, 748-768 (1993).
- [31] L. S. Schulman and R. W. Ziolkowski, *Path Integral Asymptotics in the Absence of Classical Paths* (1989), in [145].
- [32] N. Yamada, *Speakable and Unspeakable in the Tunneling Time Problem*, Phys. Rev. Lett. **83**, 3350-3353 (1999).
- [33] N. Yamada, *Unified Derivation of Tunneling Times from Decoherence Functionals*, Phys. Rev. Lett. **93**, 170401 (2004).
- [34] D. Sokolovski and L. M. Baskin, *Traversal time in quantum scattering*, Phys. Rev. A **36**, 4604-4611 (1987).

- [35] H. A. Fertig, *Traversal-Time Distribution and the Uncertainty Principle in Quantum Tunneling*, Phys. Rev. Lett. **65**, 2321-2324 (1990).
- [36] Kac's telegrapher equation model consists of a particle on a line that moves with fixed speed, $|v|$, but at random, Poisson distributed, times reverses the sign of its velocity. This process is identical to that of the Dirac equation, à la Feynman, except for the imaginary rate. The telegrapher equation is second order in time, with dissipation. It corresponds to the second order equation one gets by "squaring" the Dirac equation (with an imaginary transition rate there is no dissipation).
- [37] T. Jacobson and L. S. Schulman, *Quantum Stochastics: The Passage from a Relativistic to a Non-Relativistic Path Integral*, J. Phys. A **17**, 375-383 (1984).
- [38] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [39] S. S. Schweber, *Feynman and the visualization of space-time processes*, Rev. Mod. Phys. **58**, 449-508 (1986).
- [40] G. V. Riazanov, *The Feynman Path Integral for the Dirac Equation*, Sov. Phys. JETP **6**, 1107-1113 (1958), J. Exptl. Theoret. Phys. (U.S.S.R.) **33**, 1437-1444 (1957).
- [41] T. Jacobson, *Spinor chain path integral for the Dirac equation*, J. Phys. A **17**, 2433-2451 (1984).
- [42] T. Jacobson, *Random walk representations for spinor and vector propagators*, J. Math. Phys. **26**, 1600-1604 (1985).
- [43] M. Kac, *A Stochastic Model Related to the Telegrapher's Equation*, Rocky Mountain J. Math. **4**, 497 (1974).
- [44] B. Gaveau, T. Jacobson, M. Kac, and L. S. Schulman, *Relativistic Extension of the Analogy between Quantum Mechanics and Brownian Motion*, Phys. Rev. Lett. **53**, 419-422 (1984).
- [45] D. G. C. McKeon and G. N. Ord, *Time Reversal in Stochastic Processes and the Dirac Equation*, Phys. Rev. Lett. **69**, 3-4 (1992).
- [46] G. N. Ord, *A Reformulation of the Feynman Chessboard Model*, J. Stat. Phys. **66**, 647 (1992).
- [47] G. N. Ord and J. A. Gualtieri, *The Feynman Propagator from a Single Path*, Phys. Rev. Lett. **89**, 250403 (2002).
- [48] G. N. Ord and D. G. C. McKeon, *On the Dirac Equation in 3+1 Dimensions*, Ann. Phys. **222**, 244-253 (1993).
- [49] C. DeWitt-Morette and S. K. Foong, *Path integral solutions of wave equations with dissipation*, Phys. Rev. Lett. **62**, 2201-2204 (1989).
- [50] S. K. Foong and S. Kanno, *Properties of the telegrapher's random process with or without a trap*, Stoch. Proc. and Applic. **53**, 147-173 (1994).
- [51] T. Jacobson, *Feynman's Checkerboard and Other Games*, in *Nonlinear Equations in Classical and Quantum Field Theory*, edited by N. Sanchez (Springer, Berlin, 1985), pp. 386-395, Vol. 226, Lecture Notes in Phys.
- [52] D. Mugnai, A. Ranfagni, and R. Ruggeri, *Path-Integral Solution of the*

Telegrapher Equation: An Application to the Tunneling Time Determination, Phys. Rev. Lett. **68**, 259-262 (1992).

- [53] D. Mugnai, A. Ranfagni, R. Ruggeri, and A. Agresti, *Semiclassical Analysis of Traversal Time through Kac's Solution of the Telegrapher's Equation*, Phys. Rev. E **49**, 1771-1774 (1994).
- [54] T. Ichinose, *Path integral for the radial Dirac equation*, J. Math. Phys. **46**, 022103 (2005).
- [55] B. Gaveau and L. S. Schulman, *Dirac Equation Path Integral: Interpreting the Grassmann Variables*, Nuovo Cim. D **11**, 31-51 (1989).
- [56] B. Gaveau and L. S. Schulman, *Grassmann-valued processes for the Weyl and the Dirac equations*, Phys. Rev. D **36**, 1135-1140 (1987).
- [57] B. Gaveau and L. S. Schulman, *A projector path integral for the Dirac equation and the spin derivation of space*, Ann. Phys. **284**, 1-9 (2000).
- [58] As to what "analytic expression" means, well that's more difficult. It should mean expressible in simple and well-known functions, which leaves you to define "simple" . . .
- [59] H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics* (Addison-Wesley, Reading, Mass., 2002), 3rd ed.
- [60] I. H. Duru and H. Kleinert, *Solution of the Path Integral for the H-Atom*, Phys. Lett. B **84**, 185-188 (1979).
- [61] I. H. Duru and H. Kleinert, *Quantum Mechanics of H-Atom from Path Integrals*, Fortschr. Phys. **30**, 401-435 (1982).
- [62] B. Gaveau and L. S. Schulman, *Explicit time-dependent Schrodinger propagators*, J. Phys. A **19**, 1833-1846 (1986).
- [63] L. S. Schulman, *Applications of the Propagator for the Delta Function Potential* (1986), in [144].
- [64] M. J. Goovaerts, A. Babcenko, and J. T. Devreese, *A new expansion method in the Feynman path integral formalism: Application to a one-dimensional delta-function potential*, J. Math. Phys. **14**, 554-669 (1973).
- [65] H. R. Jauslin, *Exact Propagator and eigenfunctions for multistable models with arbitrarily prescribed N lowest eigenvalues*, J. Phys. A **21**, 2337-2350 (1988).
- [66] S. Scarlatti and A. Teta, *Derivation of the time-dependent propagator for the three-dimensional Schrödinger equation with one point interaction*, J. Phys. A **23**, L1033-L1035 (1990).
- [67] J. B. Keller, *Geometrical theory of diffraction*, J. Opt. Soc. Am. **52**, 116-130 (1962).
- [68] L. S. Schulman, *Ray Optics for Diffraction: A Useful Paradox in a Path Integral Context*, in *Wave Particle Dualism*, edited by S. Diner, D. Farague, G. Lochak, and F. Selleri (Reidel, Dordrecht, 1984).
- [69] L. S. Schulman, *Exact time-dependent Green's function for the half-plane barrier*, Phys. Rev. Lett. **49**, 599-601 (1982).
- [70] A similar construction can be made for opaque wedges when the wedge angle is $2\pi/\text{integer}$. The corresponding propagator is also solvable [168].

