This book presents beam physics using a unified approach, emphasizing basic concepts and analysis methods. While many existing resources in beams and accelerators are specialized to aid the professional practitioner, this text anticipates the needs of physics students. The central concepts underpinning the physics of accelerators, charged particle, and photon beams are built up from familiar, intertwining components, such as electromagnetism, relativity, and Hamiltonian dynamics. These components are woven into an illustrative set of examples that allow investigation of a variety of physical scenarios. With these tools, single particle dynamics in linear accelerators are discussed, with general methods that are naturally extended to circular accelerators. Beyond single particle dynamics, the proliferation of commonly used beam descriptions are surveyed and compared. These methods provide a powerful connection between the classical charged particle beams, and beams based on coherent waves – laser beams. Aspects of experimental techniques are introduced. Numerous exercises, and examples drawn from devices such as synchrotrons and free-electron lasers, are included to illustrate relevant physical principles.

James B. Rosenzweig is Professor of Physics at the University of California, Los Angeles, where he is also Director of the Particle Beam Physics Laboratory.

Also published by Oxford University Press

The physics of particle accelerators: an introduction
Klaus Wille

An introduction to particle accelerators
Edmund Wilson

OXFORD
UNIVERSITY PRESS

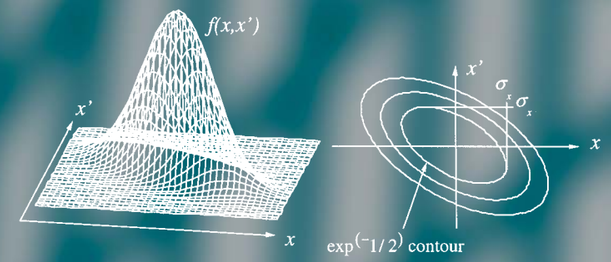
www.oup.com

ISBN 0-19-852554-0
9 780198 525547

Fundamentals of beam physics
Rosenzweig
OXFORD

Fundamentals of beam physics

James B. Rosenzweig



OXFORD

This is a selection from the beginning of the text. The full text is available through the publisher, Oxford University press.

Fundamentals of Beam Physics

J.B. ROSENZWEIG

Department of Physics and Astronomy
University of California, Los Angeles

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Oxford University Press, 2003

The moral rights of the author have been asserted

Database right Oxford University Press (maker)

First published 2003

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

A catalogue record for this title is available from the British Library

Library of Congress Cataloging in Publication Data
(Data available)

ISBN 0 19 852554 0

10 9 8 7 6 5 4 3 2 1

Typeset by Newgen Imaging Systems (P) Ltd., Chennai, India
Printed in Great Britain
on acid-free paper by The Bath Press, Avon

Preface

This book was developed as part of my efforts to build and enhance the accelerator and beam physics program in the UCLA Department of Physics and Astronomy. This program includes an active research program in advanced accelerators (particle accelerators based on lasers and/or plasmas) and free-electron lasers (conversely, lasers based on accelerators and particle beams) and is essentially a student-oriented enterprise. Its focus is primarily on graduate education and training, but the laboratory activities of the group also include a strong showing of undergraduate students. The desire to introduce both incoming graduate students and interested undergraduates to the physics of beams led me to develop a course, Physics 150, to formally provide the background needed to enter the field.

In teaching this senior-level course the first few times, I found that I was relying on a combination of excerpts from a variety of accelerator and laser physics texts and my own notes. Because the hybrid nature of the UCLA research program in beams is reflected in the course, it was simply not possible to use a single text for the course's source material. As one might imagine, in the mixing of written references, the notation, level, and assumed background varied widely from reading to reading. In addition, many existing texts and references in this area are geared towards a practitioner of accelerator physics working at a major accelerator facility. Thus this type of text has an emphasis that is heavy on the physics, engineering methods, and technical jargon specific to large accelerators at the high-energy frontier. The needs of this professional reader are inherently a bit different than the senior level university student, however. As such, previous texts typically have given less orientation to basic physics concepts than the university student needs in order to be properly introduced to the subject. My desire to clarify the written introduction of beam physics as it is practiced at UCLA to undergraduate students led directly to the production of this book.

The contents of this book were also flavored by my desire to create the compact introduction to beam physics that I wish I could have had—hopefully the reader will benefit from the resulting weight put on the points I have found to be least clear or intuitive in my journeys in the field. The present book is therefore written with the student constantly in mind, and has been structured to give a unified discussion of a variety of subjects that may seem to be, on the surface, disparate. The intent of the book is to provide a coherent introduction to the ideas and concepts behind the physics of particle beams. As such, the book begins, after some introductory historical and conceptual comments, with a review of relativity and mechanics. This discussion is intended to build up our sets of physics tools, by placing a few standard approaches to modern dynamics, such as Hamiltonian and phase space-based analyses, in the context

of relativistic motion. We then give a presentation of charged particle dynamics in various combinations of simple magnetostatic and electrostatic field configurations, providing another unifying set of basic tools for understanding more complex scenarios in beam physics. Also, at a higher conceptual level, we examine the physics of circular accelerators using simple extensions of the principles developed first in the context linear accelerators, and then use similar approaches to analyzing both transverse optics and acceleration (longitudinal) dynamics. The adopted emphasis on fundamental, unified tools is motivated by the challenges of modern accelerators and their applications as encountered in the laboratory. In present, state-of-the-art beam physics labs, the experimental systems display an increasing wealth of physical phenomena, that require a physicist's insight to understand.

One of the more unique aspects of this text is that its unified approach is extended to include a discussion of the connection between the methods of charged particle beam optics and descriptions of the physics of paraxial light beams such as lasers. This unification of concepts, between wave-based light beams and classical charged particle beams, is also motivated by experimental challenges arising from two complementary sources. The first is that lasers are increasingly used as critical components in accelerators—for example, they are used to produce intense, picosecond electron pulses in devices termed photoinjectors, two of which are found at UCLA. The second was already mentioned above. In advanced accelerators and free-electron lasers, the concepts of accelerators, particle beams, and lasers are, in fact, merged. These cutting edge subjects in beam physics are what provide the intellectual impetus behind my research program at UCLA. It should not be surprising that such subjects find their way into this introductory text in a number of different ways.

This book is also structured so that an abbreviated course consisting of the first four chapters may give an introduction to single particle dynamics in accelerators. Additionally, Chapter 5 introduces the physics of beam distributions (collections of many particles), a subject that is quite necessary if one wishes to apply this text in practice. Further, the notions developed in Chapters 1–5 can then be used to give the basis for the material on photon beams in Chapter 8. Chapters 6 and 7, which discuss the technical subjects of magnets as well as waveguides and accelerator cavities, can stand virtually by themselves. They do, however, complete the set of basic material offered here, and it is hoped that they prove useful in practice as an introductory guide to the design and use of accelerator laboratory components.

To aid in streamlining the approach to learning from this text, sections that contain general "review" material are marked with the $\star$ symbol next to their title. For an advanced student, these sections (the reviews of relativity and mechanics fall into this category) might be omitted on first reading. Other sections are marked by an asterix (*), and contain material which can be considered in some way tangential to the main exposition of topics in beam physics. Such sections, while forming important components of the book as a whole, and may be referenced elsewhere in the text, may contain material that is too lengthy or deep for a fast initial reading. In an alternative approach, much of the material in both these special sections would be included in appendices. Here they are included in the main body of text, both to improve the logical flow of the book, and to allow illustrative exercises for the student to be included.

The exercises in this book, included at the end of each chapter, are meant to be an integral part of the exposition, as some important topics are actually covered in the exercises. In order to provide a guide to approaching the exercises, and to help emphasize their importance, worked solutions to roughly one-third of the problems are included in Appendix A.

The subjects introduced in this book are related in both obvious and subtle ways. To aid in tying threads of the text together for the reader, a short summary is included at each chapter end.

Numerous acknowledgments are in order. One must first find your interest sparked by a field of inquiry; my present colleague David Cline provided the spark for me when I was yet a student. My initial training in accelerator theory was most heavily influenced by Fred Mills (then at Fermilab), and his style can be seen in many of the approaches taken to analyzing beam dynamics in this text. Other friends, mentors and collaborators of particular note, who have given me stimulus to go deeper into the subjects presented here are: Pisin Chen, Richard Cooper, Luca Serafini, and Jim Simpson. A special debt of gratitude is due to my close colleague, Claudio Pellegrini, who has, with our students and post-docs, built the UCLA beam physics program with me.

Perhaps the highest level of thanks must go, however to those students at UCLA who have provided components of the background, motivation, and critical feedback to finish this text. A no-doubt incomplete listing of the graduate students (many of whom are now professional colleagues) follows: Ron Agusstson, Scott Anderson, Gerard Andonian, Kip Bishofberger, Salime Boucher, Nick Barov, Eric Colby, Xiadong Ding, Joel England, Spencer Hartman, Mark Hogan, Pietro Musumeci, Sven Reiche, Soren Telfer, Andrei Terebilo, Matt Thompson, Gil Travish and Aaron Tremaine. Special thanks are due to the undergraduate students who have aided in the editing of this text, Pauline Lay and Maria Perrelli.

As a university professor with a large, active research program, my textbook writing has generally been performed after the "day job" is done. Therefore, I must also thank my family (my wife, Judy, and children, Max, Julia and Ian) for their patience, and occasional cheers—mainly for attempts at artistic graphics—as this project began to unfold at home.

James Rosenzweig
Los Angeles, 2002

Contents

1 Introduction to beam physics **1**
- 1.1 History and uses of particle accelerators 2
- 1.2 System of units, and the Maxwell equations⋆ 5
- 1.3 Variational methods and phase space⋆ 6
- 1.4 Dynamics with special relativity and electromagnetism⋆ 12
- 1.5 Hierarchy of beam descriptions 19
- 1.6 The design trajectory, paraxial rays, and change of independent variable 22
- 1.7 Summary and suggested reading 25

Exercises 27

2 Charged particle motion in static fields **30**
- 2.1 Charged particle motion in a uniform magnetic field 30
- 2.2 Circular accelerator 32
- 2.3 Focusing in solenoids 35
- 2.4 Motion in a uniform electric field 37
- 2.5 Motion in quadrupole electric and magnetic fields 40
- 2.6 Motion in parallel, uniform electric and magnetic fields 42
- 2.7 Motion in crossed uniform electric and magnetic fields* 44
- 2.8 Motion in a periodic magnetic field* 45
- 2.9 Summary and suggested reading 47

Exercises 48

3 Linear transverse motion **52**
- 3.1 Weak focusing in circular accelerators 52
- 3.2 Matrix analysis of periodic focusing systems 54
- 3.3 Visualization of motion in periodic focusing systems 60
- 3.4 Second-order (ponderomotive) focusing 63
- 3.5 Matrix description of motion in bending systems 67
- 3.6 Evolution of the momentum dispersion function 70
- 3.7 Longitudinal motion and momentum compaction 73
- 3.8 Linear transformations in six-dimensional phase space* 75

3.9 Summary and suggested reading 76
Exercises 77

4 Acceleration and longitudinal motion 81
4.1 Acceleration in periodic electromagnetic structures 81
4.2 Linear acceleration in traveling waves 83
4.3 Violently accelerating systems 86
4.4 Gentle accelerating systems 87
4.5 Adiabatic capture 90
4.6 The moving bucket 93
4.7 Acceleration in circular machines: the synchrotron 95
4.8 First-order transverse effects of accelerating fields* 99
4.9 Second-order transverse focusing in electromagnetic accelerating fields* 102
4.10 Summary and suggested reading 106
Exercises 107

5 Collective descriptions of beam distributions 110
5.1 Equilibrium distributions 111
5.2 Moments of the distribution function 114
5.3 Linear transformations of the distribution function 115
5.4 Linear transformations of the second moments 120
5.5 The rms envelope equation 124
5.6 Differential equation description of Twiss parameter evolution 126
5.7 Oscillations about equilibrium 128
5.8 Collective description of longitudinal beam distributions 130
5.9 The Twiss parameters in periodic transverse focusing systems 137
5.10 Summary and suggested reading 145
Exercises 146

6 Accelerator technology I: magnetostatic devices 150
6.1 Magnets based on current distributions 150
6.2 Magnets based on ferromagnetic material 153
6.3 Saturation and power loss in ferromagnets 156
6.4 Devices using permanent magnet material 159
6.5 Summary and suggested reading 164
Exercises 165

7 Accelerator technology II: waveguides and cavities 167
7.1 Electromagnetic waves in free space⋆ 167
7.2 Electromagnetic waves in conducting media⋆ 168
7.3 Electromagnetic waves in waveguides⋆ 171
7.4 Resonant cavities 177
7.5 Equivalent circuits for waveguides and cavities 183
7.6 Cavity coupling to the waveguide and the beam 191

7.7 Intra-cavity (cell-to-cell) coupling 193
7.8 Traveling wave versus standing wave cavity structures 197
7.9 Electromagnetic computational design tools* 199
7.10 Summary and suggested reading 202
Exercises 203

8 Photon beams 207
8.1 Photons 208
8.2 Ray optics 210
8.3 Optical resonators 213
8.4 Gaussian light beams 215
8.5 Hermite–Gaussian beams 220
8.6 Production of coherent radiation: lasers 223
8.7 Electromagnetic radiation from relativistic charged particles 227
8.8 Damping due to synchrotron radiation* 229
8.9 Undulator radiation and the free-electron laser* 230
8.10 Summary and suggested reading 235
Exercises 237

Appendix A: Selected problem solutions 241

Appendix B: Advanced topics addressed in volume II 289

Index 291

Introduction to beam physics

1

This textbook, the first of a planned two-set volume on the physics of accelerators and beams, is intended to provide a comprehensive introduction to the physical principles underlying the theory and application of particle beam, accelerator, and photon beam physics, within the context of a one semester (or two quarter) undergraduate course. Its emphasis lies in providing the basic conceptual and analytical tools underpinning further study in the field. The course of study is presented from a unified viewpoint, with connections drawn between what may at first seem to be disparate topics made wherever possible.

The book begins, in this chapter, with an overview of the basic concepts needed to start the discussion of particle beams—collections of charged particles all traveling in nearly the same direction with the nearly the same (possibly relativistic) speed—and accelerators. These concepts include Lagrangian and Hamiltonian approaches to mechanics, and how these methods are applied naturally in the context of relativistic charged particle motion. After this formal re-introduction to some powerful analysis tools, we then proceed to examine the motion of charged particles in static electric and magnetic fields, with the purpose of acquiring a basic understanding of the relevant categories of motion. From these building blocks, we then take up a series of topics in particle beam physics: linear transverse oscillations, acceleration and longitudinal motion in linear and circular devices, and envelope descriptions of beams. In this initial volume, these topics are discussed from the viewpoint of collections of nearly non-interacting particles, where the forces generated by the particles' collective electromagnetic fields are too small to be of interest. On the other hand, much of modern accelerator physics is concerned with intense beams that have very strong self-forces, and display characteristics of plasmas (ionized gases); the physics of such systems is beyond the scope of this text, but will be addressed in the following volume. A description of the topics covered in this second volume is given in Appendix B.

After the introductory survey of particle and beam dynamics in the first five chapters of this volume, we subsequently examine some aspects of relevant technologies. In particular, we concentrate on the features of physics and engineering methods used in accelerator magnet and electromagnetic accelerating systems most directly related to the material presented on charged particle motion. Our investigation is then extended to include the comparisons between single particle and collective descriptions of charged particle beam optics on the one hand, and ray and wave optics in coherent electromagnetic (light) beams, such as lasers, on the other. With the introduction of electromagnetic radiation in the text, the discussion progresses to encompass aspects of charged particle radiation processes and their effect on charged particle motion.