- [71] F. W. Wiegel, *Introduction to Path-Integral Methods in Physics and Polymer Science* (World Scientific, Singapore, 1986).
- [72] D. P. L. Castrigiano and F. Stärk, *New aspects of the path integrational treatment of the Coulomb potential*, J. Math. Phys. **30**, 2785-2788 (1989).
- [73] R. Ho and A. Inomata, *Exact-Path-Integral Treatment of the Hydrogen Atom*, Phys. Rev. Lett. **48**, 231-234 (1982).
- [74] A. O. Barut, A. Inomata, and G. Junker, *Path integral treatment of the hydrogen atom in a curved space of constant curvature*, J. Phys. A **20**, 6271-6280 (1987).
- [75] J. M. Cai, P. Y. Cai, and A. Inomata, *Path-integral treatment of the Hulthen potential*, Phys. Rev. A **34**, 4621-4628 (1986).
- [76] A. Inomata, *Exact Path-Integration for the Two-Dimensional Coulomb Problem*, Phys. Lett. A **87**, 387-390 (1982).
- [77] A. Young and C. DeWitt-Morette, *Time Substitutions in Stochastic Processes as a Tool in Path Integration*, Ann. Phys. **169**, 140-166 (1986).
- [78] A. O. Caldeira and A. J. Leggett, *Influence of Dissipation on Quantum Tunneling in Macroscopic Systems*, Phys. Rev. Lett. **46**, 211-214 (1981).
- [79] A. O. Caldeira and A. J. Leggett, *Quantum Tunnelling in a Dissipative System*, Ann. Phys. **149**, 374-456 (1983).
- [80] H. Dekker, *Quantum Coherence and Tunnelling in a Double-Well Potential in a Thermal Environment: Dynamics of the Weakly Coupled Spin-Boson System*, Physica **144A**, 453-480 (1987).
- [81] A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger, *Dynamics of the dissipative two-state system*, Rev. Mod. Phys. **59**, 1-85 (1987).
- [82] H. Grabert, P. Schramm, and G.-L. Ingold, *Quantum Brownian Motion: The Functional Integral Approach*, Phys. Rep. **168**, 115-207 (1988).
- [83] T. Dittrich, P. Hänggi, G.-L. Ingold, B. Kramer, G. Schön, and W. Zwerger, eds., *Quantum transport and dissipation* (Wiley-VCH, Weinheim, 1998).
- [84] J. T. Devreese, *Polarons*, Encyclopedia of Applied Physics **14**, 383 (1996), Arxiv listing: cond-mat/0004497.
- [85] F. Brosens, S. N. Klomm, and J. T. Devreese, *Variational path-integral treatment of a translation invariant many-polaron system*, Phys. Rev. B **71**, 214301 (2005).
- [86] J. T. Devreese, *Polarons and bipolarons in oxides*, Z. Phys. B **104**, 601-604 (1997).
- [87] J. Tempere, M. Wouters, and J. T. Devreese, *Path-integral mean-field description of the vortex state in the BEC-to-BCS crossover*, Phys. Rev. A **71**, 033631 (2005).
- [88] D. K. Campbell, S. Flach, and Y. S. Kivshar, *Localizing Energy Through Nonlinearity and Discreteness*, Phys. Today, pp. 43-49 (January 2004).
- [89] L. S. Schulman, D. Tolkinov, and E. Mihokova, *Structure and time-dependence of quantum breathers*, J. Chem. Phys. (2005), to appear.

- [90] T. L. Gill and W. W. Zachary, *Foundations for relativistic quantum theory. I Feynman's operator calculus and the Dyson conjecture*, J. Math. Phys. **43**, 69-93 (2002). See also [150], where other work was reported.
- [91] N. Kumano-go, *Feynman path integrals as analysis on path space by time slicing approximation*, Bull. Sci. Math. **128**, 197-251 (2004).
- [92] D. Fujiwara and N. Kumano-go, *Smooth functional derivatives in Feynman path integrals by time slicing approximation*, Bull. Sci. Math. **129**, 57-79 (2005).
- [93] T. Nakamura, *Path space measure for the 3+1 dimensional Dirac equation in momentum space*, J. Math. Phys. **41**, 5209 (2000).
- [94] A. Robinson, *Non-standard Analysis* (Princeton Univ. Press, Princeton, 1996).
- [95] L. S. Schulman and B. Gaveau, *Checkerboard Path Integral: Use and Extension*, in [148].
- [96] S. M. Flatté, S. A. Reynolds, and R. Dashen, *Path-integral treatment of intensity behavior for rays in a sound channel*, J. Acoust. Soc. Amer. **82**, 967 (1987).
- [97] J. L. Codona, D. B. Creamer, S. M. Flatté, R. G. Frehlich, and F. S. Henyey, *Moment-equation and path-integral techniques for wave propagation in random media*, J. Math. Phys. **27**, 171 (1986).
- [98] S. M. Flatté, *The Schrödinger equation in classical physics*, Am. J. Phys. **54**, 1088 (1986).
- [99] R. B. Schlottmann, *A path integral formulation of acoustic wave propagation*, Geophys. J. Int. **137**, 353-363 (1999).
- [100] L. Fishman, *One-Way Wave Equation Modeling in Two-Way Wave Propagation Problems*, in *Mathematical Modelling of Wave Phenomena 2002*, edited by B. Nilsson and L. Fishman (Vaxjo University Press, Vaxjo, Sweden, 2004).
- [101] L. Fishman, *Applications of Directional Wavefield Decomposition, Phase Space, and Path Integral Methods to Seismic Wave Propagation and Inversion*, Pure Appl. Geophysics **159**, 1637-1679 (2002).
- [102] V. Galdi, V. Pierro, and I. M. Pinto, *Path Integral Computation of Lowest Order Modes in Arbitrary-Shaped Inhomogeneous Waveguides*, IEEE Microwave & Guided Wave Lett. **7**, 402-404 (1997).
- [103] V. Galdi, V. Pierro, and I. M. Pinto, *A generalized Donsker-Kac formula to compute the fundamental modes in complex loaded waveguides*, Electromagnetics **18**, 367-382 (1998).
- [104] B. Budaev and D. Bogy, *Analysis of 1D Wave Scattering by the Random Walk Method*, J. Acoust. Soc. Am. **111**, 2555-2560 (2002).
- [105] L. S. Schulman, *Approximate Topologies*, J. Math. Phys. **12**, 304-308 (1971).
- [106] R. Capozza, D. Giuliano, P. Lucignano, and A. Tagliacozzo, *Quantum interference of electrons in a ring: tuning of the geometrical phase*, preprint (2005), Arxiv listing: cond-mat/0505273.

- [107] C. C. Bernido and A. Inomata, *Path integrals with a periodic constraint: The Aharonov-Bohm effect*, J. Math. Phys. **22**, 715-718 (1981).
- [108] D. Giuliano, P. Sodano, and A. Tagliacozzo, *Adiabatic control of the electron phase in a quantum dot*, Phys. Rev. B **67**, 155317 (2005).
- [109] R. Sundrum and L. J. Tassie, *Nonabelian Aharonov-Bohm effects, Feynman paths, and topology*, J. Math. Phys. **27**, 1566-1570 (1986).
- [110] G. S. Canright and S. M. Girvin, *Fractional Statistics: Quantum Possibilities in Two Dimensions*, Science **247**, 1197-1205 (1990).
- [111] G. S. Canright and M. D. Johnson, *Fractional Statistics: alpha to beta*, J. Phys. A **27**, 3579 (1994).
- [112] H. Cerdeira, S. Lundqvist, D. Mugnai, A. Ranfagni, V. Sa-yakanit, and L. S. Schulman, *Lectures on Path Integration: Trieste 1991* (World Scientific, Singapore, 1993), workshop and conference held at ICTP, Trieste, 1991.
- [113] R. P. Feynman and H. Kleinert, *Effective classical partition functions*, Phys. Rev. A **34**, 5080-5084 (1986).
- [114] R. Giachetti and V. Tognetti, *Variational Approach to Quantum Statistical Mechanics of Nonlinear Systems with Application to Sine-Gordon Chains*, Phys. Rev. Lett. **55**, 912-915 (1985).
- [115] R. Giachetti and V. Tognetti, *Variational approach to the thermodynamics of a quantum sine-Gordon field with out-of-plane fluctuations*, Phys. Rev. B **36**, 5512 (1987).
- [116] H. Kleinert, *Improving the variational approach to path integrals*, Phys. Lett. B **280**, 251-260 (1992).
- [117] R. J. Creswick, *A Numerical Method for Evaluating Feynman Path Integrals*, Mod. Phys. Lett. B **9**, 693-699 (1995). See also Creswick's article in [148].
- [118] A. Garg and M. Stone, *Bohr-Sommerfeld Quantization of Spin Hamiltonians*, Phys. Rev. Lett. **92**, 010401 (2004).
- [119] M. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer-Verlag, New York, 1990).
- [120] P. Cvitanovic, R. Artuso, R. Mainieri, G. Tanner, G. Vattay, N. Whelan, and A. Wirzba, *Chaos: Classical and Quantum* (online, 2004). Can be downloaded from <http://chaosbook.org>.
- [121] S. Tomsovic and E. J. Heller, *Semiclassical Dynamics of Chaotic Motion: Unexpected Long-Time Accuracy*, Phys. Rev. Lett. **67**, 664-667 (1991).
- [122] L. S. Schulman, *Accuracy of the semiclassical approximation for the time dependent propagator*, J. Phys. A **27**, 1703-1721 (1994).
- [123] K. Takatsuka, *Phase-space representation of quantum mechanics and its relation to phase-space path integrals*, Phys. Rev. A **39**, 5961-5973 (1989).
- [124] G. Roepstorff, *Path Integral Approach to Quantum Physics: An Introduction* (Springer, Berlin, 1994).
- [125] M. Kac, *Integration in Function Spaces and Some of Its Applications* (Accademia Nazionale Dei Lincei Scuola Normale Superiore, Pisa, 1980).

- [126] J. Glimm and A. Jaffe, *Quantum Physics: a Functional Integral Point of View* (Springer-Verlag, NY, 1981).
- [127] U. Weiss, *Quantum Dissipative Systems* (World Scientific, Singapore, 1999), 2nd ed.
- [128] M. Roncadelli and A. Defendi, *I Camini di Feynman* (Università degli Studi di Pavia, Pavia, Italy, 1992).
- [129] Y. Choquet-Bruhat and C. DeWitt-Morette, *Analysis, Manifolds and Physics* (Elsevier, Amsterdam, 2000).
- [130] A. Ranfagni, D. Mugnai, P. Moretti, and M. Cetica, *Trajectories and Rays: The Path-Summation in Quantum Mechanics and Optics* (World Scientific, Singapore, 1990).
- [131] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, New York, 1984). See also by the first author, *Functional Integration and Partial Differential Equations* (Princeton Univ. Press, Princeton, 1985).
- [132] F. Steiner and C. Grosche, *A Table of Feynman Path Integrals* (Springer, Berlin, 1995).
- [133] V. N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics* (Reidel, Dordrecht, 1983).
- [134] P. Ramond, *Field Theory: A Modern Primer* (Benjamin, Reading, Mass., 1981).
- [135] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).
- [136] L. D. Faddeev and A. A. Slavnov, *Gauge Fields: Introduction to Quantum Theory* (Benjamin, Reading, Mass., 1980).
- [137] R. J. Rivers, *Path Integral Methods in Quantum Field Theory* (Cambridge Univ. Press, Cambridge, 1987).
- [138] D. M. Ceperley, *Path integrals in the theory of condensed helium*, Rev. Mod. Phys. **67**, 279-355 (1995).
- [139] D. C. Khandekar and S. V. Lawande, *Feynman Path Integrals: Some Exact Results and Applications*, Phys. Rep. **137**, 115-229 (1986).
- [140] P. Gaspard, D. Alonso, and I. Burghardt, *New Ways of Understanding Semiclassical Quantization*, in *Advances in Chemical Physics*, Vol. XC, edited by I. Prigogine and S. A. Rice (Wiley, New York, 1995), pp. 105-360.
- [141] R. G. Littlejohn, *The Semiclassical Evolution of Wave Packets*, Phys. Rep. **138**, 193-291 (1986).
- [142] R. Shankar, *Renormalization-group approach to interacting fermions*, Rev. Mod. Phys. **66**, 129-192 (1994).
- [143] A. Inomata, *Path Integrals from meV to MeV* (1983). Only informal notes are available.
- [144] M. C. Gutzwiller, A. Inomata, J. R. Klauder, and L. Streit, *Path Integrals from meV to MeV* (World Scientific, Singapore, 1986).
- [145] V. Sa-yakanit, W. Srirakool, J.-O. Berananda, M. C. Gutzwiller, A. In-

- omata, S. Lundqvist, J. R. Klauder, and L. S. Schulman, *Path Integrals from meV to MeV* (World Scientific, Singapore, 1989).
- [146] H. Grabert, A. Inomata, L. S. Schulman, and U. Weiss, *Path Integrals from meV to MeV: Tutzing '92* (World Scientific, Singapore, 1993).
- [147] V. Yarunin and M. Smolyrev, *Proc. joint conf. "Path Integrals: Theory and Applications" and "Path Integrals from meV to MeV"* (Joint Inst. Nucl. Res., Dubna, 1996).
- [148] R. Casalbouni, R. Giachetti, V. Tognetti, R. Vaia, and P. Verrucchi, *Path Integrals from peV to TeV: 50 Years after Feynman's Paper* (World Scientific, Singapore, 1999).
- [149] F. Brosens, V. M. Fomin, S. Foulon, and V. N. Gladiln, *Path Integrals from Quarks to Galaxies. Program and Abstracts* (Univ. Antwerp, Antwerp, 2002), with selected papers published in *phys. stat. sol. (b)* **237**(1).
- [150] The eighth conference has already taken place and the proceedings are slated to be published by the Joint Inst. Nucl. Res., Dubna.
- [151] D. C. Khandekar and S. V. Lawande, *Path Integral Methods and their Applications* (Indian Phys. Assoc., Bombay, 1990).
- [152] A. Ranfagni, D. Mugnai, P. Moretti, and M. Cetica, *Feynman's Quantum Mechanics 40 Years after Its Proposal* (Soc. Ital. Fis., Bologna, 1989).
- [153] D. Mugnai, A. Ranfagni, and L. S. Schulman, *Tunneling and its Implications* (World Sci., Singapore, 1997), Proc. Adriatico Research Conf.
- [154] V. Sa-yakanit, J.-O. Berananda, and W. Sitrakool, *Path Integrals in Physics* (World Scientific, Singapore, 1994).
- [155] M. L. Glasser, *Summation Over Feynman Histories: Charged Particle in a Uniform Magnetic Field*, *Phys. Rev.* **133**, B831-B834 (1964).
- [156] L. S. Schulman, *Green's Function for an Electron in a Lattice*, *Phys. Rev.* **188**, 1139-1142 (1969).
- [157] G. A. Bliss, *Calculus of Variations* (Open Court, for the Math. Assoc. of Amer., La Salle, Ill., 1925).
- [158] L. S. Schulman, *An Exact Propagator and its Correspondence Limits at a Caustic*, *Phys. Rev. A* **35**, 4956-4962 (1987).
- [159] R. Friedberg and J. M. Luttinger, *Density of electronic energy levels in disordered systems*, *Phys. Rev. B* **12**, 4460 (1975).
- [160] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, *Phys. Rev.* **80**, 440-457 (1950).
- [161] R. P. Feynman, *An Operator Calculus Having Applications in Quantum Electrodynamics*, *Phys. Rev.* **84**, 108-128 (1951).
- [162] R. Bellman, *A Brief Introduction to Theta Functions* (Holt, Rinehart and Winston, New York, 1961).
- [163] F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Frederick Ungar, New York, 1955); Dover reprinting, 1990.
- [164] M. Hausner and J. T. Schwartz, *Lie Groups Lie Algebras* (Gordon and Breach, New York, 1968).

- [165] G. Hochschild, *The Structure of Lie Groups* (Holden-Day, San Francisco, 1965).
- [166] M. W. Reinsch, *A simple expression for the terms in the Baker-Campbell-Hausdorff series*, J. Math. Phys. **41**, 2434-2442 (2000).
- [167] S. O. Lundqvist, A. Ranfagni, V. Sa-Yakanit, and L. S. Schulman, *Path Summation: Achievements and Goals* (World Sci., Singapore, 1988), Proc. Adriatico Research Conf.
- [168] C. DeWitt-Morette, S. G. Low, L. S. Schulman, and A. Y. Shiekh, *Wedges I*, Found. Phys. **16**, 311-349 (1986), Wheeler Festschrift.

* * *

See pages 349–359 for the indexes to this book.

The table of contents for the supplement is on page 362.

A CATALOG OF SELECTED
DOVER BOOKS
IN SCIENCE AND MATHEMATICS



CATALOG OF DOVER BOOKS

Astronomy

BURNHAM'S CELESTIAL HANDBOOK, Robert Burnham, Jr. Thorough guide to the stars beyond our solar system. Exhaustive treatment. Alphabetical by constellation: Andromeda to Cetus in Vol. 1; Chamaeleon to Orion in Vol. 2; and Pavo to Vulpecula in Vol. 3. Hundreds of illustrations. Index in Vol. 3. 2,000pp. 6 $\frac{1}{2}$ x 9 $\frac{1}{4}$.

Vol. I: 0-486-23567-X

Vol. II: 0-486-23568-8

Vol. III: 0-486-23673-0

EXPLORING THE MOON THROUGH BINOCULARS AND SMALL TELESCOPES, Ernest H. Cherrington, Jr. Informative, profusely illustrated guide to locating and identifying craters, rills, seas, mountains, other lunar features. Newly revised and updated with special section of new photos. Over 100 photos and diagrams. 240pp. 8 $\frac{1}{2}$ x 11.

0-486-24491-1

THE EXTRATERRESTRIAL LIFE DEBATE, 1750-1900, Michael J. Crowe. First detailed, scholarly study in English of the many ideas that developed from 1750 to 1900 regarding the existence of intelligent extraterrestrial life. Examines ideas of Kant, Herschel, Voltaire, Percival Lowell, many other scientists and thinkers. 16 illustrations. 704pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$.

0-486-40675-X

THEORIES OF THE WORLD FROM ANTIQUITY TO THE COPERNICAN REVOLUTION, Michael J. Crowe. Newly revised edition of an accessible, enlightening book recreates the change from an earth-centered to a sun-centered conception of the solar system. 242pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$.