1.1 History and uses of particle accelerators **2**
1.2 System of units, and the Maxwell equations⋆ **5**
1.3 Variational methods and phase space⋆ **6**
1.4 Dynamics with special relativity and electromagnetism⋆ **12**
1.5 Hierarchy of beam descriptions **19**
1.6 The design trajectory, paraxial rays, and change of independent variable **22**
1.7 Summary and suggested reading **25**
Exercises **27**

1.1 History and uses of particle accelerators

The history of particle accelerators is one of physics and technology at the cutting edge, as the desire to use increasingly higher-energy particles for basic research in physics has led to vigorous innovation and experimentation. These efforts have caused explosive growth in the field of particle accelerators, which at the dawn of the new millennium has established itself as a fundamental area of research in its own right, with its own research journals and societies. It is, however, a cross-disciplinary field, having many connections to other sub-disciplines within the world of physics. Areas of inquiry in which particle beam physicists have made a significant impact, or have borrowed techniques from, over the years include high-energy physics, nuclear physics, nonlinear dynamics, medical physics, plasma physics, and coherent radiation and X-ray sources. Particle beams are now an indispensable tool in these endeavors, with thousands of practitioners worldwide using them, and billions of dollars per year being spent by governments and industry to develop and improve them.

It was not always such! Let us review the history of particle accelerators in order to provide a context for our discussion and to introduce some concepts and terminology. This is not meant to be a self-contained discussion—some concepts are mentioned and can only be defined later—we only wish to sketch a description of the successive generations of accelerators so that, when we discuss them further in the text, the reader has an idea of the conceptual and chronological roles of these devices.

The history of the particle accelerator begins in the mid-nineteenth century with the development of the cathode ray tube, in which electrons are accelerated across a vacuum gap with an applied electrostatic potential. The need to create cathode ray tubes spurred the advance of many aspects of modern experimental physics, such as voltage sources and vacuum techniques. The cathode rays themselves were the subject of intense scrutiny, which led to the actual discovery of the electron and its properties. Notable experiments in this regard include the determination of the electron's charge-to-mass ratio by J.J. Thomson, and the discovery of the photoelectric effect by Lenard and Millikan.

The cathode ray tube evolved[1] over time and technological development to the electrostatic accelerator, which, instead of kV potentials, gave rise to MV potentials and the creation of electron beams with relativistic velocity. The technological innovations associated with electrostatic accelerators included several inventions such as the belt-charged Van der Graaf accelerator and the cascaded-voltage Cockcroft–Walton generator. These devices could be used to accelerate both electrons and heavier, ionized particles—allowing the birth of nuclear physics, and playing a key role in the quantum revolution of the 1920s and 1930s. During this time, radio-frequency linear accelerators were also studied, but did not become prominent tools in physics research until later.

The advancement of particle accelerators definitively hit its stride with the invention of the initial circular accelerators: the ion accelerator known as the cyclotron, and an electron accelerator termed the betatron. The betatron was proposed as early as 1924 by Wideroe and was made into experimental reality by Kerst and Serber in 1940. This device introduced acceleration based on electromagnetic induction and provided a demonstration of the principle of weak focusing (giving rise to simple transverse betatron oscillations, discussed in Sections 2.2 and 3.1). This transverse focusing effect was also developed in

[1] The cathode ray tube evolved most importantly from the general viewpoint into the television display and computer monitor. Much of this text was being written on a lap-top computer with a liquid crystal display, however, so this ancient accelerator technology may, after a century of use, be losing its dominance in this application.

the cyclotron, a machine with an expanding circular geometry (see Fig. 2.14 associated with Ex. 2.3). The cyclotron was also notably the first device to show acceleration based on resonance of particle motion with time-varying electromagnetic fields. Additionally, the cyclotron lends its name to the frequency of oscillation upon which this resonance is based, the well-known cyclotron frequency. This frequency is not constant (see Section 2.1) but turns out to vary noticeably when the particle becomes relativistic. For heavy particles, going above a few 100 MeV of kinetic energy required the invention of the synchrocyclotron, in which the frequency of the applied fields is varied in time. The cyclotron concept is still employed in many nuclear and medical accelerators, but for higher energies, such devices could not be used. This is due to iron-based magnets that must be used throughout the machine to bend the particles in circular or spiral orbits. At a certain point one cannot keep building larger magnets, due to the complication and expense involved.

After the Second World War (in which the cyclotron played a part in development of the atomic bomb[2]), radar technology pushed the invention of the radio-frequency linear accelerator, by allowing microwave powers high enough to directly accelerate particles in electromagnetic cavities. The Alvarez drift-tube linear accelerator (linac), a standing wave structure (see Figs 1.1 and 4.13), was the first of this category of accelerator, followed by the periodically loaded traveling wave (e.g. Fig. 4.1) structures typical of modern linacs. The traveling wave linac has allowed the construction of a 50-GeV electron accelerator at Stanford in which the quark structure of matter was first observed. Higher-energy linacs are now on the horizon, and may be the next frontier tool for discoveries in particle physics.

[2]The principles of cyclotron motion were employed in radioactive isotope separation.

Fig. 1.1 The interior of an Alvarez drift-tube linear accelerator cavity, showing the drift tubes and their supports.

The circular accelerator also underwent a "revolution" in the post-war period due to the invention of the synchrotron. The synchrotron, in which the concept of phase focusing, or phase stability (characterized by longitudinal—in the direction of nominal beam motion—synchrotron oscillations), was fully developed, is a merging of the linac, in that it employs radio-frequency acceleration, with the circular accelerator and its associated bend magnets. In the synchrotron, unlike the cyclotron, particles always stay on approximately the same radius orbit. This is also true of the betatron, but, since (see Ex. 2.2) the acceleration in the betatron arises from electromagnetic induction, the entire interior area bounded by the particle orbit must have a time-varying magnetic field. The synchrotron, however, is free of the constraints on magnetic field of both the betatron and the cyclotron, so the bend magnets need only be placed near that orbit, not the entire device. This innovation, along with the implementation of alternating gradient focusing (also termed strong focusing, as opposed to the weak focusing of the betatron), has allowed very large energy synchrotrons to be built. One such device is the 0.9-TeV (1 TeV $= 10^{12}$ eV) Tevatron at Fermi National Accelerator laboratory outside of Chicago with a radius of 1 km. One of course could not imagine the cost associated with using iron in the entire interior of this device! In fact the modern electromagnets employed in the Tevatron, which are shown in Fig. 1.2, are superconducting and as such do not rely on iron to achieve high fields.

Fig. 1.2 Bend magnets associated with the Tevatron collider, presently the world's largest highest-energy synchroton.

The Tevatron is an example of a synchrotron that is operated as a collider, in which counter-propagating beams of equal energy particles and antiparticles are squeezed into sub-mm-sized collision regions located inside of the huge, sophisticated particle detectors used to analyze the debris produced in hard

Fig. 1.3 Aerial view of the Tevatron collider complex at Fermilab.

collisions. At the Tevatron, the top quark was recently discovered in such a proton–antiproton ($p\bar{p}$) collider experiment; at the European laboratory CERN, which built the first $p\bar{p}$ collider, the W and Z intermediate vector bosons were discovered some 15 years previously in a similar manner. An aerial view of the entire Tevatron complex at Fermilab is shown in Fig. 1.3. The Tevatron injection system includes a charged particle source, linear accelerator, and two smaller-energy synchrotrons, as well as the collider ring itself. It also has beamlines through which high-energy protons extracted from the rings can be directed onto fixed targets, allowing experiments based on creation of secondary beams that consist of more exotic particles, such as muons and neutrinos.

The synchrotron also lends its name to the radiation produced by charged particles as they bend in magnetic fields—synchrotron radiation. This radiation is both a curse and a blessing. As an energy loss mechanism that has a strong dependence on the ratio of the particle energy to its rest energy (see Section 8.7), it practically limits the energy of electron synchrotrons to that currently achieved, around 100 GeV. On the other hand, synchrotron radiation derived from multi-GeV electron synchrotrons is the preferred source of hard x-rays for research purposes today, with over a dozen such major facilities (synchrotron light sources) world-wide. Synchrotron radiation also forms the physical basis of the free-electron laser; it can produce coherent radiation in both long- and short-wavelength regimes that are inaccessible to present laser sources based on quantum systems. Both the need to create collisions in high-energy colliders, and the desire to make a high-intensity free-electron laser imply that the beams involved must be not only energetic, but of very high quality. A measure of this quality is the phase space density of the beam, which is introduced in Section 1.5.

Today, particle accelerators, while a mature field, present considerable challenges to the physicist who must use and improve these tools. These challenges arise from the need in elementary particle experiments to move to ever increasing energies, a trend that is placed in doubt by the cost of future machines. As the present high-energy frontier machines cost well in excess of $\$10^9$, accelerator physicists are in the process of exploring much more compact and powerful accelerators based on new physical principles. These new acceleration techniques may include use of lasers, plasmas, or ultra-high-intensity charged particle beams themselves. Accelerators also promise to play a critical future role in short-wavelength radiation production, inertial fusion, advanced fission schemes, medical diagnosis, surgery and therapy, food sterilization, and transmutation of nuclear waste. These goals present new challenges worthy of the short, yet accomplished, history of the field.

Even with the present level of sophistication, the subject of accelerators can be initially approached in a straightforward way. The fundamental aspects of particle motion in accelerators can be appreciated from examination of simple configurations magnetostatic (or, less commonly, electrostatic) fields, which may be used to focus and guide the particles, and confined electromagnetic fields that allow acceleration. Moreover, analysis of charged particle dynamics in these physical systems has certain general characteristics, which are discussed in the remainder of the chapter. We begin this discussion by writing the basic equations governing the electromagnetic field, and then proceed to review aspects of methods in mechanics—Lagrangians and Hamiltonians, as well as special relativity. Based upon this discussion, we then introduce the

description of beams as distributions in phase space. We finish the present chapter by examining the notion of the design trajectory and analysis of nearby "paraxial" trajectories.

We note that the following three sections are indicated (by the ⋆ symbol) as review, and therefore optional for a first reading of the text. In fact, most readers should benefit from the material presented, either as a review for those who are familiar with the methods discussed, or as a focused introduction to the uninitiated. In any case, the results contained in Sections 1.2–1.4 will be referenced often in the remainder of the text, and will thus need to be seriously examined sooner or later.

1.2 System of units, and the Maxwell equations⋆

In order to construct our analyses, we must begin by choosing a system of units. While classical electromagnetism in general, and particle beam physics in particular, are a bit more compactly written in cgs units, we use mks or SI units in this text. This is for two main reasons: (a) ease of translation of the results into laboratory situations, and (b) familiarity of undergraduate physics students, as well as engineers, with the mks system.

Several basic equations need to be introduced in units-specific context. These include the Maxwell equations:

$$\vec{\nabla} \cdot \vec{B} = 0, \tag{1.1}$$

$$\vec{\nabla} \cdot \vec{D} = \rho_{\mathrm{e}}, \tag{1.2}$$

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{J}_{\mathrm{e}}, \tag{1.3}$$

and

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{1.4}$$

where ρ_{e} and $\vec{J}_{\mathrm{e}}$ are the free electric charge density and current density, respectively, that are related by the equation of continuity

$$\vec{\nabla} \cdot \vec{J}_{\mathrm{e}} + \frac{\partial \rho_{\mathrm{e}}}{\partial t} = 0. \tag{1.5}$$

We will also make use of the following relations between the electromagnetic fields, the scalar potential ϕ_{e} and the vector potential $\vec{A}$,

$$\vec{E} = -\vec{\nabla}\phi_{\mathrm{e}} - \frac{\partial \vec{A}}{\partial t}, \tag{1.6}$$

$$\vec{B} = \vec{\nabla} \times \vec{A}. \tag{1.7}$$

For completeness, we must also include the constitutive equations,

$$\vec{D} = \varepsilon(\vec{D})\vec{E} \quad \text{and} \quad \vec{B} = \mu(\vec{H})\vec{H}, \tag{1.8}$$

where $\varepsilon(\vec{D})$ and $\mu(\vec{H})$ are the electric permittivity and the magnetic permeability of a material, respectively. We will not encounter the first of Eq. (1.8) again in

the text (except in the context of vacuum electrodynamics, where $\varepsilon(\vec{D}) = \varepsilon_0$, the permittivity of free space) until we discuss propagation of light in Chapter 8. The latter of Eq. (1.8) will be revisited when we discuss the design principles of electromagnets based on ferric materials. We note here that the constants $\varepsilon_0 = 8.85 \times 10^{-12}\mathrm{C}^2/\mathrm{N\,m}^2$ and $\mu_0 = 4\pi \times 10^{-7}\mathrm{N/A}^2$ are related to the speed of light c by

$$c = (\varepsilon_0 \mu_0)^{-1/2} = 2.998 \times 10^8 \mathrm{m/s}. \tag{1.9}$$

In particle beam physics, one often wishes to use MeV (10^6 eV=1.6×10^{-13} J) as the unit of energy. In this case it is useful, when making calculations of applied acceleration due to an electric field, to quote the electric force (acceleration gradient) qE in terms of eV/m by simply absorbing the charge q, an integer multiple of e, into the units. This same position may be adopted in the context of applied magnetic forces if one notes that the force qvB also has units of eV/m when one absorbs the charge q and multiplies the magnetic field B in tesla (T) by the velocity v in m/s. Note that this implies that the commonly encountered level of 1 T static magnetic field is equivalent to a 299.8 MV/m static electric field in force for a relativistic ($v \approx c$) charged particle. This electric field exceeds typical breakdown limits on metallic surfaces by nearly two orders of magnitude, giving partial explanation to the predominance of magnetostatic devices over electrostatic devices for manipulation of charged particle beams.

When considering the self-forces of a collection of charged particles, the combination of constants $e^2/4\pi\varepsilon_0$ often arises. This quantity may be converted to our desired units by writing

$$\frac{e^2}{4\pi\varepsilon_0} = r_\mathrm{c} m_0 c^2, \tag{1.10}$$

where r_c is defined as the classical radius of the (assumed $|q| = e$) and m_0 is the rest mass of the particle. In the case of the electron, we have a rest energy m_0c^2 in useful "high-energy physics" units of $m_\mathrm{e}c^2 = 0.511$ MeV and a classical radius of $r_\mathrm{e} = 2.82 \times 10^{-15}$ m. Thus, we may write $e^2/4\pi\varepsilon_0 = 1.44 \times 10^{-15}$ MeV m.

1.3 Variational methods and phase space⋆

The study of beam physics is based on the understanding of relativistic motion of charged particles under the influence of electromagnetic fields. Such fields are constrained by the relations shown in Eqs (1.1)–(1.7). Given the $\vec{E}$ and $\vec{B}$ fields, the analysis of charged particle dynamics can be performed, perhaps most naturally, using only differential equations derived from the Lorentz force equation,

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q(\vec{E} + \vec{v} \times \vec{B}). \tag{1.11}$$

While we will base many of our discussions of charged particle motion in this book on the Lorentz force equation, more powerful methods are also available that use variational principles, that is, Lagrangian and Hamiltonian analyses. These methods, which have traditionally been introduced at the graduate level, are now increasingly taught in undergraduate-level mechanics courses. The power of variational methods is found in their rigor, and in the clarity of the results obtained when such approaches are applied to problems naturally

formulated in difficult coordinate systems, such as curvilinear or accelerating systems. Even in difficult cases, variational methods give a straightforward formalism that reliably yields the correct equations of motion.