0-486-41444-2

A HISTORY OF ASTRONOMY, A. Pannekoek. Well-balanced, carefully reasoned study covers such topics as Ptolemaic theory, work of Copernicus, Kepler, Newton, Eddington's work on stars, much more. Illustrated. References. 521pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$.

0-486-65994-1

A COMPLETE MANUAL OF AMATEUR ASTRONOMY: TOOLS AND TECHNIQUES FOR ASTRONOMICAL OBSERVATIONS, P. Clay Sherrod with Thomas L. Koed. Concise, highly readable book discusses: selecting, setting up and maintaining a telescope; amateur studies of the sun; lunar topography and occultations; observations of Mars, Jupiter, Saturn, the minor planets and the stars; an introduction to photoelectric photometry; more. 1981 ed. 124 figures. 25 halftones. 37 tables. 335pp. 6 $\frac{1}{2}$ x 9 $\frac{1}{4}$.

0-486-40675-X

AMATEUR ASTRONOMER'S HANDBOOK, J. B. Sidgwick. Timeless, comprehensive coverage of telescopes, mirrors, lenses, mountings, telescope drives, micrometers, spectroscopes, more. 189 illustrations. 576pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. (Available in U.S. only.)

0-486-24034-7

STARS AND RELATIVITY, Ya. B. Zel'dovich and I. D. Novikov. Vol. 1 of *Relativistic Astrophysics* by famed Russian scientists. General relativity, properties of matter under astrophysical conditions, stars, and stellar systems. Deep physical insights, clear presentation. 1971 edition. References. 544pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$.

0-486-69424-0

CATALOG OF DOVER BOOKS

Chemistry

THE SCEPTICAL CHYMIST: THE CLASSIC 1661 TEXT, Robert Boyle. Boyle defines the term "element," asserting that all natural phenomena can be explained by the motion and organization of primary particles. 1911 ed. viii+232pp. 5% x 8%. 0-486-42825-7

RADIOACTIVE SUBSTANCES, Marie Curie. Here is the celebrated scientist's doctoral thesis, the prelude to her receipt of the 1903 Nobel Prize. Curie discusses establishing atomic character of radioactivity found in compounds of uranium and thorium; extraction from pitchblende of polonium and radium; isolation of pure radium chloride; determination of atomic weight of radium; plus electric, photographic, luminous, heat, color effects of radioactivity. ii+94pp. 5% x 8%. 0-486-42550-9

CHEMICAL MAGIC, Leonard A. Ford. Second Edition, Revised by E. Winston Grundmeier. Over 100 unusual stunts demonstrating cold fire, dust explosions, much more. Text explains scientific principles and stresses safety precautions. 128pp. 5% x 8%. 0-486-67628-5

THE DEVELOPMENT OF MODERN CHEMISTRY, Aaron J. Ihde. Authoritative history of chemistry from ancient Greek theory to 20th-century innovation. Covers major chemists and their discoveries. 209 illustrations. 14 tables. Bibliographies. Indices. Appendices. 851pp. 5% x 8%. 0-486-64235-6

CATALYSIS IN CHEMISTRY AND ENZYMOLOGY, William P. Jencks. Exceptionally clear coverage of mechanisms for catalysis, forces in aqueous solution, carbonyl- and acyl-group reactions, practical kinetics, more. 864pp. 5% x 8%. 0-486-65460-5

ELEMENTS OF CHEMISTRY, Antoine Lavoisier. Monumental classic by founder of modern chemistry in remarkable reprint of rare 1790 Kerr translation. A must for every student of chemistry or the history of science. 539pp. 5% x 8%. 0-486-64624-6

THE HISTORICAL BACKGROUND OF CHEMISTRY, Henry M. Leicester. Evolution of ideas, not individual biography. Concentrates on formulation of a coherent set of chemical laws. 260pp. 5% x 8%. 0-486-61053-5

A SHORT HISTORY OF CHEMISTRY, J. R. Partington. Classic exposition explores origins of chemistry, alchemy, early medical chemistry, nature of atmosphere, theory of valency, laws and structure of atomic theory, much more. 428pp. 5% x 8%. (Available in U.S. only.) 0-486-65977-1

GENERAL CHEMISTRY, Linus Pauling. Revised 3rd edition of classic first-year text by Nobel laureate. Atomic and molecular structure, quantum mechanics, statistical mechanics, thermodynamics correlated with descriptive chemistry. Problems. 992pp. 5% x 8%. 0-486-65622-5

FROM ALCHEMY TO CHEMISTRY, John Read. Broad, humanistic treatment focuses on great figures of chemistry and ideas that revolutionized the science. 50 illustrations. 240pp. 5% x 8%. 0-486-28690-8

CATALOG OF DOVER BOOKS

Engineering

DE RE METALLICA, Georgius Agricola. The famous Hoover translation of greatest treatise on technological chemistry, engineering, geology, mining of early modern times (1556). All 289 original woodcuts. 6 $\frac{1}{2}$ x 11. 0-486-60006-8

FUNDAMENTALS OF ASTRODYNAMICS, Roger Bate et al. Modern approach developed by U.S. Air Force Academy. Designed as a first course. Problems, exercises. Numerous illustrations. 455pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-60061-0

DYNAMICS OF FLUIDS IN POROUS MEDIA, Jacob Bear. For advanced students of ground water hydrology, soil mechanics and physics, drainage and irrigation engineering and more. 335 illustrations. Exercises, with answers. 784pp. 6 $\frac{1}{2}$ x 9 $\frac{1}{4}$. 0-486-65675-6

THEORY OF VISCOELASTICITY (Second Edition), Richard M. Christensen. Complete consistent description of the linear theory of the viscoelastic behavior of materials. Problem-solving techniques discussed. 1982 edition. 29 figures. xiv+364pp. 6 $\frac{1}{2}$ x 9 $\frac{1}{4}$. 0-486-42880-X

MECHANICS, J. P. Den Hartog. A classic introductory text or refresher. Hundreds of applications and design problems illuminate fundamentals of trusses, loaded beams and cables, etc. 334 answered problems. 462pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-60754-2

MECHANICAL VIBRATIONS, J. P. Den Hartog. Classic textbook offers lucid explanations and illustrative models, applying theories of vibrations to a variety of practical industrial engineering problems. Numerous figures. 233 problems, solutions. Appendix. Index. Preface. 436pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-64785-4

STRENGTH OF MATERIALS, J. P. Den Hartog. Full, clear treatment of basic material (tension, torsion, bending, etc.) plus advanced material on engineering methods, applications. 350 answered problems. 323pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-60755-0

A HISTORY OF MECHANICS, René Dugas. Monumental study of mechanical principles from antiquity to quantum mechanics. Contributions of ancient Greeks, Galileo, Leonardo, Kepler, Lagrange, many others. 671pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-65632-2

STABILITY THEORY AND ITS APPLICATIONS TO STRUCTURAL MECHANICS, Clive L. Dym. Self-contained text focuses on Koiter postbuckling analyses, with mathematical notions of stability of motion. Basing minimum energy principles for static stability upon dynamic concepts of stability of motion, it develops asymptotic buckling and postbuckling analyses from potential energy considerations, with applications to columns, plates, and arches. 1974 ed. 208pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-42541-X

METAL FATIGUE, N. E. Frost, K. J. Marsh, and L. P. Pook. Definitive, clearly written, and well-illustrated volume addresses all aspects of the subject, from the historical development of understanding metal fatigue to vital concepts of the cyclic stress that causes a crack to grow. Includes 7 appendixes. 544pp. 5 $\frac{1}{2}$ x 8 $\frac{1}{2}$. 0-486-40927-9

CATALOG OF DOVER BOOKS

ROCKETS, Robert Goddard. Two of the most significant publications in the history of rocketry and jet propulsion: "A Method of Reaching Extreme Altitudes" (1919) and "Liquid Propellant Rocket Development" (1936). 128pp. 5% x 8%. 0-486-42537-1

STATISTICAL MECHANICS: PRINCIPLES AND APPLICATIONS, Terrell L. Hill. Standard text covers fundamentals of statistical mechanics, applications to fluctuation theory, imperfect gases, distribution functions, more. 448pp. 5% x 8%. 0-486-65390-0

ENGINEERING AND TECHNOLOGY 1650-1750: ILLUSTRATIONS AND TEXTS FROM ORIGINAL SOURCES, Martin Jensen. Highly readable text with more than 200 contemporary drawings and detailed engravings of engineering projects dealing with surveying, leveling, materials, hand tools, lifting equipment, transport and erection, piling, bailing, water supply, hydraulic engineering, and more. Among the specific projects outlined—transporting a 50-ton stone to the Louvre, erecting an obelisk, building timber locks, and dredging canals. 207pp. 8% x 11%. 0-486-42232-1

THE VARIATIONAL PRINCIPLES OF MECHANICS, Cornelius Lanczos. Graduate level coverage of calculus of variations, equations of motion, relativistic mechanics, more. First inexpensive paperbound edition of classic treatise. Index. Bibliography. 418pp. 5% x 8%. 0-486-65067-7

PROTECTION OF ELECTRONIC CIRCUITS FROM OVERVOLTAGES, Ronald B. Standler. Five-part treatment presents practical rules and strategies for circuits designed to protect electronic systems from damage by transient overvoltages. 1989 ed. xxiv+434pp. 6% x 9%. 0-486-42552-5

ROTARY WING AERODYNAMICS, W. Z. Stepniewski. Clear, concise text covers aerodynamic phenomena of the rotor and offers guidelines for helicopter performance evaluation. Originally prepared for NASA. 537 figures. 640pp. 6% x 9%. 0-486-64647-5

INTRODUCTION TO SPACE DYNAMICS, William Tyrrell Thomson. Comprehensive, classic introduction to space-flight engineering for advanced undergraduate and graduate students. Includes vector algebra, kinematics, transformation of coordinates. Bibliography. Index. 352pp. 5% x 8%. 0-486-65113-4

HISTORY OF STRENGTH OF MATERIALS, Stephen P. Timoshenko. Excellent historical survey of the strength of materials with many references to the theories of elasticity and structure. 245 figures. 452pp. 5% x 8%. 0-486-61187-6

ANALYTICAL FRACTURE MECHANICS, David J. Unger. Self-contained text supplements standard fracture mechanics texts by focusing on analytical methods for determining crack-tip stress and strain fields. 336pp. 6% x 9%. 0-486-41737-9

STATISTICAL MECHANICS OF ELASTICITY, J. H. Weiner. Advanced, self-contained treatment illustrates general principles and elastic behavior of solids. Part 1, based on classical mechanics, studies thermoelastic behavior of crystalline and polymeric solids. Part 2, based on quantum mechanics, focuses on interatomic force laws, behavior of solids, and thermally activated processes. For students of physics and chemistry and for polymer physicists. 1983 ed. 96 figures. 496pp. 5% x 8%. 0-486-42260-7

CATALOG OF DOVER BOOKS

Mathematics

FUNCTIONAL ANALYSIS (Second Corrected Edition), George Bachman and Lawrence Narici. Excellent treatment of subject geared toward students with background in linear algebra, advanced calculus, physics and engineering. Text covers introduction to inner-product spaces, normed, metric spaces, and topological spaces; complete orthonormal sets, the Hahn-Banach Theorem and its consequences, and many other related subjects. 1966 ed. 544pp. 6 $\frac{1}{4}$ x 9 $\frac{1}{4}$. 0-486-40251-7

ASYMPTOTIC EXPANSIONS OF INTEGRALS, Norman Bleistein & Richard A. Handelsman. Best introduction to important field with applications in a variety of scientific disciplines. New preface. Problems. Diagrams. Tables. Bibliography. Index. 448pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-65082-0

VECTOR AND TENSOR ANALYSIS WITH APPLICATIONS, A. I. Borisenko and I. E. Tarapov. Concise introduction. Worked-out problems, solutions, exercises. 257pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-63833-2

AN INTRODUCTION TO ORDINARY DIFFERENTIAL EQUATIONS, Earl A. Coddington. A thorough and systematic first course in elementary differential equations for undergraduates in mathematics and science, with many exercises and problems (with answers). Index. 304pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-65942-9

FOURIER SERIES AND ORTHOGONAL FUNCTIONS, Harry F. Davis. An incisive text combining theory and practical example to introduce Fourier series, orthogonal functions and applications of the Fourier method to boundary-value problems. 570 exercises. Answers and notes. 416pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-65973-9

COMPUTABILITY AND UNSOLVABILITY, Martin Davis. Classic graduate-level introduction to theory of computability, usually referred to as theory of recursive functions. New preface and appendix. 288pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-61471-9

ASYMPTOTIC METHODS IN ANALYSIS, N. G. de Bruijn. An inexpensive, comprehensive guide to asymptotic methods—the pioneering work that teaches by explaining worked examples in detail. Index. 224pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-64221-6

APPLIED COMPLEX VARIABLES, John W. Dettman. Step-by-step coverage of fundamentals of analytic function theory—plus lucid exposition of five important applications: Potential Theory; Ordinary Differential Equations; Fourier Transforms; Laplace Transforms; Asymptotic Expansions. 66 figures. Exercises at chapter ends. 512pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-64670-X

INTRODUCTION TO LINEAR ALGEBRA AND DIFFERENTIAL EQUATIONS, John W. Dettman. Excellent text covers complex numbers, determinants, orthonormal bases, Laplace transforms, much more. Exercises with solutions. Undergraduate level. 416pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-65191-6

RIEMANN'S ZETA FUNCTION, H. M. Edwards. Superb, high-level study of landmark 1859 publication entitled "On the Number of Primes Less Than a Given Magnitude" traces developments in mathematical theory that it inspired. xiv+315pp. 5 $\frac{1}{4}$ x 8 $\frac{1}{4}$. 0-486-41740-9

CATALOG OF DOVER BOOKS

CALCULUS OF VARIATIONS WITH APPLICATIONS, George M. Ewing. Applications-oriented introduction to variational theory develops insight and promotes understanding of specialized books, research papers. Suitable for advanced undergraduate/graduate students as primary, supplementary text. 352pp. 5% x 8 $\frac{1}{2}$.
0-486-64856-7

COMPLEX VARIABLES, Francis J. Flanigan. Unusual approach, delaying complex algebra till harmonic functions have been analyzed from real variable viewpoint. Includes problems with answers. 364pp. 5% x 8 $\frac{1}{2}$.
0-486-61388-7

AN INTRODUCTION TO THE CALCULUS OF VARIATIONS, Charles Fox. Graduate-level text covers variations of an integral, isoperimetrical problems, least action, special relativity, approximations, more. References. 279pp. 5% x 8 $\frac{1}{2}$.
0-486-65499-0

COUNTEREXAMPLES IN ANALYSIS, Bernard R. Gelbaum and John M. H. Olmsted. These counterexamples deal mostly with the part of analysis known as "real variables." The first half covers the real number system, and the second half encompasses higher dimensions. 1962 edition. xxiv+198pp. 5% x 8 $\frac{1}{2}$. 0-486-42875-3

CATASTROPHE THEORY FOR SCIENTISTS AND ENGINEERS, Robert Gilmore. Advanced-level treatment describes mathematics of theory grounded in the work of Poincaré, R. Thom, other mathematicians. Also important applications to problems in mathematics, physics, chemistry and engineering. 1981 edition. References. 28 tables. 397 black-and-white illustrations. xvii + 666pp. 6 $\frac{1}{2}$ x 9 $\frac{1}{4}$.
0-486-67539-4

INTRODUCTION TO DIFFERENCE EQUATIONS, Samuel Goldberg. Exceptionally clear exposition of important discipline with applications to sociology, psychology, economics. Many illustrative examples; over 250 problems. 260pp. 5% x 8 $\frac{1}{2}$.
0-486-65084-7

NUMERICAL METHODS FOR SCIENTISTS AND ENGINEERS, Richard Hamming. Classic text stresses frequency approach in coverage of algorithms, polynomial approximation, Fourier approximation, exponential approximation, other topics. Revised and enlarged 2nd edition. 721pp. 5% x 8 $\frac{1}{2}$.
0-486-65241-6

INTRODUCTION TO NUMERICAL ANALYSIS (2nd Edition), F. B. Hildebrand. Classic, fundamental treatment covers computation, approximation, interpolation, numerical differentiation and integration, other topics. 150 new problems. 669pp. 5% x 8 $\frac{1}{2}$.
0-486-65363-3

THREE PEARLS OF NUMBER THEORY, A. Y. Khinchin. Three compelling puzzles require proof of a basic law governing the world of numbers. Challenges concern van der Waerden's theorem, the Landau-Schnirelmann hypothesis and Mann's theorem, and a solution to Waring's problem. Solutions included. 64pp. 5% x 8 $\frac{1}{2}$.
0-486-40026-3

THE PHILOSOPHY OF MATHEMATICS: AN INTRODUCTORY ESSAY, Stephan Körner. Surveys the views of Plato, Aristotle, Leibniz & Kant concerning propositions and theories of applied and pure mathematics. Introduction. Two appendices. Index. 198pp. 5% x 8 $\frac{1}{2}$.
0-486-25048-2