We now give a short review of these methods, which will also prepare us, in a quite natural way, to discuss the roles of electromagnetic fields and special relativity in classical mechanics. This review is meant to clarify these subjects to the reader who is already conversant in variational methods and relativity. For one who has not studied these subjects before, the following discussion (the remainder of this chapter) may serve as an introduction, albeit a steep one, which may be supplemented by material recommended in the bibliography. It may be remarked that beam physics provides some of the most elegant and illustrative uses of advanced methods in dynamics, as well as the role of relativity in these dynamics, that are encountered in modern physics. Thus, even if this text serves as a first introduction to these subjects, it will be a physically relevant and, hopefully, rewarding discussion.

The discussion of variational methods nearly always is initiated by introduction of the Lagrangian, which in non-relativistic mechanics is given by

$$L(\vec{x}, \dot{\vec{x}}) = T - V. \tag{1.12}$$

Most commonly, the potential energy V is a function of the position (coordinates) $\vec{x}$ and the kinetic energy T is a function of the velocity $\dot{\vec{x}}$. Note that we use the notation $\vec{x}$ to indicate the set of M generalized coordinates[3] x_i ($i = 1, \ldots, M$), and the associated velocities are, thus, defined as $\dot{\vec{x}} \equiv \mathrm{d}\vec{x}/\mathrm{d}t$ (the compact notation $\dot{(\cdot)} \equiv \mathrm{d}/\mathrm{d}t$ will be used in this text to indicate a total time derivative). The application of Lagrangian formalism, and the Hamiltonian formalism that is based upon it, to forces not derivable from a scalar potential V (such as magnetic forces) is discussed in the next section.

[3] A generalized coordinate is often a simple Cartesian distance (x, y, or z), but may also be an angle, as naturally found in cylindrical or spherical polar coordinate systems.

The equations of motion are derived from the Lagrangian by Hamilton's principle, or the principle of extreme action,

$$\delta \int_{t_1}^{t_2} L \, \mathrm{d}t = 0. \tag{1.13}$$

The variation of coordinate and velocity components in the integral in Eq. (1.13), when at an extremum, yields a recipe that gives the Lagrange–Euler equations of motion,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0. \tag{1.14}$$

The power of these equations is first and foremost in that they rigorously generate forces of constraint and "fictitious forces" such as those arising from centripetal acceleration. This is a significant accomplishment, but one that is eventually overshadowed by the use of the Lagrangian to form the basis of constructing a Hamiltonian function,

$$H(\vec{x}, \vec{p}) \equiv \vec{p} \cdot \dot{\vec{x}} - L, \tag{1.15}$$

where the canonical momenta are defined through the Lagrangian by

$$p_i \equiv \frac{\partial L}{\partial \dot{x}_i}. \tag{1.16}$$

These momenta (momentum components) are new dependent variables in the formalism, replacing the role of the velocity components in the Lagrangian analysis. In the most familiar example, that of non-relativistic motion in Cartesian coordinates, the kinetic energy is $T = \frac{1}{2}m_0\dot{\vec{x}}^2 = \vec{p}^2/2m_0$, and the momenta are $p_i = m_0\dot{x}_i$, as expected.

In the Hamiltonian formalism, Hamilton's principle gives twice the number of equations of motion,

$$\dot{x}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial x_i}, \tag{1.17}$$

as the Lagrange–Euler equations. The first of Eq. (1.17) defines the velocity components in terms of the canonical momentum components; the second, governing the time evolution of the momentum, is a generalization of Eq. (1.11). The canonical momentum components and corresponding coordinates have a nearly symmetrical relationship with each other, and pairs of such variables are termed canonically conjugate. The space $(\vec{x}, \vec{p})$ of all such pairs is termed phase space. It should be noted that the canonical momentum is not necessarily identical to the more familiar mechanical momentum employed in Eq. (1.11). This point is returned to in Section 1.4.

The Hamiltonian formalism allows constants of the motion to be derived with little difficulty. From Eq. (1.17), it is apparent that if the Hamiltonian is independent of the coordinate, then the conjugate momentum component is a constant of the motion. Likewise, if the Hamiltonian is independent of the momentum component, then the conjugate coordinate is a constant of the motion. Further, the Hamiltonian obeys the relation

$$\dot{H} = \frac{\partial H}{\partial t}, \tag{1.18}$$

and therefore H is a constant of the motion if it is not explicitly dependent on the time t.

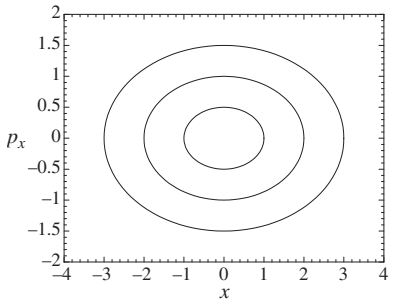


Fig. 1.4 Phase plane plot for simple harmonic oscillator orbits, corresponding to three different values of the Hamiltonian $H = 1/8, 1/2$, and 2, with $m = 1$ and, $\omega = 0.5$. The trajectories of the oscillator lie along ellipses described by these constant H curves.

If the Hamiltonian is a constant of the motion, we can most often identify this constant as the total energy U of the system. The invariance of H allows one of the main tools of particle beam physics to be employed—the drawing of the so-called phase space maps. The phase space is the $2M$-dimensional space of all M pairs of coordinates and their canonically conjugate momenta. Phase space maps, an example of which is displayed in Fig. 1.4, are of a (x_i, p_i) trajectory drawn in two-dimensional projections of the full phase space. As such, these representations may more properly be called phase plane maps. Of course, in fully three-dimensional accelerators, we have $M = 3$, and there are three phase planes in which maps are drawn. In particle beam physics, in fact, one often deals with motion in which the variables in one phase plane are very nearly independent of any other phase plane variable. This state of affairs, in addition to the inherent ease of two-dimensional (as opposed to higher dimensional) visualization, makes phase plane descriptions popular as a tool for understanding particle beam dynamics.

The creation of phase plane maps such as Fig. 1.4 is accomplished analytically by using a time-independent Hamiltonian and plotting $H(x_i, p_i) =$ constant curves. In more complicated cases, phase plane maps are created by numerical solution of the equations of motion. The concept of phase space is central to

the field of particle beam physics and certain results, such as the invariance of phase space density (see discussion Section 1.4) can only be clearly discussed in the context of Hamiltonian formalism.

Perhaps the most familiar example of this mapping technique is the one-dimensional non-relativistic simple harmonic oscillator. In this case (for simplicity indicating the coordinate $x_1 \Rightarrow x$) the one-dimensional Hamiltonian is of the form

$$H = \frac{1}{2m}[p_x^2 + m^2\omega^2 x^2], \tag{1.19}$$

where $\omega^2 = K/m$, and K is the oscillator strength or spatial gradient of the restoring force, $K = -F_x/x$. The phase plane maps associated with simple harmonic motion are thus ellipses in the two-dimensional (x, p_x) phase plane, as shown by the examples in Fig. 1.4. Note that the Hamiltonian alone does not indicate the direction in which the system traces out the ellipse, but examination of the force and velocity direction does—the direction of motion is clearly clockwise in phase space for this system.

The area of the phase plane ellipse is proportional to the value of the Hamiltonian associated with each trajectory, and is therefore also a constant of the motion. This area is given by

$$\oint p_x \,\mathrm{d}x = \oint p_x \dot{x} \,\mathrm{d}t = \oint (H + L) \,\mathrm{d}t = U\tau. \tag{1.20}$$

where τ is the period of the oscillation. We shall see in Chapter 5 that the area associated with a closed trajectory in phase space forms a central place in the theory particle beam dynamics.

The phase plane map is of great use in visualizing the motion of charged particles beyond simple harmonic orbits (see Ex. 1.3) and is profitably employed even in cases when the Hamiltonian is not a constant of the motion. In the case of a time-varying Hamiltonian, one may not trivially generate plots like Fig. 1.4, but must often solve the equations of motion (Eq.(1.17)) first. Furthermore, if one solves these equations in such a case, it may not be illuminating, but rather confusing (e.g. Fig. 3.6), to use continuous lines in phase space to illustrate the motion as it advances continuously in time. For systems typical of circular accelerators, the Hamiltonian varies periodically in time t, however. A valuable strategy for phase plane plotting in this case is taking periodic "snap-shots" and plotting the instantaneous position in the phase plane once per Hamiltonian (not oscillation) period. This type of map is termed a Poincare plot (e.g. Fig. 3.7) and is discussed further in Chapter 3.

There are also manipulations of the phase space or phase plane variables that can be undertaken to create a description where the Hamiltonian is a constant of the motion in the new variables, where it was not constant in the old variables. To see the utility of this approach, consider an explicitly time-dependent Hamiltonian, in which the potential arises from a traveling wave so that the Hamiltonian can be written

$$H = \frac{1}{2m}[p_x^2 + G(x - v_\varphi t)], \tag{1.21}$$

where v_φ is the phase velocity of the wave in the x direction and G is an arbitrary function. The simplest way to make the Hamiltonian into a constant of the motion is to perform a mathematical transformation of the system

description. Let us examine such a transformation of the coordinates, the Galilean transformation, where

$$\tilde{x} = x - v_\varphi t. \tag{1.22}$$

Now we must transform the Hamiltonian so that the coordinate's equation of motion remains correct. With the new canonical momentum set equal to the old, $\tilde{p}_x = p_x$,

$$\dot{\tilde{x}} = -\frac{\partial \tilde{H}}{\partial \tilde{p}_x} = \frac{\tilde{p}_x}{m} - v_\varphi = v - v_\varphi, \tag{1.23}$$

as expected for a Galilean transformation. To generate Eq. (1.23) as a correct canonical equation of motion (i.e. one derivable from Eq. (1.17)), the new Hamiltonian $\tilde{H}$ must transform from the old Hamiltonian H as

$$\tilde{H}(\tilde{x}, \tilde{p}_x) = H(\tilde{x}, \tilde{p}_x) - v_\phi \tilde{p}_x = \frac{1}{2m}[\tilde{p}_x^2 + G(\tilde{x})] - v_\phi \tilde{p}_x. \tag{1.24}$$

Now the new Hamiltonian $\tilde{H}(\tilde{x}, \tilde{p}_x)$ is explicitly independent of t and is thus a constant of the motion. The trajectory of the charged particle in this wave potential can, therefore, be visualized, as before, with a phase plane map created by the simple algebraic relationship between $\tilde{x}, \tilde{p}_x$, and $\tilde{H}$. For example, we may use Eq. (1.24) with a moving simple harmonic oscillator potential, $G(\tilde{x}) = \frac{1}{2}K\tilde{x}^2$. This leads to

$$\begin{aligned}\tilde{H}(\tilde{x}, \tilde{p}_x) &= \frac{1}{2m}\left[\tilde{p}_x^2 + m^2\omega^2\tilde{x}^2 - 2p_\varphi \tilde{p}_x\right] \\ &= \frac{1}{2m}\left[(\tilde{p}_x - p_\varphi)^2 + m^2\omega^2\tilde{x}^2\right] - T_\varphi,\end{aligned} \tag{1.25}$$

where we have defined p_φ and T_φ as the (non-relativistic) momentum and kinetic energy associated with a particle of mass m traveling at the phase velocity v_φ. In this case, the constant $\tilde{H}$ curves associated with the motion shift upward in $\tilde{p}_x$ by p_φ, when compared those shown in Fig. 1.4, to as shown in Fig. 1.5.

Note that the phase plane plots for the moving simple harmonic oscillator potential can be made to look identical to the stationary potential plot by use of

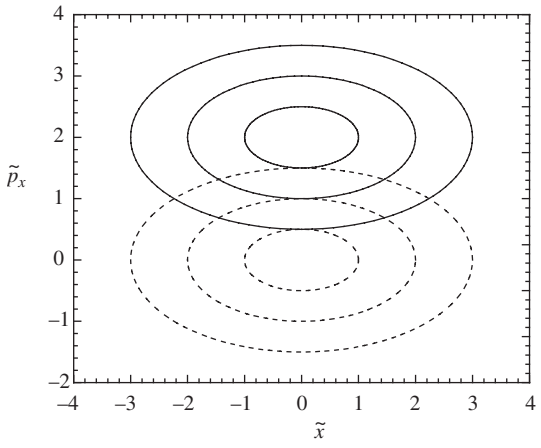


Fig. 1.5 Phase plane plot for moving simple harmonic oscillator orbits corresponding to same limits in $\tilde{x}$ as those in x found in Fig. 1.4. Here $p_\varphi = 2$, with $m = 1$ and $\omega = 0.5$. The curves corresponding to the moving potential are represented by solid lines, and their counterparts from Fig. 1.4 are shown in dashed lines.

our Galilean transformation from x to $\tilde{x}$ and, further, by plotting $\delta p \equiv \tilde{p}_x - p_\varphi$. Each of the curves in the $\tilde{x}$ frame is associated with values (total energies) of the new Hamiltonian $\tilde{H} = H - T_\varphi$. The constant $-T_\varphi$ contains no information about the system's dynamics (cf. Eq. (1.17)), however, and may be ignored. At this point, we need to clarify that the transformation given by Eqs (1.22)–(1.24) is a purely mathematical change of variables, not a change of physical frame. We will make use of this type of mathematical transformation while discussing acceleration in traveling electromagnetic waves (Chapter 4). The physical change of frame is described, of course, not by a Galilean transformation but by a Lorentz transformation, as discussed in Section 1.4.

The type of variable transformation illustrated by Eqs (1.22)–(1.24) is termed a canonical transformation because it preserves the canonically conjugate relationship between the coordinate and the momentum. In general, one does not use such an ad hoc way of deriving the transformation but a more rigorous method based on generating functions, which are discussed in advanced mechanics textbooks. These functions come in a variety of types, depending on the variables to be transformed. The generating function always is dependent on a pair of variables per phase plane, some combination of the old and new canonically conjugate variables.