CATALOG OF DOVER BOOKS

INTRODUCTORY REAL ANALYSIS, A.N. Kolmogorov, S. V. Fomin. Translated by Richard A. Silverman. Self-contained, evenly paced introduction to real and functional analysis. Some 350 problems. 403pp. 5% x 8%. 0-486-61226-0

APPLIED ANALYSIS, Cornelius Lanczos. Classic work on analysis and design of finite processes for approximating solution of analytical problems. Algebraic equations, matrices, harmonic analysis, quadrature methods, much more. 559pp. 5% x 8%. 0-486-65656-X

AN INTRODUCTION TO ALGEBRAIC STRUCTURES, Joseph Landin. Superb self-contained text covers "abstract algebra": sets and numbers, theory of groups, theory of rings, much more. Numerous well-chosen examples, exercises. 247pp. 5% x 8%. 0-486-65940-2

QUALITATIVE THEORY OF DIFFERENTIAL EQUATIONS, V. V. Nemytskii and V.V. Stepanov. Classic graduate-level text by two prominent Soviet mathematicians covers classical differential equations as well as topological dynamics and ergodic theory. Bibliographies. 523pp. 5% x 8%. 0-486-65954-2

THEORY OF MATRICES, Sam Perlis. Outstanding text covering rank, nonsingularity and inverses in connection with the development of canonical matrices under the relation of equivalence, and without the intervention of determinants. Includes exercises. 237pp. 5% x 8%. 0-486-66810-X

INTRODUCTION TO ANALYSIS, Maxwell Rosenlicht. Unusually clear, accessible coverage of set theory, real number system, metric spaces, continuous functions, Riemann integration, multiple integrals, more. Wide range of problems. Undergraduate level. Bibliography. 254pp. 5% x 8%. 0-486-65038-3

MODERN NONLINEAR EQUATIONS, Thomas L. Saaty. Emphasizes practical solution of problems; covers seven types of equations. "... a welcome contribution to the existing literature...."—*Math Reviews*. 490pp. 5% x 8%. 0-486-64232-1

MATRICES AND LINEAR ALGEBRA, Hans Schneider and George Phillip Barker. Basic textbook covers theory of matrices and its applications to systems of linear equations and related topics such as determinants, eigenvalues and differential equations. Numerous exercises. 432pp. 5% x 8%. 0-486-66014-1

LINEAR ALGEBRA, Georgi E. Shilov. Determinants, linear spaces, matrix algebras, similar topics. For advanced undergraduates, graduates. Silverman translation. 387pp. 5% x 8%. 0-486-63518-X

ELEMENTS OF REAL ANALYSIS, David A. Sprecher. Classic text covers fundamental concepts, real number system, point sets, functions of a real variable, Fourier series, much more. Over 500 exercises. 352pp. 5% x 8%. 0-486-65385-4

SET THEORY AND LOGIC, Robert R. Stoll. Lucid introduction to unified theory of mathematical concepts. Set theory and logic seen as tools for conceptual understanding of real number system. 496pp. 5% x 8%. 0-486-63829-4

CATALOG OF DOVER BOOKS

TENSOR CALCULUS, J.L. Synge and A. Schild. Widely used introductory text covers spaces and tensors, basic operations in Riemannian space, non-Riemannian spaces, etc. 324pp. 5% x 8%. 0-486-63612-7

ORDINARY DIFFERENTIAL EQUATIONS, Morris Tenenbaum and Harry Pollard. Exhaustive survey of ordinary differential equations for undergraduates in mathematics, engineering, science. Thorough analysis of theorems. Diagrams. Bibliography. Index. 818pp. 5% x 8%. 0-486-64940-7

INTEGRAL EQUATIONS, F. G. Tricomi. Authoritative, well-written treatment of extremely useful mathematical tool with wide applications. Volterra Equations, Fredholm Equations, much more. Advanced undergraduate to graduate level. Exercises. Bibliography. 238pp. 5% x 8%. 0-486-64828-1

FOURIER SERIES, Georgi P. Tolstov. Translated by Richard A. Silverman. A valuable addition to the literature on the subject, moving clearly from subject to subject and theorem to theorem. 107 problems, answers. 336pp. 5% x 8%. 0-486-63317-9

INTRODUCTION TO MATHEMATICAL THINKING, Friedrich Waismann. Examinations of arithmetic, geometry, and theory of integers; rational and natural numbers; complete induction; limit and point of accumulation; remarkable curves; complex and hypercomplex numbers, more. 1959 ed. 27 figures. xii+260pp. 5% x 8%. 0-486-63317-9

POPULAR LECTURES ON MATHEMATICAL LOGIC, Hao Wang. Noted logician's lucid treatment of historical developments, set theory, model theory, recursion theory and constructivism, proof theory, more. 3 appendixes. Bibliography. 1981 edition. ix + 283pp. 5% x 8%. 0-486-67632-3

CALCULUS OF VARIATIONS, Robert Weinstock. Basic introduction covering isoperimetric problems, theory of elasticity, quantum mechanics, electrostatics, etc. Exercises throughout. 326pp. 5% x 8%. 0-486-63069-2

THE CONTINUUM: A CRITICAL EXAMINATION OF THE FOUNDATION OF ANALYSIS, Hermann Weyl. Classic of 20th-century foundational research deals with the conceptual problem posed by the continuum. 156pp. 5% x 8%. 0-486-67982-9

CHALLENGING MATHEMATICAL PROBLEMS WITH ELEMENTARY SOLUTIONS, A. M. Yaglom and I. M. Yaglom. Over 170 challenging problems on probability theory, combinatorial analysis, points and lines, topology, convex polygons, many other topics. Solutions. Total of 445pp. 5% x 8%. Two-vol. set.

Vol. I: 0-486-65536-9 Vol. II: 0-486-65537-7

INTRODUCTION TO PARTIAL DIFFERENTIAL EQUATIONS WITH APPLICATIONS, E. C. Zachmanoglou and Dale W. Thoe. Essentials of partial differential equations applied to common problems in engineering and the physical sciences. Problems and answers. 416pp. 5% x 8%. 0-486-65251-3

THE THEORY OF GROUPS, Hans J. Zassenhaus. Well-written graduate-level text acquaints reader with group-theoretic methods and demonstrates their usefulness in mathematics. Axioms, the calculus of complexes, homomorphic mapping, p -group theory, more. 276pp. 5% x 8%. 0-486-40922-8

Math—Decision Theory, Statistics, Probability

ELEMENTARY DECISION THEORY, Herman Chernoff and Lincoln E. Moses. Clear introduction to statistics and statistical theory covers data processing, probability and random variables, testing hypotheses, much more. Exercises. 364pp. 5% x 8½. 0-486-65218-1

STATISTICS MANUAL, Edwin L. Crow et al. Comprehensive, practical collection of classical and modern methods prepared by U.S. Naval Ordnance Test Station. Stress on use. Basics of statistics assumed. 288pp. 5% x 8½. 0-486-60599-X

SOME THEORY OF SAMPLING, William Edwards Deming. Analysis of the problems, theory and design of sampling techniques for social scientists, industrial managers and others who find statistics important at work. 61 tables. 90 figures. xvii +602pp. 5% x 8½. 0-486-64684-X

LINEAR PROGRAMMING AND ECONOMIC ANALYSIS, Robert Dorfman, Paul A. Samuelson and Robert M. Solow. First comprehensive treatment of linear programming in standard economic analysis. Game theory, modern welfare economics, Leontief input-output, more. 525pp. 5% x 8½. 0-486-65491-5

PROBABILITY: AN INTRODUCTION, Samuel Goldberg. Excellent basic text covers set theory, probability theory for finite sample spaces, binomial theorem, much more. 360 problems. Bibliographies. 322pp. 5% x 8½. 0-486-65252-1

GAMES AND DECISIONS: INTRODUCTION AND CRITICAL SURVEY, R. Duncan Luce and Howard Raiffa. Superb nontechnical introduction to game theory, primarily applied to social sciences. Utility theory, zero-sum games, n-person games, decision-making, much more. Bibliography. 509pp. 5% x 8½. 0-486-65943-7

INTRODUCTION TO THE THEORY OF GAMES, J. C. C. McKinsey. This comprehensive overview of the mathematical theory of games illustrates applications to situations involving conflicts of interest, including economic, social, political, and military contexts. Appropriate for advanced undergraduate and graduate courses; advanced calculus a prerequisite. 1952 ed. x+372pp. 5% x 8½. 0-486-42811-7

FIFTY CHALLENGING PROBLEMS IN PROBABILITY WITH SOLUTIONS, Frederick Mosteller. Remarkable puzzlers, graded in difficulty, illustrate elementary and advanced aspects of probability. Detailed solutions. 88pp. 5% x 8½. 65355-2

PROBABILITY THEORY: A CONCISE COURSE, Y. A. Rozanov. Highly readable, self-contained introduction covers combination of events, dependent events, Bernoulli trials, etc. 148pp. 5% x 8½. 0-486-63544-9

STATISTICAL METHOD FROM THE VIEWPOINT OF QUALITY CONTROL, Walter A. Shewhart. Important text explains regulation of variables, uses of statistical control to achieve quality control in industry, agriculture, other areas. 192pp. 5% x 8½. 0-486-65232-7

Math—Geometry and Topology

ELEMENTARY CONCEPTS OF TOPOLOGY, Paul Alexandroff. Elegant, intuitive approach to topology from set-theoretic topology to Betti groups; how concepts of topology are useful in math and physics. 25 figures. 57pp. 5% x 8%. 0-486-60747-X

COMBINATORIAL TOPOLOGY, P. S. Alexandrov. Clearly written, well-organized, three-part text begins by dealing with certain classic problems without using the formal techniques of homology theory and advances to the central concept, the Betti groups. Numerous detailed examples. 654pp. 5% x 8%. 0-486-40179-0

EXPERIMENTS IN TOPOLOGY, Stephen Barr. Classic, lively explanation of one of the byways of mathematics. Klein bottles, Moebius strips, projective planes, map coloring, problem of the Koenigsberg bridges, much more, described with clarity and wit. 43 figures. 210pp. 5% x 8%. 0-486-25933-1

THE GEOMETRY OF RENÉ DESCARTES, René Descartes. The great work founded analytical geometry. Original French text, Descartes's own diagrams, together with definitive Smith-Latham translation. 244pp. 5% x 8%. 0-486-60068-8

EUCLIDEAN GEOMETRY AND TRANSFORMATIONS, Clayton W. Dodge. This introduction to Euclidean geometry emphasizes transformations, particularly isometries and similarities. Suitable for undergraduate courses, it includes numerous examples, many with detailed answers. 1972 ed. viii+296pp. 6% x 9%. 0-486-43476-1

PRACTICAL CONIC SECTIONS: THE GEOMETRIC PROPERTIES OF ELLIPSES, PARABOLAS AND HYPERBOLAS, J. W. Downs. This text shows how to create ellipses, parabolas, and hyperbolas. It also presents historical background on their ancient origins and describes the reflective properties and roles of curves in design applications. 1993 ed. 98 figures. xii+100pp. 6% x 9%. 0-486-42876-1

THE THIRTEEN BOOKS OF EUCLID'S ELEMENTS, translated with introduction and commentary by Sir Thomas L. Heath. Definitive edition. Textual and linguistic notes, mathematical analysis. 2,500 years of critical commentary. Unabridged. 1,414pp. 5% x 8%. Three-vol. set.

Vol. I: 0-486-60088-2 Vol. II: 0-486-60089-0 Vol. III: 0-486-60090-4

SPACE AND GEOMETRY: IN THE LIGHT OF PHYSIOLOGICAL, PSYCHOLOGICAL AND PHYSICAL INQUIRY, Ernst Mach. Three essays by an eminent philosopher and scientist explore the nature, origin, and development of our concepts of space, with a distinctness and precision suitable for undergraduate students and other readers. 1906 ed. vi+148pp. 5% x 8%. 0-486-43909-7

GEOMETRY OF COMPLEX NUMBERS, Hans Schwerdtfeger. Illuminating, widely praised book on analytic geometry of circles, the Moebius transformation, and two-dimensional non-Euclidean geometries. 200pp. 5% x 8%. 0-486-63830-8

DIFFERENTIAL GEOMETRY, Heinrich W. Guggenheimer. Local differential geometry as an application of advanced calculus and linear algebra. Curvature, transformation groups, surfaces, more. Exercises. 62 figures. 378pp. 5% x 8%. 0-486-63433-7

CATALOG OF DOVER BOOKS

History of Math

THE WORKS OF ARCHIMEDES, Archimedes (T. L. Heath, ed.). Topics include the famous problems of the ratio of the areas of a cylinder and an inscribed sphere; the measurement of a circle; the properties of conoids, spheroids, and spirals; and the quadrature of the parabola. Informative introduction. clxxxvi+326pp. 5% x 8½.
0-486-42084-1

A SHORT ACCOUNT OF THE HISTORY OF MATHEMATICS, W. W. Rouse Ball. One of clearest, most authoritative surveys from the Egyptians and Phoenicians through 19th-century figures such as Grassman, Galois, Riemann. Fourth edition. 522pp. 5% x 8½.
0-486-20630-0

THE HISTORY OF THE CALCULUS AND ITS CONCEPTUAL DEVELOPMENT, Carl B. Boyer. Origins in antiquity, medieval contributions, work of Newton, Leibniz, rigorous formulation. Treatment is verbal. 346pp. 5% x 8½.
0-486-60509-4

THE HISTORICAL ROOTS OF ELEMENTARY MATHEMATICS, Lucas N. H. Bunt, Phillip S. Jones, and Jack D. Bedient. Fundamental underpinnings of modern arithmetic, algebra, geometry and number systems derived from ancient civilizations. 320pp. 5% x 8½.
0-486-25563-8

A HISTORY OF MATHEMATICAL NOTATIONS, Florian Cajori. This classic study notes the first appearance of a mathematical symbol and its origin, the competition it encountered, its spread among writers in different countries, its rise to popularity, its eventual decline or ultimate survival. Original 1929 two-volume edition presented here in one volume. xxviii+820pp. 5% x 8½.
0-486-67766-4

GAMES, GODS & GAMBLING: A HISTORY OF PROBABILITY AND STATISTICAL IDEAS, F. N. David. Episodes from the lives of Galileo, Fermat, Pascal, and others illustrate this fascinating account of the roots of mathematics. Features thought-provoking references to classics, archaeology, biography, poetry. 1962 edition. 304pp. 5% x 8½. (Available in U.S. only)
0-486-40023-9

OF MEN AND NUMBERS: THE STORY OF THE GREAT MATHEMATICIANS, Jane Muir. Fascinating accounts of the lives and accomplishments of history's greatest mathematical minds—Pythagoras, Descartes, Euler, Pascal, Cantor, many more. Anecdotal, illuminating. 30 diagrams. Bibliography. 256pp. 5% x 8½.
0-486-28973-7

HISTORY OF MATHEMATICS, David E. Smith. Nontechnical survey from ancient Greece and Orient to late 19th century; evolution of arithmetic, geometry, trigonometry, calculating devices, algebra, the calculus. 362 illustrations. 1,355pp. 5% x 8½. Two-vol. set.
Vol. I: 0-486-20429-4 Vol. II: 0-486-20430-8

A CONCISE HISTORY OF MATHEMATICS, Dirk J. Struik. The best brief history of mathematics. Stresses origins and covers every major figure from ancient Near East to 19th century. 41 illustrations. 195pp. 5% x 8½.
0-486-60255-9

CATALOG OF DOVER BOOKS

Physics

OPTICAL RESONANCE AND TWO-LEVEL ATOMS, L. Allen and J. H. Eberly. Clear, comprehensive introduction to basic principles behind all quantum optical resonance phenomena. 53 illustrations. Preface. Index. 256pp. 5% x 8%. 0-486-65533-4

QUANTUM THEORY, David Bohm. This advanced undergraduate-level text presents the quantum theory in terms of qualitative and imaginative concepts, followed by specific applications worked out in mathematical detail. Preface. Index. 655pp. 5% x 8%. 0-486-65969-0

ATOMIC PHYSICS (8th EDITION), Max Born. Nobel laureate's lucid treatment of kinetic theory of gases, elementary particles, nuclear atom, wave-corpuscles, atomic structure and spectral lines, much more. Over 40 appendices, bibliography. 495pp. 5% x 8%. 0-486-65984-4

A SOPHISTICATE'S PRIMER OF RELATIVITY, P. W. Bridgman. Geared toward readers already acquainted with special relativity, this book transcends the view of theory as a working tool to answer natural questions: What is a frame of reference? What is a "law of nature"? What is the role of the "observer"? Extensive treatment, written in terms accessible to those without a scientific background. 1983 ed. xlviii+172pp. 5% x 8%. 0-486-42549-5