As an example, consider use of a generating function to transform the one-dimensional simple harmonic oscillator Hamiltonian, Eq. (1.19) to a particularly interesting new form. The well-known solutions for the motion in such a system are given by

$$x(t) = x_m \cos(\omega t + \theta_0), \quad \text{and} \quad p_x(t) = m\omega x_m \sin(\omega t + \theta_0). \tag{1.26}$$

We would like to transform the Hamiltonian to one that reflects the constant rate of advance in argument of the cosine function in Eq. (1.26), so we propose that the new coordinate be chosen as $\theta = \omega t + \theta_0$. In this case, we can use a generating function of the form $F(x, \theta)$ to transform the simple harmonic oscillator problem into a more useful form. According to Hamilton's principle, Eq. (1.27) must yield the following formal properties:

$$p_x = \frac{\partial}{\partial x} F(x, \theta), \quad J = \frac{\partial}{\partial \theta} F(x, \theta), \quad H' = H + \frac{\partial}{\partial t} F(x, \theta), \tag{1.27}$$

where J is the new momentum and H' is the new Hamiltonian. We can deduce from Eq. (1.27) that a proper generating function is given by $F(x, \theta) = \frac{1}{2} m\omega x^2 \cot(\theta)$. Using $F(x, \theta)$ to obtain the new momentum and Hamiltonian, we have

$$H' = J\omega, \tag{1.28}$$

which is a constant of the motion. Since ω is a constant, the momentum J, known as the action, is also a constant. The new canonically conjugate pair are termed action-angle variables. The action-angle description is important for analyzing perturbations to simple harmonic systems, a commonly encountered problem in particle beam physics. The action is, comparing Eqs (1.18) and (1.25), simply related to the area enclosed by the phase space trajectory,

$$J = \frac{1}{2\pi} \oint p_x \, \mathrm{d}x. \tag{1.29}$$

The action is also generally known to be an adiabatic invariant, in that when the parameters of an oscillatory system are changed slowly, the action remains a

constant. This can be illustrated by writing the differential equation for a slowly varying oscillator (with the mass factor set to $m = 1$) as

$$\ddot{x} + K(t)x = 0, \tag{1.30}$$

and substituting an assumed form of the solution

$$x(t) = C\sqrt{a(t)}\cos(\psi(t)). \tag{1.31}$$

After some manipulations, we obtain the relations

$$\dot{\psi} = \frac{1}{a} \quad \text{and} \quad \ddot{a} - \frac{\dot{a}^2 + 4}{2a} + 2Ka = 0. \tag{1.32}$$

The substitution given in Eq. (1.31) is commonly found in the theory of time-dependent oscillators, which is quite important in particle beam physics—it is explicitly used in an analysis in Chapter 5. The solution of the first equation in (1.32) is formally

$$\psi = \int \frac{\mathrm{d}t}{a(t)}, \tag{1.33}$$

while the second of these equations' solution can be examined approximately. Assuming the first and second time derivatives of a are small ($\ddot{a}/a \ll K$ and $(\dot{a}/a)^2 \ll K$)[4], one has simply

$$a(t) \cong (K(t))^{-1/2} \quad \text{or} \quad x(t) = C(K(t))^{-1/4}\cos(\psi(t)). \tag{1.34}$$

[4]These conditions give a quantitave definition of the term "adiabatic".

The momentum corresponding to this approximate solution is

$$p_x(t) = \dot{x} \cong C(K(t))^{1/4}\sin(\psi(t)). \tag{1.35}$$

The area of the phase space ellipse, whose semimajor and semiminor axes are the maximum excursion in x and p_x, respectively, is again the value of the action at that time,

$$J = \tfrac{1}{2}p_{x,\max}x_{\max} = \tfrac{1}{2}C^2. \tag{1.36}$$

The action is independent of time and dependent only on the initial conditions (taken at $t = 0$) through the constant C^2. Thus, we have shown the adiabatic invariance of the action J for oscillators whose strength is slowly varying, a result pertinent to discussions in future chapters.

We emphasize at this juncture the primary role of the momenta in Hamiltonian methods, as opposed to the velocity components found in the Lagrangian formalism. This is inherent in both the structure of Hamiltonian and relativistic analyses, as is reviewed in Section 1.4.

1.4 Dynamics with special relativity and electromagnetism⋆

The Hamiltonian formulation of dynamics is naturally suited to analyzing relativistic systems. We shall see that this is because in the canonical approach to dynamics the roles of the coordinates, momenta, time, and energy have a rigorously defined relationship with each other. The relationships between

these dynamical variables actually become clearer after one studies relativistic dynamics, where the ways in which all such variables transform from one inertial frame to another are emphasized.

The point of departure for the present discussion is precisely this transformation, which should be familiar to any reader of this text, the Lorentz transformation. This relation, which governs the transformation of coordinates and time from one inertial reference frame to another moving at constant speed $v_f = \beta_f c$ with respect to the first frame (along what we choose to be the z-axis), is written as

$$x' = x, \quad y' = y, \quad z' = \gamma_\mathrm{f}(z - \beta_\mathrm{f} ct), \quad ct' = \gamma_\mathrm{f}(ct - \beta_\mathrm{f} z), \tag{1.37}$$

where the Lorentz factor $\gamma_\mathrm{f} = (1 - \beta_\mathrm{f}^2)^{-1/2}$. This Lorentz transformation acts upon the space-time four-vector $\vec{X} \equiv (x, y, z, ct)$, and preserves the length, or norm, of the four-vector. This norm, termed a Lorentz invariant because it is frame independent, is specified by the quantity

$$|\vec{X}|^2 = x^2 + y^2 + z^2 - (ct)^2. \tag{1.38}$$

The invariance of the norm of the space-time four-vector is often the starting point of the derivation of the Lorentz transformation, as it indicates that the phase velocity c of spherical light waves in vacuum is independent of inertial reference frame.

The invariance property of the norm of $\vec{X}$ is therefore entirely equivalent to the property that the four-vector transforms between frames under the rules of a Lorentz transformation. Thus, a four-vector can be defined equivalently either as an object that obeys Lorentz transformations or one in which its norm, as defined by Eq. (1.38), is conserved during such a transformation. The invariance of four-vector norm is a key tool in performing analyses of relativistic dynamics.

The absolute value of $|\vec{X}|^2$, which refers to the "distance" in space-time between two events (or implicitly, one event and the origin), can be positive, negative, or zero. If it is positive, it is termed "space-like", as one may always transform to a frame where the events occur at the same time, but at a separated distance. If it is negative, it is termed "time-like", as one may always transform to a frame in which the events occur at the same point in space, but at separate times. Space-like pairs of events cannot be causally connected, because they are too far separated in space-time for light to propagate between them. If the norm of $\vec{X}$ is zero, the two events are exactly connected by a signal traveling at the speed of light. In this case, the events are said to be on each other's light cone.

Using Lorentz transformations of space and time, it can be trivially shown that properties of waves—the wave numbers (spatial frequencies) k_i and (temporal) frequency ω—form a four-vector. This is intuitively so, since the wave number simply measures spatial intervals while the frequency measures intervals in time. As an illustration of this derivation, consider a plane electromagnetic wave moving in the positive z-direction, with functional form $\cos[k_z z - \omega t]$. If one begins in a frame moving with velocity $\beta_\mathrm{f} c$ in the z-direction, the inverse transformation

$$\begin{aligned} &\cos\left[k_z \gamma_\mathrm{f}(z' + \beta_\mathrm{f} ct') - \omega\gamma_\mathrm{f}\left(t' + \beta_\mathrm{f}\frac{z'}{c}\right)\right] \\ &\quad = \cos\left[\gamma_\mathrm{f}\left(k_z - \beta_\mathrm{f}\frac{\omega}{c}\right)z' - \gamma_\mathrm{f}(\omega - \beta_\mathrm{f} k_z c)t'\right] \end{aligned}$$

can be deduced from Eq. (1.37). We can thus see that k_z and ω (Lorentz) transform as z and t, respectively. Further, we know, therefore, that the norm of the four-vector $(\vec{k}c, \omega) = (k_x c, k_x c, k_x c, \omega)$ is

$$\sum_i k_i^2 c^2 - \omega^2 = \text{const.} \tag{1.39}$$

In the case where the constant on the right-hand side of Eq. (1.39) is zero, this can be recognized as the dispersion relation for vacuum electromagnetic waves. From a quantum-mechanical viewpoint, such waves correspond to massless photons (cf. Section 7.1). When the constant in Eq. (1.39) is not zero, a quantum-mechanical interpretation indicates that we are examining an object of non-vanishing rest mass, or rest energy.

The quantum mechanical indentifications of free particle momenta and energy in terms of wave properties, $p_i = \hbar k_i$ and $U = \hbar\omega$, lead us to the conclusion that momentum and energy must also form a four-vector, $\vec{P} \equiv (\vec{p}c, U) = \hbar(\vec{k}c, \omega)$, with the invariant

$$\vec{P}^2 \equiv \vec{p}^2 c^2 - U^2 = \sum_i p_i^2 c^2 - U^2 = \text{const.} \tag{1.40}$$

We must allow for the possibility of Lorentz transformation into the rest frame of the particle, in which case $\vec{p} = 0$ and the invariant can be identified as the rest energy of the particle. Thus,

$$\vec{p}^2 c^2 - U^2 = -(m_0 c^2)^2, \tag{1.41}$$

where m_0 is the rest mass of the particle. Note the norm of the momentum–energy four-vector is always negative (or "energy-like"), because the square of the rest energy is positive definite.

As an example of the utility of Lorentz transformations of the $(\vec{k}c, \omega)$ four-vector, we consider the process known as relativistic Thomson backscattering where a relativistic electron collides head-on with a photon (quantum of light, see Section 8.1), yielding a reversal of photon direction and an increase in the photon energy and momentum. This process is illustrated in Fig. 1.6, through a diagram of the initial and final momentum vectors of the electron and photon. The term Thomson is somewhat imprecisely[5] applied to this scattering process whenever the change in the electron momentum during collision is negligible. In the frame traveling with the electron, $\beta_f = v/c$ (v is the velocity of the electron), an oncoming photon of laboratory frequency ω has an observed frequency given by Lorentz transformation, $\omega' = \omega\gamma_f(1 + \beta_f)$. If we assume that this photon suffers a reversal of its momentum vector direction but no change in amplitude during collision, a second Lorentz transformation of the $(\vec{k}c, \omega)$ four-vector yields $\omega_s = \omega'\gamma_f(1 + \beta_f) = \omega\gamma_f^2(1 + \beta_f)^2$. Thus, for a highly relativistic ($\beta_f \approx 1$) electron, the frequency (energy) of backscattered light is increased, $\omega_s \cong 4\gamma_f^2\omega$. This scattering process is explored further in Exercise 1.6.

[5] It is more precise to term this process "inverse Compton scattering", as in the end we see that the photon energy, as observed in the laboratory frame, increases. This is the opposite of what happens in Compton scattering of photons off of electrons that are initially at rest. See Exercise 1.6 for further discussion of this point.

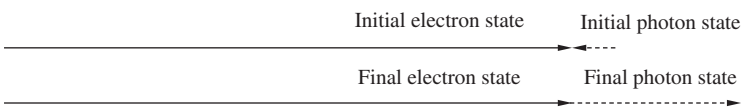


Fig. 1.6 Diagram of electron and photon momenta in initial and final states of the Thomson backscattering process. The violation of momentum conservation is exaggerated in this picture.

While we have now established the four-vector relationship between the energy and momentum of a physical system, we have not described what these quantities are in the familiar terms of mass and velocity. In doing so, we must obtain the well-known expressions of the non-relativistic limit, where the momentum $\vec{p} = m\vec{v}$ and energy $U = \vec{p}^2/2m + \text{const}$. We must also preserve the most general relationship between the momentum and energy,

$$\mathrm{d}U = \vec{v} \cdot \mathrm{d}\vec{p}, \quad \text{or less specifically,} \quad \frac{\mathrm{d}U}{\mathrm{d}p} = v, \tag{1.42}$$

where $p = |\vec{p}|$ and $v = |\vec{v}|$. This result is derived by noting that the energy change on a particle is equal to the work performed on it, $\mathrm{d}U = \vec{F} \cdot \mathrm{d}\vec{l}$, where the force is assumed to still obey Newton's third law, $\vec{F} = \mathrm{d}\vec{p}/\mathrm{d}t$, and the differential length is $\mathrm{d}\vec{l} = \vec{v} \cdot \mathrm{d}t$. Differentiating Eq. (1.40) and combining with Eq. (1.42), we also obtain

$$\vec{v} = \frac{\vec{p}c^2}{U}. \tag{1.43}$$

Then, solving for the energy as a function of the velocity, we may write

$$\left(\left(\frac{v}{c}\right)^2 - 1\right) U^2 = -(m_0c^2)^4 \quad \text{or} \quad U = \gamma m_0 c^2, \tag{1.44}$$

where we have now defined the Lorentz factor associated with the particle motion as $\gamma \equiv (1 - (\vec{v}^2/c^2))^{-1/2}$. The Lorentz factor of a particle is, therefore, its total (mechanical) energy normalized to its rest energy, and the condition $\gamma \gg 1$ implies a particle that travels at nearly the speed of light. For electrons, having rest energy $m_\mathrm{e}c^2 = 0.511\,\mathrm{MeV}$, it is very easy to obtain a particle that travels nearly at the speed of light—megavolt-class electrostatic accelerators can accomplish this feat. However, for the other most commonly accelerated particle, the proton (with rest energy $m_\mathrm{p}c^2 = 938\,\mathrm{MeV}$), it is relatively difficult to impart enough energy (several GeV) to make the particle relativistic.

Using Eqs (1.43) and (1.44), we also now have an expression for the momentum vector,

$$\vec{p} = \gamma m_0 \vec{v} \equiv \vec{\beta}\gamma m_0 c, \tag{1.45}$$

which allows Eq. (1.40) to be written, after removing the common factor of m_0c^2 in all quantities, as

$$\gamma^2 = \vec{\beta}^2\gamma^2 + 1. \tag{1.46}$$

Equation (1.40) is valid not only for single particle systems, but also for general systems of many ($j = 1, \ldots, N$) objects, in which case we have

$$\left(\sum_j \vec{p}_j\right)^2 c^2 - \left(\sum_j U_j\right)^2 = \text{constant}. \tag{1.47}$$

For such systems it is still true that, if any of the objects have non-zero rest mass, one may transform to a coordinate system in which the total momentum of the system vanishes, $\sum_j p_j = 0$.

In this frame, the total energy of the system is obviously minimized. This fact allows straightforward calculation of the available energy for particle creation

in high-energy physics experiments. Such a calculation serves as an illustrative example of the use of a Lorentz invariant norm.

In colliders, where charged particles and their antiparticles counter-circulate in rings (or collide after accelerating in opposing linacs), the colliding species, in general, have equal and opposite momentum. Therefore, the first term in Eq. (1.47) vanishes and all of the particle energy ($2U$) is available for creation of new particles. Thus, the Z^0 particle (with a rest energy of 91.8 GeV, 1 GeV = 10^9 eV) has been studied in detail using the LEP collider at CERN, by using electron and positron beams accelerated to 45.9 GeV and then collided. Before the era of the colliding beam machines, however, the frontier energies for exploring creation of new particles took place in **fixed-target** experiments where the beam particles struck stationary target particles. In the fixed target collision, one may calculate the Lorentz invariant on the right-hand side of Eq. (1.47) by evaluating the left-hand side in the lab frame,

$$p_{\mathrm{b}}^2c^2 - (U_{\mathrm{b}} + m_{\mathrm{t}}c^2)^2 = -m_{\mathrm{t}}^2c^4 - m_{\mathrm{t}}^2c^4 - 2\gamma_{\mathrm{b}}m_{\mathrm{b}}m_{\mathrm{t}}c^4 = \text{constant}, \quad (1.48)$$

where the subscripts b and t indicate beam and target particles, respectively. In the center of momentum frame, however, the total momentum vanishes and the constant in Eq. (1.48) is seen to set a maximum on the total rest energy of particles created in the collision,

$$\sum_i m_{p,i}c^2 \leq \sqrt{2\gamma_{\mathrm{b}}m_{\mathrm{b}}m_{\mathrm{t}}c^2}. \quad (1.49)$$

The maximum energy for particle creation in Eq. (1.49) occurs when the beam and target particles annihilate and the newly created particles are at rest in the center of momentum frame. Equation (1.49) clearly indicates why colliding beam machines are so important in exploring the energy frontier. At Fermilab, proton–antiproton collisions occur between counter-propagating beams of 900 GeV with up to 1.8 TeV available for creation of new particles.[6] If one, instead, substitutes a stationary proton as the target particle with a 900 GeV incident particle, the available energy for particle creation is only 41 GeV!

[6]Because protons and antiprotons are hadrons, composed of substituent particles (quarks and gluons), the effective (likely to observe) energy available for particle creation is considerably smaller than the full beam particle energy, and the "physics reach" of a 1.8-TeV hadron collider may be only in the several hundreds of GeV.

Equations (1.44) and (1.45) have introduced the relativistically correct momentum and energy. If one substitutes the relativistically correct form of the momenta into Eq. (1.9), it should be emphasized that this Lorentz force relation remains valid by construction. It is of interest to examine this vector equation of motion for the momentum in the case (discussed in detail in Section 2.1) where only a magnetic field is present. In such a scenario the energy of the particle is constant, and we may write

$$\gamma m_0 \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = q(\vec{v} \times \vec{B}). \quad (1.50)$$

Equation (1.45) displays, upon comparison with the non-relativistic version of Eq. (1.9), an effect known as the **transverse relativistic mass increase**, where the inertial mass under transverse (normal to the velocity) acceleration effectively behaves as though $m_0 \rightarrow \gamma m_0$. Note that this is also true for electric forces that are instantaneously transverse. For cases involving energy-changing acceleration (in which the acceleration is parallel to the velocity vector), the situation is different, as illustrated in Section 2.4.

With the results in this section thus far, obtained by emphasizing the invariance of the norms of a variety of four-vectors under Lorentz transformation, it is straightforward to find other relations more typically derived directly by use of Lorentz transformations. For example, let us examine the addition of velocities. Using the notation $v_\mathrm{f} = \beta_\mathrm{f} c$ to designate the relative velocity of a new frame, the velocity in the new frame of a particle whose velocity is v (parallel to v_f) in the original frame is

$$v' = \frac{p'c^2}{U'} = \frac{\gamma_\mathrm{f} c(pc - \beta_\mathrm{f} U)}{\gamma_\mathrm{f}(U - \beta_\mathrm{f} pc)} = \frac{\gamma_\mathrm{f}\gamma m_0 c^3(\beta - \beta_\mathrm{f})}{\gamma_\mathrm{f}\gamma m_0 c^2(1 - \beta_\mathrm{f}\beta)} = \frac{\beta - \beta_\mathrm{f}}{1 - \beta_\mathrm{f}\beta} c. \quad (1.51)$$

This derivation is, perhaps, more transparent than the more standard version based on Lorentz transformation of the components of $\vec{X}$, because it begins with more powerful concepts. It also points to an important facet of the theory of special relativity—velocities do not play a central role as useful descriptions of the motion since they do not form a part of a four-vector.

The momenta and energy, on the other hand, do form a four-vector. This is an interesting state of affairs within the context of particle dynamics, because these same quantities play a primary role in the Hamiltonian formulation of mechanics. We now discuss how the two concepts, Hamiltonian and four-vector dynamics, relate to each other in the context of charged particle dynamics in electromagnetic fields.

We begin by noting that the Lorentz force (Eq. (1.9)) acting on a charged particle is written, in terms of potentials, as

$$\begin{aligned} \vec{F}_\mathrm{L} &= \frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q(\vec{E} + \vec{v} \times \vec{B}) \\ &= q\left[-\vec{\nabla}\phi_\mathrm{e} - \frac{\partial \vec{A}}{\partial t} + \vec{v} \times (\vec{\nabla} \times \vec{A})\right] \\ &= q\left[-\vec{\nabla}\phi_\mathrm{e} - \frac{\partial \vec{A}}{\partial t} - (\vec{v} \cdot \vec{\nabla})\vec{A}\right] = q\left[-\vec{\nabla}\phi_\mathrm{e} - \frac{\mathrm{d}\vec{A}}{\mathrm{d}t}\right]. \end{aligned} \quad (1.52)$$

In the last line of this expression, we have used the definition of the total (partial plus convective) time derivative.

The forces in electromagnetic fields are most generally derived not only from a scalar (electrostatic) potential ϕ, but from an electromagnetic vector potential $\vec{A}$ as well; thus, we must generalize our approach to Hamiltonian analysis. In particular, we note that Eq. (1.52) indicates that the equations of motion for the momenta are not simply derivable from a conservative potential energy function. If we define the canonical momenta to be $p_{c,i} = p_i + qA_i$, however, the Hamilton equation of motion for this canonical momentum is correctly obtained by the prescription

$$\frac{\mathrm{d}p_{c,i}}{\mathrm{d}t} = -\frac{\partial H}{\partial x_i} = -q\frac{\partial \phi_\mathrm{e}}{\partial x_i}, \quad (1.53)$$

where the Hamiltonian contains only a conservative (electrostatic) potential energy. At this point, what we have considered (á la Newton) to be the momentum in the problem can now be seen to be a mechanical as opposed

to canonical momentum. The symmetry between the total "canonical" energy (value of the Hamiltonian) and the canonical momentum is clear—the total energy $H = U + V = U + q\phi_e$ (the numerical value of the Hamiltonian function) also has a portion arising from a potential (in this case, $q\phi_e$). Note that the remainder of the total energy (the "mechanical" component) also includes a rest energy component, $U = \gamma m_0 c^2 = T + m_0 c^2$, that is, the sum of rest and kinetic energies.

From this discussion, it should also be clear that the scalar and vector potentials, since they comprise components of the momentum–energy four-vector, also form a four-vector, $(\vec{A}c, \phi)$. Further, they are paired with the mechanical momenta and energy through the definitions of canonical momenta and energy, $p_{c,i} = p_i + qA_i$ and $H = U + q\phi$. To complete our survey of electrodynamic four-vector quantities, we note that the equations governing the potentials can be written as

$$\left[\vec{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\begin{Bmatrix}\vec{A}\\ \phi_e\end{Bmatrix} = -\begin{Bmatrix}\mu_0\vec{J}_e\\ \dfrac{\rho_e}{\varepsilon_0}\end{Bmatrix}. \tag{1.54}$$

Since the potentials form a four-vector, the sources $(\vec{J}_e, \rho_e c)$ form one as well. This result could also have been alternatively derived directly from charge conservation, and Lorentz transformation of lengths (and thus charge density) and velocities.

As we have just discussed sources and potentials associated with electromagnetic fields, it is appropriate at this point to examine the transformation between inertial frames of these fields. Since the potentials form a four-vector $(\vec{A}c, \phi_e)$, an obvious starting point of the discussion is to discuss the Lorentz transformation of this four-vector. This relation, governing the transformation of potentials from one frame to another that moves at speed $\beta_f c$ along what we again choose to be the z-axis is written as

$$cA'_x = cA_x, \quad cA'_y = cA_y, \quad cA'_z = \gamma_f(cA_z - \beta_f\phi_e), \quad \phi'_e = \gamma_f(\phi_e - \beta_f cA_z). \tag{1.55}$$

The quantities A_x, A_y, A_z, and ϕ are all, in principle, functions of x, y, z, and t. In the new (primed) frame, the spatio-temporal dependence of these quantities must be expressed in terms of the primed variables, found by substitutions obtained from the inverse Lorentz transformation

$$x = x', \quad y = y', \quad z = \gamma_f(z' + \beta_f ct'), \quad t = \gamma_f(ct' + \beta_f z'). \tag{1.56}$$

The fields are obtained from this expression using the following relations:

$$\vec{E}' = -\vec{\nabla}'\phi'_e - \frac{\partial\vec{A}'}{\partial t'}, \qquad \vec{B}' = \vec{\nabla}' \times \vec{A}'. \tag{1.57}$$

The transformation of the electromagnetic fields described by Eqs(1.55)–(1.57) are often written as

$$\begin{aligned}\vec{E}'_\perp &= \gamma_f(\vec{E}_\perp + \vec{v}_f \times \vec{B}_\perp), \qquad \vec{E}'_{||} = \vec{E}_{||},\\ \vec{B}'_\perp &= \gamma_f\left(\vec{B}_\perp - \frac{1}{c^2}\vec{v}_f \times \vec{E}_\perp\right), \qquad \vec{B}'_{||} = \vec{B}_{||},\end{aligned} \tag{1.58}$$

where the symbols || and $\perp$ indicate the components of the field parallel to and perpendicular to the direction of Lorentz transformation of the frame $\vec{v}_{\rm f}$, respectively.

Now we can return to the derivation of relativistic Lagrangian and Hamiltonian mechanics with electromagnetic fields. With our definition of canonical momenta, we can proceed to construct the Lagrangian by integrating the expression that defines these momenta,

$$p_{c,i} = \frac{\partial L}{\partial \dot{x}_i} = \gamma m_0 \dot{x}_i + eA_i, \tag{1.59}$$

with respect to the spatial coordinates. We thus obtain

$$L(\vec{x}, \dot{\vec{x}}) = -\frac{m_0 c^2}{\gamma} + q\vec{A} \cdot \vec{v} - q\phi_{\rm e}(\vec{x}), \tag{1.60}$$

where we have allowed the presence in the Lagrangian of a conservative potential dependent only on $\vec{x}$, and identified it as the negative of the electrostatic potential energy $-q\phi_{\rm e}(\vec{x})$.

The relativistically correct Hamiltonian is obtained from Eq. (1.60) by use of the definition

$$\begin{aligned} H = \vec{p}_{\rm c} \cdot \dot{\vec{x}} - L &= \frac{\vec{p}_{\rm c} \cdot (\vec{p}_{\rm c} - q\vec{A})}{\gamma m_0} - L \\ &= \frac{(\vec{p}_{\rm c} - q\vec{A})^2}{\gamma m_0} + \frac{m_0 c^2}{\gamma} + q\phi_{\rm e}(\vec{x}). \end{aligned} \tag{1.61}$$

By multiplying this expression by $\gamma m_0 c^2$ and using $H - q\phi = \gamma m_0 c^2$, we arrive at

$$(H - q\phi_{\rm e})^2 = (\vec{p}_{\rm c} - q\vec{A})^2 c^2 + (m_0 c^2)^2, \tag{1.62}$$

or

$$H = \sqrt{(\vec{p}_{\rm c} - q\vec{A})^2 c^2 + (m_0 c^2)^2} + q\phi_{\rm e}. \tag{1.63}$$

We note that Eq. (1.62) could have been obtained by direct substitution of canonical definitions into Eq. (1.41) governing the norm of the mechanical momentum–energy four-vector, that is,

$$(\vec{p}_{\rm c} - q\vec{A})^2 c^2 - (H - q\phi_e)^2 = -(m_0 c^2)^2. \tag{1.64}$$

1.5 Hierarchy of beam descriptions

The methods for analyzing single particle dynamics given in Section 1.4 represent the first step in understanding the physics of charged particle beams. A real beam is made up not of a single particle, however, but a collection of many (N) particles. The second step towards describing the dynamics of an actual beam, therefore, is to consider a collection of N points in phase space, as illustrated in the phase plane plot shown in Fig. 1.7. It is not obvious how to proceed with the description of such a system, where the phase space has $2NM$ variables. It is therefore now necessary to discuss a hierarchy of descriptions that begin with single particle dynamics.

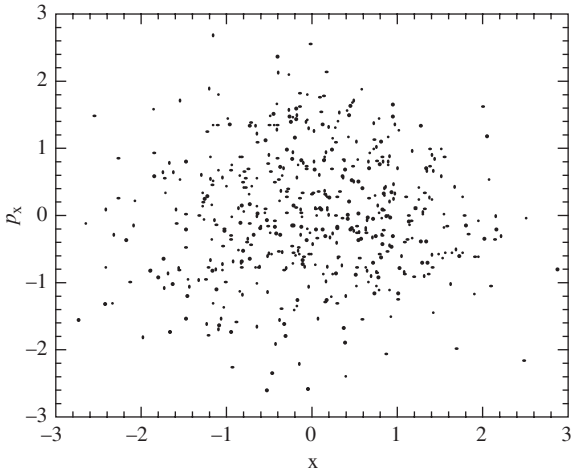


Fig. 1.7 Distribution of particles in (x, p_x) phase plane.

For many particle beams, the density of particles in phase space is small enough that the particles are essentially a non-interacting ensemble, with both macroscopic and microscopic electromagnetic fields created by the particles themselves contributing insignificantly to the motion. In this case, one only needs to solve the single particle equations of motion in the presence of applied forces, and then proceed to produce a collective description of the beam ensembles' evolution based on the known single-particle orbits. This book assumes the validity of this type of description, which straightforwardly leads to analyses based the beam's distribution. In real charged particle beams, as well as in the modeling of such beams in multi-particle computations, this distribution is discrete, as illustrated in Fig. 1.7. On the other hand, for analytical approaches, the distribution is viewed as a smooth probability function[7] in a $2M$-dimensional phase space $f(\vec{x}, \vec{p}, t)$, that is, the number of particles found in a differential phase space volume $\mathrm{d}V = \mathrm{d}^3\vec{x}\,\mathrm{d}^3\vec{p}$ in the neighborhood of a phase space location $\vec{x}, \vec{p}$ at a time t is simply given by $f(\vec{x}, \vec{p}, t)\,\mathrm{d}V$. While the computational approach to multi-particle dynamics is beyond the scope of this book, analytical approaches based on the distribution function $f(\vec{x}, \vec{p}, t)$ will be introduced in Chapter 5. One result concerning phase space distributions deserves prominent discussion at this point, however, the conservation of phase space density, or Liouville's theorem.

[7]With a smooth phase space distribution, the charge and current distributions associated with such a distribution are also continuous and smooth. The fields derived from the smooth charge/current densities may be termed macroscopic. Deviations from these approximate fields (near an individual particle) may be termed microscopic.

To begin, we write the total time derivative of the phase space distribution function,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \dot{\vec{x}} \cdot \vec{\nabla}_{\vec{x}} f + \dot{\vec{p}} \cdot \vec{\nabla}_{\vec{p}} f, \tag{1.65}$$

where the second and third terms on the right-hand side of Eq. (1.62) are the convective derivatives in phase space, derived simply by the chain rule. (Note: the subscript on the gradient operators indicates differentiation with respect to either coordinate or momentum components.) If the forces are derivable from

a Hamiltonian, then

$$\begin{aligned}\frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial f}{\partial t} + \sum_i \left(\frac{\mathrm{d}x_i}{\mathrm{d}t}\frac{\partial f}{\partial x_i} + \frac{\mathrm{d}p_i}{\mathrm{d}t}\frac{\partial f}{\partial p_i} \right) \\ &= \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial H}{\partial p_i}\frac{\partial f}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial f}{\partial p_i} \right) \\ &= \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial H}{\partial p_i}\frac{\partial H}{\partial x_i}\frac{\mathrm{d}f}{\mathrm{d}H} - \frac{\partial H}{\partial x_i}\frac{\partial H}{\partial p_i}\frac{\mathrm{d}f}{\mathrm{d}H} \right) = \frac{\partial f}{\partial t}. \end{aligned} \tag{1.66}$$

Because we have assumed that the forces are derived from a Hamiltonian, it can be seen that the summation in Eq. (1.66) vanishes, and $\mathrm{d}f/\mathrm{d}t = \partial f/\partial t$. In Eq. (1.65), it should be understood that the total derivative is taken by moving about phase space according to the particle's equations of motion at a given phase space point $(\vec{x}, \vec{p})$. If neither creation nor destruction of the particles are allowed, we also have, by drawing a small constant differential volume $\mathrm{d}V$ around a given particle, $f(\vec{x}, \vec{p}, t) = \mathrm{d}V^{-1}$ in the neighborhood of this point. Then we have the result that $\partial f/\partial t = 0$ and

$$\frac{\mathrm{d}f}{\mathrm{d}t} = 0. \tag{1.67}$$

This result is termed Liouville's theorem, and it states that the phase space density encountered as one travels with a particle in a Hamiltonian system is conserved. The derivation of Eq. (1.67) may seem to be tautological, since we stated that $\partial f/\partial t = 0$ without referring to any property of Hamiltonian. Equation (1.67) is more illustrative than it seems, however, because the properties of the Hamilton equations guarantee that the density of any volume of phase space whose boundary follows these equations is constant. An alternative point of view, suggested by Eq. (1.65), is that phase space itself is incompressible. Perhaps the most common statement derived from interpretation of Eq. (1.67) is that the volume occupied by particles in phase space is conserved. For systems in which the motion of all phase planes is uncorrelated, the motion in the separate phase planes is independent. Then one can state that the emittance or area occupied by particles in a phase plane is conserved. The conservation of emittance, which is discussed further in Chapter 5, plays a very important role in the theory and design of particle accelerators.