AN INTRODUCTION TO HAMILTONIAN OPTICS, H. A. Buchdahl. Detailed account of the Hamiltonian treatment of aberration theory in geometrical optics. Many classes of optical systems defined in terms of the symmetries they possess. Problems with detailed solutions. 1970 edition. xv + 360pp. 5% x 8%. 0-486-67597-1

PRIMER OF QUANTUM MECHANICS, Marvin Chester. Introductory text examines the classical quantum bead on a track: its state and representations; operator eigenvalues; harmonic oscillator and bound bead in a symmetric force field; and bead in a spherical shell. Other topics include spin, matrices, and the structure of quantum mechanics; the simplest atom; indistinguishable particles; and stationary-state perturbation theory. 1992 ed. xiv+314pp. 6% x 9%. 0-486-42878-8

LECTURES ON QUANTUM MECHANICS, Paul A. M. Dirac. Four concise, brilliant lectures on mathematical methods in quantum mechanics from Nobel Prize-winning quantum pioneer build on idea of visualizing quantum theory through the use of classical mechanics. 96pp. 5% x 8%. 0-486-41713-1

THIRTY YEARS THAT SHOOK PHYSICS: THE STORY OF QUANTUM THEORY, George Gamow. Lucid, accessible introduction to influential theory of energy and matter. Careful explanations of Dirac's anti-particles, Bohr's model of the atom, much more. 12 plates. Numerous drawings. 240pp. 5% x 8%. 0-486-24895-X

ELECTRONIC STRUCTURE AND THE PROPERTIES OF SOLIDS: THE PHYSICS OF THE CHEMICAL BOND, Walter A. Harrison. Innovative text offers basic understanding of the electronic structure of covalent and ionic solids, simple metals, transition metals and their compounds. Problems. 1980 edition. 582pp. 6% x 9%. 0-486-66021-4

CATALOG OF DOVER BOOKS

HYDRODYNAMIC AND HYDROMAGNETIC STABILITY, S. Chandrasekhar. Lucid examination of the Rayleigh-Benard problem; clear coverage of the theory of instabilities causing convection. 704pp. 5% x 8%. 0-486-64071-X

INVESTIGATIONS ON THE THEORY OF THE BROWNIAN MOVEMENT, Albert Einstein. Five papers (1905-8) investigating dynamics of Brownian motion and evolving elementary theory. Notes by R. Fürth. 122pp. 5% x 8%. 0-486-60304-0

THE PHYSICS OF WAVES, William C. Elmore and Mark A. Heald. Unique overview of classical wave theory. Acoustics, optics, electromagnetic radiation, more. Ideal as classroom text or for self-study. Problems. 477pp. 5% x 8%. 0-486-64926-1

GRAVITY, George Gamow. Distinguished physicist and teacher takes reader-friendly look at three scientists whose work unlocked many of the mysteries behind the laws of physics: Galileo, Newton, and Einstein. Most of the book focuses on Newton's ideas, with a concluding chapter on post-Einsteinian speculations concerning the relationship between gravity and other physical phenomena. 160pp. 5% x 8%. 0-486-42563-0

PHYSICAL PRINCIPLES OF THE QUANTUM THEORY, Werner Heisenberg. Nobel Laureate discusses quantum theory, uncertainty, wave mechanics, work of Dirac, Schroedinger, Compton, Wilson, Einstein, etc. 184pp. 5% x 8%. 0-486-60113-7

ATOMIC SPECTRA AND ATOMIC STRUCTURE, Gerhard Herzberg. One of best introductions; especially for specialist in other fields. Treatment is physical rather than mathematical. 80 illustrations. 257pp. 5% x 8%. 0-486-60115-3

AN INTRODUCTION TO STATISTICAL THERMODYNAMICS, Terrell L. Hill. Excellent basic text offers wide-ranging coverage of quantum statistical mechanics, systems of interacting molecules, quantum statistics, more. 523pp. 5% x 8%. 0-486-65242-4

THEORETICAL PHYSICS, Georg Joos, with Ira M. Freeman. Classic overview covers essential math, mechanics, electromagnetic theory, thermodynamics, quantum mechanics, nuclear physics, other topics. First paperback edition. xxiii + 885pp. 5% x 8%. 0-486-65227-0

PROBLEMS AND SOLUTIONS IN QUANTUM CHEMISTRY AND PHYSICS, Charles S. Johnson, Jr. and Lee G. Pedersen. Unusually varied problems, detailed solutions in coverage of quantum mechanics, wave mechanics, angular momentum, molecular spectroscopy, more. 280 problems plus 139 supplementary exercises. 430pp. 6% x 9%. 0-486-65236-X

THEORETICAL SOLID STATE PHYSICS, Vol. I: Perfect Lattices in Equilibrium; Vol. II: Non-Equilibrium and Disorder, William Jones and Norman H. March. Monumental reference work covers fundamental theory of equilibrium properties of perfect crystalline solids, non-equilibrium properties, defects and disordered systems. Appendices. Problems. Preface. Diagrams. Index. Bibliography. Total of 1,301pp. 5% x 8%. Two volumes. Vol. I: 0-486-65015-4 Vol. II: 0-486-65016-2

WHAT IS RELATIVITY? L. D. Landau and G. B. Rumer. Written by a Nobel Prize physicist and his distinguished colleague, this compelling book explains the special theory of relativity to readers with no scientific background, using such familiar objects as trains, rulers, and clocks. 1960 ed. vi+72pp. 5% x 8%. 0-486-42806-0

CATALOG OF DOVER BOOKS

A TREATISE ON ELECTRICITY AND MAGNETISM, James Clerk Maxwell. Important foundation work of modern physics. Brings to final form Maxwell's theory of electromagnetism and rigorously derives his general equations of field theory. 1,084pp. 5% x 8%. Two-vol. set. Vol. I: 0-486-60636-8 Vol. II: 0-486-60637-6

QUANTUM MECHANICS: PRINCIPLES AND FORMALISM, Roy McWeeny. Graduate student-oriented volume develops subject as fundamental discipline, opening with review of origins of Schrödinger's equations and vector spaces. Focusing on main principles of quantum mechanics and their immediate consequences, it concludes with final generalizations covering alternative "languages" or representations. 1972 ed. 15 figures. xi+155pp. 5% x 8%. 0-486-42829-X

INTRODUCTION TO QUANTUM MECHANICS With Applications to Chemistry, Linus Pauling & E. Bright Wilson, Jr. Classic undergraduate text by Nobel Prize winner applies quantum mechanics to chemical and physical problems. Numerous tables and figures enhance the text. Chapter bibliographies. Appendices. Index. 468pp. 5% x 8%. 0-486-64871-0

METHODS OF THERMODYNAMICS, Howard Reiss. Outstanding text focuses on physical technique of thermodynamics, typical problem areas of understanding, and significance and use of thermodynamic potential. 1965 edition. 238pp. 5% x 8%. 0-486-69445-3

THE ELECTROMAGNETIC FIELD, Albert Shadowitz. Comprehensive undergraduate text covers basics of electric and magnetic fields, builds up to electromagnetic theory. Also related topics, including relativity. Over 900 problems. 768pp. 5% x 8%. 0-486-65660-8

GREAT EXPERIMENTS IN PHYSICS: FIRSTHAND ACCOUNTS FROM GALILEO TO EINSTEIN, Morris H. Shamos (ed.). 25 crucial discoveries: Newton's laws of motion, Chadwick's study of the neutron, Hertz on electromagnetic waves, more. Original accounts clearly annotated. 370pp. 5% x 8%. 0-486-25346-5

EINSTEIN'S LEGACY, Julian Schwinger. A Nobel Laureate relates fascinating story of Einstein and development of relativity theory in well-illustrated, nontechnical volume. Subjects include meaning of time, paradoxes of space travel, gravity and its effect on light, non-Euclidean geometry and curving of space-time, impact of radio astronomy and space-age discoveries, and more. 189 b/w illustrations. xiv+250pp. 8% x 9%. 0-486-41974-6

STATISTICAL PHYSICS, Gregory H. Wannier. Classic text combines thermodynamics, statistical mechanics and kinetic theory in one unified presentation of thermal physics. Problems with solutions. Bibliography. 532pp. 5% x 8%. 0-486-65401-X

Paperbound unless otherwise indicated. Available at your book dealer, online at www.doverpublications.com, or by writing to Dept. GI, Dover Publications, Inc., 31 East 2nd Street, Mineola, NY 11501. For current price information or for free catalogues (please indicate field of interest), write to Dover Publications or log on to www.doverpublications.com and see every Dover book in print. Dover publishes more than 500 books each year on science, elementary and advanced mathematics, biology, music, art, literary history, social sciences, and other areas.