Here, we concentrate on cases where Liouville's theorem holds, since we will almost exclusively consider motion due to applied forces that are derivable from a Hamiltonian. Deviations from this physical scenario form much of the following volume on advanced subjects in beam physics so we restrict comment on such topics to a few general statements. When multi-particle interaction effects become important, there are two distinct regimes to consider. The first can be described as one in which the collective macroscopic fields arising from the bulk beam charge and current density is the dominant self-interaction mechanism of the beam. These self-fields are most often termed space-charge when they arise from the near-field of the beam's charge distribution and wake-fields when they arise from the beam's collectively radiated fields. The evolution of the beam distribution can be dominated by space-charge fields, in which case one can use a description that is based on the notions of cold fluid motion

developed in the field of plasma physics. In this scenario, simple collective (plasma) oscillations are encountered. When only macroscopic beam fields are important, these fields may, in principle, be included into the Hamiltonian, Eq. (1.67) is unchanged, and the implications of Liouville's theorem still hold.

On the other hand, microscopic binary collisions may strongly affect the beam distribution evolution, and diffusion (heating) may occur in phase space. There are a number of methods for dealing with this complication, but most are beyond the scope of the present discussion. However, one method is relatively easy to appreciate in the context of this section. An explicit non-Hamiltonian term may be introduced on the right-hand side of Eq. (1.67), which accounts for the time derivative of the distribution due to collisions (the effects of microscopic fields), so that $\mathrm{d}f/\mathrm{d}t = \partial f/\partial t|_{\mathrm{coll}}$. In this generalized form, the revised version of Eq. (1.67) is termed the Boltzmann equation. Thus, Eq. (1.67), which is commonly known (when the full derivative is written as in Eq. (1.65)) as the Vlasov equation, is also referred to as the collisionless Boltzmann equation. The non-macroscopic-Hamiltonian physics arising from emission of radiation by particle beams gives rise to dissipation and damping of phase space trajectories. The description of the effect of such phenomena on the distribution function may be treated with a full Boltzmann equation approach, or though an analysis based on the Fokker–Planck equation.

1.6 The design trajectory, paraxial rays, and change of independent variable

While a rigorous analysis of classical motion can be performed, as discussed above, using canonical momenta and coordinates within the confines of the Hamiltonian formalism, one often finds it useful in practice to use a more physically transparent description. This is obtained by use of paraxial rays, vector representations of the local trajectory which, by definition, have an angle with respect to a design trajectory that is much smaller than unity. Trajectories of interest in beam physics are always paraxial—one must confine the beam inside of small, near-axis regions, such as the drift tubes shown in Fig. 1.1.

Both the paraxial ray and the design trajectory are illustrated in Fig. 1.8. In this figure, an example of a design trajectory, defined as the ideally preferred trajectory—a locally straight or curved line— through the system, is displayed. As we shall see, one defines the coordinate system for analyzing beam physics problems locally by use of the design trajectory. As a result, the coordinate systems we encounter in this text may naturally be locally Cartesian, or locally curvilinear (i.e. inside of bend magnets). Curvilinear coordinate systems are obviously found in circular accelerators where the design trajectory is closed upon itself and is often referred to as the design orbit.

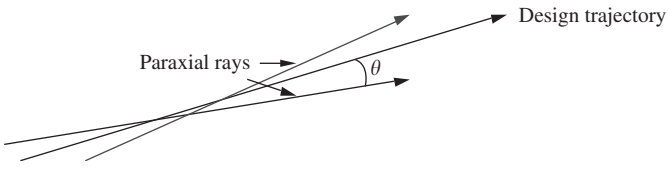


Fig. 1.8 Design trajectory and examples of small angle paraxial rays.

Figure 1.8 also displays examples of paraxial rays. The ray is a useful visualization tool and can be described mathematically by a coordinate offset, and an angle θ. In a locally Cartesian coordinate system,[8] we take the distance along the design trajectory to be z. In this system, a horizontal offset is designated by x and the horizontal projection of the angle is θ_x. This angle is given in terms of the momenta as

$$\tan\theta_x = \frac{p_x}{p_z} = \frac{v_x}{v_z}. \tag{1.68}$$

Analogous definitions hold for the other transverse offset and angle, which we term vertical and indicated by the variable y. We are ultimately interested in a description of particle dynamics that uses the distance along the design trajectory as the independent variable. The reason for this is straightforward; in an optics system (for charged particles or photons) the forces encountered are always specified in space and not in time. Thus, a spatial description is more efficient and natural.

In order to use z as the independent variable, we must be able to write equations of motion in terms of z. This is accomplished by writing total time derivatives (written in compact notation as $(\cdot) \equiv \mathrm{d}/\mathrm{d}t$) in terms of total spatial derivative in z,

$$(\,)' \equiv \frac{\mathrm{d}}{\mathrm{d}z} = \frac{1}{v_z}\frac{\mathrm{d}}{\mathrm{d}t}. \tag{1.69}$$

The derivative of a horizontal offset with respect to z is given by

$$x' = \frac{\mathrm{d}x}{\mathrm{d}z} = \tan(\theta_x). \tag{1.70}$$

When analyzing beams, which are by definition are collections of rays localized in offset and angle near the design orbit, it can be seen from Eq. (1.70) that the transverse momentum (normal to the design orbit) is much smaller than the longitudinal momentum, $p_{x,y} \ll p_z \cong |\vec{p}|$. As a consequence, we may in most cases use the small angle, or paraxial, approximation, which allows us to write a series of useful approximate expressions,

$$x' = \tan(\theta_x) \cong \theta_x \cong \sin(\theta_x) \ll 1. \tag{1.71}$$

It is not immediately apparent how this assumption restricts the offset from the design orbit. As we shall see, all particle beam optics systems that focus and control the rays fundamentally resemble simple harmonic oscillators, where the transverse force is linear in offset, $F_x = -Kx$. Using this analogy, we can write a model equation for the transverse offset

$$\ddot{x} + \omega^2 x = 0, \tag{1.72}$$

where $\omega^2 = K/\gamma m_0$. Further, Eq. (1.72) can be cast in terms of a ray description, using Eq. (1.70), to obtain

$$x'' + k^2 x = 0. \tag{1.73}$$

Here, $k \equiv \omega/v_z$ is the characteristic oscillation wavenumber of the optics system. The solutions of Eq. (1.73) are of the form $x = x_\mathrm{m}\cos(kz+\phi)$, yielding an approximate angle of $kx_\mathrm{m}\sin(kz+\phi)$, and so the paraxial approximation is obeyed for offsets $x_\mathrm{m} \ll k^{-1}$. Another restriction on the validity of a paraxial description is relevant in the case of a curvilinear design orbit. In this case, we

[8] When the design orbit is derived from rectilinear motion, we will choose z as the distance along the design orbit, in part to easily connect to field descriptions written in cylindrical coordinates. When the design orbit is bent, we will emphasize the curvilinear nature of the coordinate system by using the variable s as the distance along the design orbit.

must require that $x_m \ll R$ (where R is the local radius of curvature of the design trajectory) for the paraxial approximation to hold.

When one uses paraxial equations of the physically transparent form shown in Eqs (1.72) and (1.73), the analysis strays somewhat from a rigorous phase space description. In fact, one often uses this paraxial formalism instead of a canonically correct approach, even to the point of replacing a momentum (e.g. p_x) with the angle (x') in phase plane plots. In this case, when we make a plot of the trajectory in, for example, the (x, x') plane, which is an example of a **trace space**, we construct what is termed a **trace space plot**. This does not introduce complications in understanding the motion at constant values of p_z, because we can always recover the transverse momentum by using $p_x = p_z x' \cong \beta\gamma m_0 c x'$. If longitudinal acceleration occurs, however, the angle is diminished and an apparent damping (so-called **adiabatic damping**, see Section 2.6) of the motion is observed.

With a change of independent variable evidently of high interest to particle beam physics, we must revisit the question of Hamiltonian analysis in this context. In order to proceed with transformation of the Hamiltonian into one with a new independent variable, we first note that the symmetry of Eqs (1.62) and (1.64) (in which the energy and momenta are on equal footing) is suggestive of a notion—the choice of the energy function as the Hamiltonian is a bit arbitrary as well. We could just as naturally have chosen one of the momenta as the Hamiltonian from which we derive equations of motion, and the form of Eq. (1.62) would remain the same, except for a minus sign. The process of changing the Hamiltonian from an energy-based function to a momentum-based function also requires that the independent variable be changed from t to the coordinate canonical to the new Hamiltonian. This can be illustrated by referring to the stationary property of the action integral (Eq. (1.11)) which gives rise to the equations of motion,

$$\delta \int_{t_1}^{t_2} L \, \mathrm{d}t = \delta \int_{t_1}^{t_2} (\vec{p}_c \cdot \dot{\vec{x}} - H) \, \mathrm{d}t = 0. \tag{1.74}$$

We can accomplish the change of independent variable, from t to z, by rewriting Eq. (1.74) as follows:

$$\delta \int_{z(t_1)}^{z(t_2)} (p_{c,x} x' + p_{c,y} y' - H t' + p_{c,z}) \, \mathrm{d}z = 0, \tag{1.75}$$

where again the prime indicates differentiation with respect to the new independent variable, $(\,)' \equiv \mathrm{d}/\mathrm{d}z$. In Eq. (1.64), the role of the new Hamiltonian is obviously played by $-p_{c,z} \equiv G$,

$$G = -p_{c,z} = \sqrt{(H - q\phi_e)^2 - (p_{c,y} - qA_y)^2 c^2 - (p_{c,x} - qA_x)^2 c^2 - (m_0 c^2)^2}, \tag{1.76}$$

the role of the third coordinate is taken by $-t$, and the role of the third momentum by H. Note that this statement implies that H and $-t$ are canonically conjugate.

This type of independent variable transformation is a canonical transformation by design, and is used in applications, such as particle beam dynamics, where the applied forces are described more naturally by functions of a spatial coordinate than by functions of time. Because the transformation is canonical,

the Hamilton's equations of motion are obtained in the usual fashion,

$$p'_{c,x} = -\frac{\partial G}{\partial x} = \frac{\partial p_{c,z}}{\partial x}, \quad p'_{c,y} = -\frac{\partial G}{\partial y} = \frac{\partial p_{c,z}}{\partial y}, \quad H' = \frac{\partial G}{\partial t} = -\frac{\partial p_{c,z}}{\partial t},$$
$$x' = \frac{\partial G}{\partial p_{c,x}} = -\frac{\partial p_{c,z}}{\partial p_{c,x}}, \quad y' = \frac{\partial G}{\partial p_{c,y}} = -\frac{\partial p_{c,z}}{\partial p_{c,y}}, \quad t' = -\frac{\partial G}{\partial H} = \frac{\partial p_{c,z}}{\partial H}. \tag{1.77}$$

For problems in which one of the spatial coordinates (z in particle beam physics) is taken to be the independent variable, it is assumed that the motion can be followed monotonically in this coordinate. Note that this is always the case for time as an independent variable (as far as we know!). If the trajectory describing the motion is not a monotonic function of the independent variable, then it is not uniquely described by this variable. Fortunately, in particle beams this assumption is also always correct.

1.7 Summary and suggested reading

In this chapter, we have motivated much of the contents of this book by introducing particle accelerators in their scientific and historical context. In order to build up the tools needed to analyze the dynamics of charged particle beams, we have reviewed methods in Lagrangian and Hamiltonian dynamics as well as special relativity, in a unified way. These general subjects gave way to concepts more specific to describing the motion charged particles in beams. We have examined the notion of phase space, and the conservation of its density in Hamiltonian systems—the Liouville theorem. We introduced the concept of the design trajectory, an ideal trajectory through an accelerator or transport system, which allows nearby trajectories (paraxial rays) to be defined. The design trajectory also gives one the freedom to analyze the motion using distance along such a trajectory as the independent variable, instead of time. This way of approaching description of trajectories is natural in beam optics and other problems in beam physics.

The subject of classical mechanics can be reviewed in a recommended number of texts:

1. K.R. Symon, *Mechanics* (Addison-Wesley, 1971). A classic and readable undergraduate text that introduces variational methods.
2. J.B. Marion and S.T. Thornton, *Classical Dynamics of Particles and Systems* (Harcourt Brace & Company, 1995). A more expansive treatment of mechanics for advanced undergraduate readers.
3. H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (Addison-Wesley, 2002). One of the best graduate treatments of variational principles in mechanics.
4. L. Michelotti, *Intermediate Classical Mechanics with Applications to Beam Physics* (Wiley, 1995). A fairly complete introduction to modern methods in nonlinear dynamics, with an excellent array of problems and, as a bonus, has detailed examples in beam physics. Written at the graduate level.

Electromagnetic theory is central to this book, and provides the natural context for introducing special relativity. A selection of useful texts includes:

5. R. Wangsness, *Electromagnetic Fields* (Wiley, 1986). A complete, junior level introductory text.
6. R. Resnick, *Introduction to Special Relativity* (Wiley, 1968). A nice primer on special relativity.
7. M.A. Heald and J.B. Marion, *Classical Electromagnetic Radiation* (Harcourt Brace & Company, 1995). An advanced undergraduate course book.
8. J.D. Jackson, *Classical Electrodynamics* (Wiley 1975). The standard in graduate texts.
9. L.D. Landau and E.M. Lifschitz, *The Classical Theory of Fields* (Addison-Wesley, 1971). A dense, deep investigation for graduate students.
10. J. Schwinger, L.L. Deraad, and K. Milton, *Classical Electrodynamics* (Perseus, 1998). This text is a fine alternative to Jackson's, and a bit more modern.

This chapter has introduced some of the more basic notions of particle accelerators, including phase space descriptions and the design trajectory. Other recommended introductory texts, which will also serve as references to the following chapters, may include:

11. M.S. Livingston, *Particle Accelerators, Advances in Electronics I*, pp. 269–331 (Academic Press, 1948). This monograph gives an overview of the history of particle accelerators until 1948, and has an excellent reference list pointing to the primary articles in early accelerator scient. It anticipates the rise in importance of the proton synchrotron.
12. D. Edwards and M. Syphers, *An Introduction to the Physics of High Energy Accelerators* (Wiley, 1993). A fine primer on high-energy devices, and good reference book, especially on circular accelerators. Written at senior undergraduate level.
13. H. Wiedemann, *Particle Accelerator Physics I: Basic Principles and Linear Beam Dynamics* (Springer-Verlag, 1993). Very comprehensive and rigorous book, concentrating on circular accelerators. Written at the graduate level.
14. S.Y. Lee, *Accelerator Physics* (World Scientific, 1996). A comprehensive and sweeping introduction, especially in the areas of circular accelerators. Written at the graduate level.
15. M. Reiser, *Theory and Design of Charged Particle Beams* (Wiley, 1994). An in-depth treatment of beam optics and transverse collective effects, with emphasis on rigorous analytical methods. Written at the graduate level.
16. S. Humphries, Jr., *Principles of Charged Particle Acceleration* (Wiley, 1986). This is a strongly pedagogical text with good physics underpinning.
17. M. Sands, *The Physics of Electron Storage Rings: an Introduction* (Stanford, 1970). A ground-breaking book, with clear explanations of basic storage ring physics and radiation damping effects. Written at the graduate level.

There are a number of books one may examine to get a flavor for collective effects in charged particle beams, subjects that have been deemed to lie outside of this text's scope:

18. J.P. Lawson, *The Physics of Charged Particle Beams* (Clarendon Press, 1977). Very physics oriented, with excellent melding of plasma physics notions into the text. Written at the graduate level.
19. S. Humphries, Jr., *Charged Particle Beams* (Wiley, 1990). A unified presentation of the physics of high power and high brightness beams. Written at the graduate level, it is a companion to Ref. 16, by the same author.
20. R. B. Miller, *Intense Charged Particle Beams* (Plenum, 1982). A good first look at very high current beams, with collective forces emphasized. Written at the graduate level.
21. A. Chao, *Physics of Collective Beam Instabilities in High Energy Accelerators* (Wiley, 1993). A classic treatment of instabilities, very broad in scope and rigorous in approach. Written at the graduate level.

The are some general references that will aid a student or practitioner in the field:

22. A. Chao and M. Tigner, editors, *Handbook of Accelerator Physics and Engineering* (World Scientific, 1999). Not a text, but a vade mecum for the accelerator field, with summaries of basic principles, and useful formulae covering almost every conceivable aspect of accelerators. For professional use.
23. *The Particle Data Book*, an annually-updated updated reference for high-energy physics, available from Lawrence Berkeley National Laboratory at http://pdg.lbl.gov/.

Exercises

(1.1) Consider the three-dimensional simple harmonic oscillator where $T = \frac{1}{2}m\dot{\vec{x}}^2$ and $V = \frac{1}{2}K\vec{x}^2$.

(a) Construct the Lagrangian of this system in Cartesian coordinates, $\vec{x} = (x, y, z)$.
(b) Derive the canonical momenta and construct the Hamiltonian for this system. Show that the total energy of the system is conserved, along with the energy associated with motion in each phase plane.
(c) Construct the Lagrangian of this system in cylindrical coordinates (r, z, ϕ), noting that $\dot{\vec{x}} = \dot{\rho}\hat{\rho} + \dot{z}\hat{z} + \rho\dot{\phi}\hat{\phi}$ and that $\vec{x}^2 = \rho^2$.
(d) Construct the Hamilton for this system in cylindrical coordinates. From inspection of the Hamiltonian, deduce any constants of the motion associated with the angular momentum.
(e) Construct the Lagrangian of this system in spherical polar coordinates (r, θ, ϕ), noting that $\dot{\vec{x}} = \dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + r\sin\theta\dot{\phi}\hat{\phi}$ and that $\vec{x}^2 = r^2$.
(f) Construct the Hamilton for this system in spherical polar coordinates. From inspection of the Hamiltonian, deduce any constants of the motion associated with angular momenta.

(1.2) In order for the motion to be stable in the simple harmonic oscillator case illustrated in Fig. 1.4, the force must be restoring or $K > 0$. Assuming an unstable system, however, we have $K < 0$. Plot the curves corresponding to $H = -\frac{1}{8}, -\frac{1}{2}$, and -2 in (x, p_x) phase space. Note that the curves are not closed, indicating unbounded motion.

(1.3) While the forces most commonly associated with charged particle motion in accelerators are of the form of a simple

harmonic oscillator, meaning they are linearly proportional to distance from an equilibrium point (i.e. $F_x = -Kx$), other types of forces may be present. For instance, in a sextupole magnet, the force is of the form, $F_x = -ax^2$, where a is a constant.

(a) What is the Hamiltonian associated with one-dimensional motion under this applied field?
(b) Draw some representative constant H curves in (x, p_x) phase space for this Hamiltonian. Comment on whether the motion is bounded or unbounded (see Example 1.2).
(c) In an octupole-like field, the force is of the form $F_x = -ax^3$. Construct the Hamiltonian and plot some constant H curves. Consider the effect of changing the sign of the constant a and discuss whether the motion is bounded or unbounded.

(1.4) Consider an undamped oscillator consisting of a weight hanging from a spring. This spring is set in motion and has a certain action J and frequency ω. If a window to the cool outside air is opened in the room containing this oscillator, the spring becomes colder and more rigid, causing ω to slowly rise. As J is an adiabatic invariant, the total energy in the oscillator grows as the thermal energy is removed from the spring! Explain. (Hint: consider the microscopic internal degrees of freedom of the spring.)

(1.5) Show, by Taylor expansion of the particle mechanical energy, $U = \gamma m_0 c^2$ in terms of the velocity when $v \ll c$, that this total energy is approximately the sum of the particle rest energy and the non-relativistic expression for the kinetic energy.

(1.6) The Thomson backscattering analysis of photons by electrons given above is only approximate and applies when the incident photon energy is very small compared to the rest energy of the electron, $m_0 c^2 = 0.511\,\text{MeV}$. This analysis, referred as Compton scattering, should be familiar to the reader. One can thus recognize that Thomson scattering, where the scattered photon has the same frequency as the incident photon in the electron rest frame, is only an approximate description—a limit of the general Compton scattering phenomenon for low-energy photons.

(a) The usual Compton scattering analysis is performed in the electron rest frame. From the Lorentz transformation above, find an expression of electron energy $\gamma m_e c^2$ and incident (counter-propagating) photon energy $\hbar\omega$ such that the photon energy in the electron rest frame is less than 10 per cent of the electron rest energy, $\hbar\omega' < m_e c^2/10$.
(b) By first performing (in the electron rest frame) the usual Compton scattering analysis of scattered wavelength λ' as a function of angle θ', and then performing a Lorentz transformation on the results back to the lab frame, find the scattered wavelength as a function of laboratory angle.
(c) Now consider the photon to be incident in the lab frame at some arbitrary angle θ_i, where $\theta_i = 0$ is defined as the counter-propagating case. Find the angle and energy of the incident photon in the electron rest frame by Lorentz transformation. Find the energy of the backscattered light (copropagating with the electron velocity) as a function of θ_i, using a full Compton analysis.
(d) With the Thomson assumption that (in the electron rest frame) the frequency does not change during the scattering event, perform the same analysis as in part (c). For an energy in the electron rest frame of $\hbar\omega' = m_e c^2/10$, compare the backscattered energies given by the Thomson and Compton analyses.

(1.7) Using the Lorentz transformation of the $\vec{P}$ four-vector, generalize the addition of velocities expression given by Eq. (1.51) to account for any arbitrary angle (set to zero in Eq. (1.51)) between the particle and frame velocities.

(1.8) Consider a stationary uniform cylinder of charge that has charge density ρ_e up to a radius a, and then vanishes outside of this radius.

(a) From Gauss' law, find the radial electric field associated with this charge distribution.
(b) Now assume that this distribution is in motion along its symmetry axis with speed v and with respect to the lab frame. From Lorentz transformation-derived rules for determining the fields (Eq. (1.58)), find the electric and magnetic fields in the lab frame associated with the moving charge distribution.
(c) From Lorentz transformation of the charge–current four-vector, find the density ρ_e' and current density $J_{e,z}'$ associated with the moving charge distribution.
(d) Using Gauss' and Ampere's law directly, calculate the radial electric field, and the azimuthal magnetic field associated with the moving charge distribution. Compare with your answer in part (b).
(e) What is the net radial force on a particle inside of the beam (that has velocity v)?

(1.9) For a beam particle acted upon by the forces given in Exercise 1.8:

(a) Determine the scalar and vector potentials. (Hint: find the scalar potential in the rest frame of the beam first.)
(b) Construct the Lagrangian for this system.
(c) Derive the Hamiltonian for the system in Exercise 1.8(c).

(d) Derive the radial equations of motion for this particle.
(e) If you change the sign of the particle but not its velocity, the radial motion of the particle is approximately simple harmonic for motion for $\rho < a$. Under what conditions on ρ_e, v, and a is this approximation valid? Assuming these conditions, plot the phase space using constant H curves both inside and outside of the beam.

(1.10) Non-equilibrium solutions to the Vlasov equation are generally quite difficult, but it is possible to make a very powerful statement about solutions for cases in equilibrium, where $\partial f/\partial t = 0$. As an example of the analysis of a Vlasov equilibrium, show that for a time-independent Hamiltonian, $H(\vec{x}, \vec{p})$,

$$f(\vec{x}, \vec{p}) = g[H(\vec{x}, \vec{p})],$$

is an equilibrium solution to the Vlasov equation, where g is any differentiable function of the Hamiltonian. This result will be important to the discussion of equilibrium distributions in Chapter 5 (cf. Ex. 5.2).

(1.11) A practical check on the paraxial approximation is to see that the error made in assuming that the longitudinal velocity is approximately the total velocity, $v_z \cong v$, is small.

(a) Assuming $p_{x,y} \ll p_z$, write v as a function of p_x, p_y, and p_z. Taylor expand this expression to second order in p_x/p_z and p_x/p_z (x' and y').
(b) For what ratio of $\sqrt{p_x^2 + p_y^2}/p_z$ is the error in the expression $v_z \cong v$ kept to 0.1 per cent?

(1.12) For the Hamiltonian derived in Exercise 1.9 for the unlike-sign particle case:

(a) Perform a canonical transformation to use z as the independent variable.
(b) Derive the equations of motion for the radial and angular coordinates and momenta with z as the independent variable.
(c) Assuming the angular momentum of the particle is zero, construct (r, r') trace space plots for motion both inside and outside of the beam.
(d) What is the equation of motion for H? What about t?

2 Charged particle motion in static fields

2.1 Charged particle motion in a uniform magnetic field 30
2.2 Circular accelerator 32
2.3 Focusing in solenoids 35
2.4 Motion in a uniform electric field 37
2.5 Motion in quadrupole electric and magnetic fields 40
2.6 Motion in parallel, uniform electric and magnetic fields 42
2.7 Motion in crossed uniform electric and magnetic fields* 44
2.8 Motion in a periodic magnetic field* 45
2.9 Summary and suggested reading 47
Exercises 48

Many of the physical scenarios that unfold in the sometimes complex-looking world of particle beams and accelerators can be understood by first referring to a simplified situation, from which the essential physics can be extracted. The purpose of this chapter is to survey the most illustrative of these situations, where only static fields play a role. We thus examine the motion of charged particles in: uniform (dipole) magnetic and electric fields, both separately and in combination; magnetic dipole fields that have periodic longitudinal variation of the dipole orientation (undulators); and magnetic and electric quadrupole fields. These analyses will serve to both introduce physical concepts, and to develop mathematical techniques which are needed for the discussions in following chapters.

2.1 Charged particle motion in a uniform magnetic field

In a uniform (dipole) magnetostatic field, with field direction chosen along the z-axis, $\vec{B} = B_0\hat{z}$, the Lorentz force equation governing the time evolution of the momentum can be written in two components, one parallel and one perpendicular to the field,

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = 0, \qquad \frac{\mathrm{d}\vec{p}_\perp}{\mathrm{d}t} = q(\vec{v}_\perp \times \vec{B}) = \frac{qB_0}{\gamma m_0}(\vec{p}_\perp \times \hat{z}). \tag{2.1}$$

One immediately sees from Eq. (2.1) that p_z is a constant of the motion and, further, since the magnetic force does no work on the particle, the total momentum $p = \sqrt{p_z^2 + p_\perp^2}$ (and therefore the total energy $\gamma m_0 c^2$) is also constant. Thus the transverse momentum must also be constant in amplitude, but changing in direction. It is well known that this implies circular motion in the plane perpendicular to the magnetic field. In general, therefore, the motion is the superposition of a circular orbit and a uniform drift normal to the circle—in other words, it is helical. Although the circular orbit traced by the particle in the plane perpendicular to the field is familiar from non-relativistic mechanics, we now examine this motion in more detail in the context of relativistic particles.

One approach is to write the differential equations governing the transverse velocities,

$$\frac{\mathrm{d}v_x}{\mathrm{d}t} = \frac{qB_0}{\gamma m_0}v_y, \qquad \frac{\mathrm{d}v_y}{\mathrm{d}t} = -\frac{qB_0}{\gamma m_0}v_x. \tag{2.2}$$

Taking the time derivative of Eq. (2.2) and combining the results with the original expressions yields simple harmonic oscillator equations,

$$\frac{\mathrm{d}^2 v_x}{\mathrm{d}t^2} + \omega_\mathrm{c}^2 v_x = 0, \qquad \frac{\mathrm{d}^2 v_y}{\mathrm{d}t^2} + \omega_\mathrm{c}^2 v_y = 0, \tag{2.3}$$

where we have substituted in the relativistic cyclotron frequency,

$$\omega_\mathrm{c} \equiv \frac{qE_0}{\gamma m_0}. \tag{2.4}$$

The relativistically correct transverse velocities are, from Eq. (2.3), obviously harmonic functions of time, having angular frequency ω_c. This definition of the cyclotron frequency differs from its non-relativistic analogue by the presence of the factor of γ in the denominator. This factor of γ is due to the relativistic change to the inertial mass (cf. Eq. 1.50)—the particle appears to be heavier, and the "fictitious force" associated with centripetal acceleration becomes larger, $mv^2/R \rightarrow \gamma mv^2/R$. This point is explored further in the exercises.

Equations (2.2) and (2.3) taken together indicate harmonic oscillations of equal amplitude in x and y, which are 90° apart, having the general solution

$$v_x = -v_\mathrm{m} \sin(\omega_\mathrm{c} t + \phi), \qquad v_y = v_\mathrm{m} \cos(\omega_\mathrm{c} t + \phi). \tag{2.5}$$

Equation (2.5) can be integrated to find the transverse motion in its most general form:

$$x = R\cos(\omega_\mathrm{c} t + \phi) + x_0, \qquad y = R\sin(\omega_\mathrm{c} t + \phi) + y_0. \tag{2.6}$$

One can easily deduce from Eq. (2.6) that the particle orbit is a circle of radius $R = v_\mathrm{m}/\omega_\mathrm{c}$ centered at (x_0, y_0).

We have not yet established the dependence of R on other physical parameters of the system. To do this, we note that balancing of radial force and centripetal acceleration implies

$$\frac{\gamma m v_\perp^2}{R} = q v_\perp B_0, \quad \text{or} \quad p_\perp = qB_0 R. \tag{2.7}$$

This relationship between the transverse momentum component and the magnetic field is often written in a form which is useful for easy calculations,

$$p_\perp(\mathrm{MeV/c}) = 299.8 \cdot B_0(\mathrm{T}) R(\mathrm{m}). \tag{2.8}$$

Equation (2.8) recasts the second of Eq. (2.7) in the "engineering" units of high-energy accelerators.

In this discussion we have described the relativistically correct helical motion of a charged particle in a uniform magnetic field, which is the spring-board for examining two situations of present interest: the circular accelerator, and the focusing solenoid.

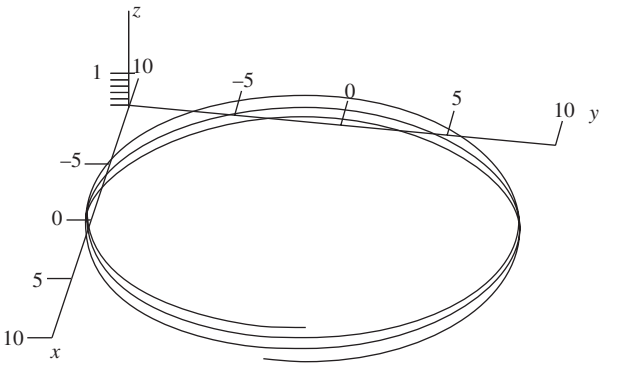


Fig. 2.1 Helical orbit with small pitch angle (0.1) in uniform magnetic field.

2.2 Circular accelerator

The results of the previous section can be applied to discussion of the circular accelerator if the pitch angle ($\theta_p = \tan^{-1}(p_z/p_\perp)$) associated with the motion is small, as is shown in Fig. 2.1. In fact, for simplicity of analysis, we must begin the discussion by assuming that the pitch angle vanishes, thus allowing us to initially ignore the out of bend plane motion. To see why this is so, we reintroduce in this context the notion of the circular design orbit, the trajectory in the accelerator that the ideal particle—the one tracing the exact trajectory that the designer desires—follows. Because this orbit must be closed in the circular accelerator, we are clearly restricted to in-plane motion, with $p_z = 0$. In the example of Fig. 2.1, we may think of the design orbit as being in the $z = 0$ plane (the bend plane), with radius $R = 5$ centered on the point $x = 0, y = 0$.

The motion of charged particle trajectories near the design orbit may be stable (tending to remain near the design orbit) or unstable (tending to diverge from the design orbit). In some circular accelerators, it is necessary that the charged particles be stored in the accelerator near the design orbit for more than 10^{10} revolutions, or turns, so it is obviously of paramount importance that motion near the design orbit be stable.

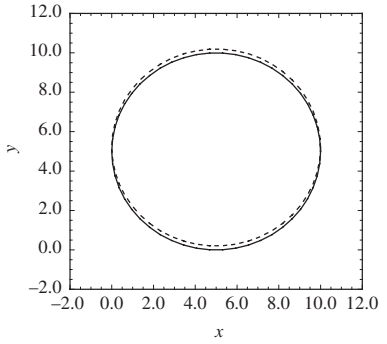


Fig. 2.2 Design (solid line) and transversely perturbed (dashed line) orbits in uniform magnetic field.

It is obvious from Fig. 2.1 that the vertical motion (along the z-axis, out of the bend plane) is not at all stable—the motion in this system is unbounded in the z-dimension. We shall return to this point in Section 3.1 when we discuss motion in a nonuniform magnetic field, and resolve the problem of unbounded motion in the dimension normal to the plane containing the design orbit.

Concentrating now on the motion in the bend plane, there are two possible ways in which this motion can be perturbed from the design orbit—first, by an error in center of curvature of the orbit, and second, by an error in the value of the radius of curvature, which is, of course, equivalent to a deviation in particle momentum from that of the ideal (design) particle. The first type of perturbation is illustrated in Fig. 2.2, which shows two orbits, with the second (perturbed) orbit having the same radius, but with center offset from the design orbit center by a small amount in y. It can be clearly seen that this perturbed orbit displays stability, as it passes through the design orbit twice per turn around the circle. Two different classes of perturbations, both leading to an error in the center of curvature, are displayed in Fig. 2.2: pure angle errors at $x = 0$, 10 and $y = 5$; pure offset errors at $x = 5$ and $y = 0$, 10. The errors at all other points

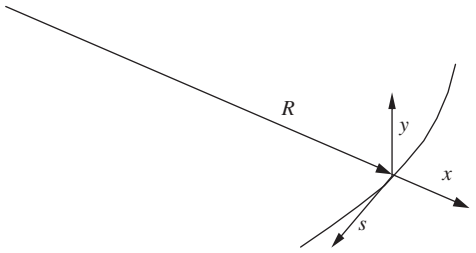


Fig. 2.3 Locally defined right-handed coordinate system used with bending design. Here s is the distance along the design orbit, x is the distance from this orbit along the radius of curvature, and y is the distance from the design orbit out of the bend plane.

along the trajectory are a superposition of angle and offset perturbations. From these observations, we may strongly suspect that the perturbed trajectory is a harmonic oscillation, but we must first develop the proper analysis tools to verify this suspicion.

As stated in Chapter 1, as we will typically take the distance along the design orbit to be the independent variable (in this case indicated by s), we implicitly wish to analyze the charged particle dynamics near the design orbit. In the present case, this orbit is specified by a certain radius of curvature R (and thus a certain momentum $p_0 = qB_0R$), and center of curvature, (x_0, y_0). With this choice of analysis geometry, we can locally define a new right-handed coordinate system (x, y, s), as shown in Fig. 2.3. In this coordinate system, x is the distance of the orbit under consideration from the design orbit, in the direction measured along the radius and normal to s. The distance y (formerly indicated by the coordinate z in Section 2.1) is measured from the design orbit to the particle orbit under consideration, in the direction out of the bend plane.[1] The choice of a right-handed system in this case is a function of the direction of the bend, and in simple circular accelerators, one is free to construct the curvilinear coordinates once and for all. On the other hand, when we encounter bends in the opposing direction, as in chicane systems (see Chapter 3), we will choose to consistently define the coordinate x, so that it is positive along the direction away from the origin of the bend. As we will also choose to leave the vertical direction unchanged in this transformation, a left-handed coordinate system will result when the bend direction is changed.

[1] This convention, in which the symbol y is defined to be the distance out of the bend plane is typical of the American literature. European beam physicists more often use the symbol z instead, but we do not follow this convention even though it connects more naturally to our previous discussion. This is because our adopted convention makes subsequent derivations somewhat easier to understand, and also because it allows the connection between linear accelerator and circular accelerator coordinate systems to become more obvious.

The coordinate system shown in Fig. 2.3 is quite similar to a cylindrical coordinate system, with x related to the radial variable ρ by the definition $x \equiv \rho - R$, s replacing the azimuthal angle ϕ ($\mathrm{d}s = R\,\mathrm{d}\phi$), and y, as previously noted, replacing z. Thus we can write the equations of motion for orbits in this system by using the Lagrange–Euler formulation (see Problem 2.1), as

$$\frac{\mathrm{d}p_\rho}{\mathrm{d}t} = \frac{\gamma m_0 v_\phi^2}{\rho} - q v_\phi B_0, \tag{2.9}$$

where $v_\phi = \rho\dot{\phi}$ is the azimuthal velocity.

Equation (2.9) can be cast as a familiar differential equation by using x as a small variable ($x \ll R$, which is also equivalent, as will be seen below, to the paraxial ray approximation) to linearize the relation. This is accomplished through use of a lowest order Taylor series expansion of the motion about the design orbit equilibrium ($p_x = p_\rho = 0$) at $\rho = R$,

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} \cong -\frac{\gamma_0 m_0 v_0^2}{R^2} x. \tag{2.10}$$

Using Eq. (2.7), we write the design radius as $R = \gamma_0 m_0 v_0 / qB_0$, to obtain

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \omega_\mathrm{c}^2 x = 0, \tag{2.11}$$

where $p_0 = \gamma_0 m_0 v_0$ is the design momentum. Equation (2.11) indicates simple harmonic motion with the same frequency as that of the cyclotron motion describing the design orbit.

Equation (2.11) is written in more standard form by using s as the independent variable, with shorthand designation of differentiation $()' \equiv \mathrm{d}/\mathrm{d}s = (1/v_0)\,\mathrm{d}/\mathrm{d}t$, as

$$x'' + \left(\frac{1}{R}\right)^2 x = 0. \tag{2.12}$$

The simple harmonic oscillations about the design orbit described by Eq. (2.12), associated with the perturbed orbits of particles having the same momentum as the design particle, are termed **betatron oscillations**, because they were first described in the context of the betatron.

These oscillations may seem to be a bit of a mystery at this point, even though the mathematical derivation leading to Eqs (2.11) and (2.12) is straightforward. The question may be asked: if the force and effective mass of the charged particles on and off of the design orbit are the same, how does the "restoring force" arise? This question can be answered most easily by looking at the description of the motion using s as the independent variable, and using the picture shown in Fig. 2.4.

The total momentum transfer of the particle on an arbitrary offset orbit, as shown in Fig. 2.4, is calculated as follows:

$$\Delta p_x = -q \int_{t_1}^{t_2} v_0 B_0 \,\mathrm{d}t = -q \int_{s_1}^{s_2} B_0 \left(1 + \frac{x}{R}\right) \mathrm{d}s. \tag{2.13}$$

The momentum transfer for offset orbits is different in this case because the integration path length is different—integration of the force equation with s as the independent variable is equivalent to using the angle along the design orbit, as parameterized by $\mathrm{d}s = R\,\mathrm{d}\theta$. Integration of the force over the offset path length for a given angular increment covers a larger differential length $\mathrm{d}s_x = (R + x)\mathrm{d}\theta$. For this reason, the focusing effect described by Eqs (2.11) and (2.12) is sometimes termed **path length focusing**. This type of focusing

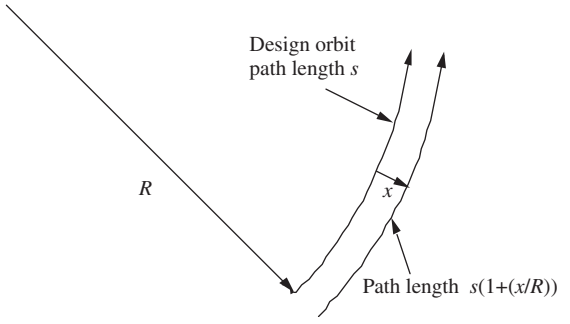


Fig. 2.4 Path length difference between design orbit and offset betatron orbit.

forms the basis of so-called weak focusing systems, which are discussed in the next chapter.

As a final comment in this cursory introduction to circular accelerators, we remind the reader that the betatron motion (which we indicate by the subscripted variable x_β) is that due only to trajectory errors for particles that have the design momentum. The displacement of an arbitrary particle from the design orbit has another important component, that due to deviations from the design momentum $\delta p \equiv p - p_0$. An analysis which treats the particle dynamics only in a first-order Taylor series in both betatron (angle and offset) and momentum errors (which requires both $x_\beta \ll R$ and $\delta p/p_0 \ll 1$) is by assumption a description which is additive in these quantities, that is,

$$x = x_\beta + \eta_x \frac{\delta p}{p_0}. \tag{2.14}$$

The coefficient η_x is termed the momentum dispersion (in this case in the x-direction, or horizontal momentum dispersion), and is, as we shall see, generally a detailed function of the magnetic field profile, with variation in s. In the case of the uniform magnetic field that we have been studying, however, it is a constant, as can be seen by the schematic shown in Fig. 2.5. Because the radius of curvature of each momentum component is linear in p, the momentum dispersion function in this case is constant,

$$\eta_x = \frac{\partial x}{\partial \left[\frac{\delta p}{p_0}\right]} = p_0 \frac{\partial R}{\partial p} = R(p_0). \tag{2.15}$$

The momentum dispersion about the design orbit is simply the radius of curvature of the design orbit (at the design momentum) in the case of a uniform magnetic field.

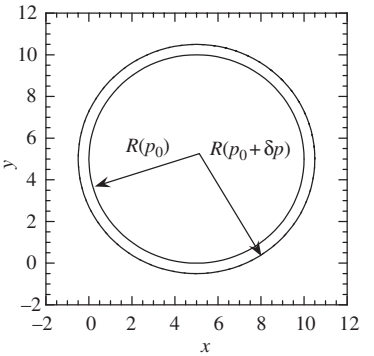


Fig. 2.5 Momentum dispersion in a uniform magnetic field.

Even though this section is devoted to discussion of charged particle motion under the influence of static magnetic fields, because we have discussed circular design orbits we refer the reader to an exercise at this point which illustrates acceleration in a time-varying magnetic field in the betatron, Exercise 2.2.

2.3 Focusing in solenoids

The motion of a charged particle in a focusing solenoid magnet is conceptually the complement of that in the circular accelerator. This can be seen by noting that in the solenoid, the design orbit is that which travels straight down the longitudinal axis of the device (the z-direction), the direction parallel to the uniform magnetic field. Thus only off-axis orbits with non-zero angular momentum—defined about the z-axis—are deflected by this solenoid field. The assumption of paraxial orbits means that $p_z \gg p_\perp$, and the pitch angle of the helical orbit is very large, as illustrated by Fig. 2.6. The key to understanding the motion of a charged particle in a focusing solenoid is to recognize how the angular momentum, which drives this helical motion, arises. To do this we must violate the assumptions of the previous two sections slightly, and ask what happens when the charged particle moves from a region where the magnetic field vanishes to one where it is uniform.

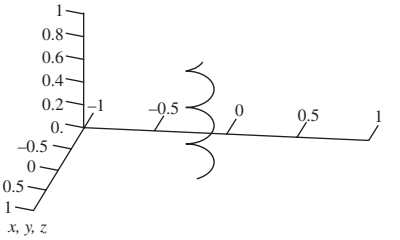


Fig. 2.6 Helical orbit typical of a solenoid, where the pitch angle is large.

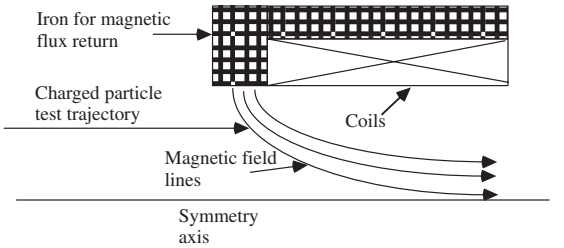


Fig. 2.7 Fringe field region of solenoid, with initially offset particle trajectory encountering radial component of magnetic field.

In this transitional region, the magnetic field must "fringe" to satisfy the divergence-free criterion, $\vec{\nabla} \cdot \vec{B} = 0$. Assuming a cylindrically symmetric geometry, this criterion becomes

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\rho B_\rho = -\frac{\partial B_z}{\partial z}. \tag{2.16}$$

Thus, as the charged particle enters the region where the solenoid field B_z rises and $\partial B_z/\partial z$ is non-vanishing, a radial component of the magnetic field is encountered. This is illustrated in Fig. 2.7, in which the fringe-field region of a solenoid is schematically shown.

In order to integrate Eq. (2.16), we assume the lowest order approximation on the form of the longitudinal field, that it is independent of radius. Because of symmetry, this approximation is good to second order in ρ, that is $B_z(\rho, z) \cong B(0, z)[1 + \alpha\rho^2 + \cdots]$, where α is a constant. Assuming in the region of interest that $|\alpha\rho^2| \ll 1$, we may write

$$B_\rho \cong -\frac{\rho}{2}\left.\frac{\partial B_z}{\partial z}\right|_{\rho=0}. \tag{2.17}$$

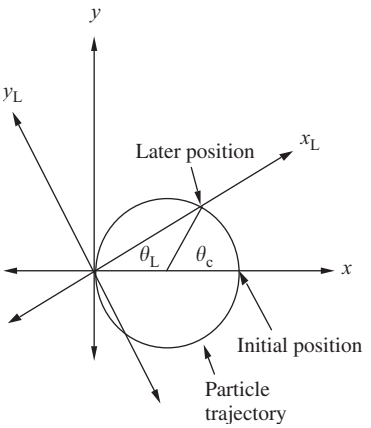


Fig. 2.8 Transverse trajectory of charged particle with no initial transverse momentum (before magnet) in a solenoid. The rotating Larmor frame is defined by the x_L-axis passing through the charged particle that begins its motion on the x-axis.

Further, using the paraxial approximation, we may obtain an excellent estimate of the total angular momentum imparted to a charged particle as it passes through the fringe field region of a solenoid by evaluating the force integral assuming a constant radial offset ρ_0,

$$\begin{aligned}\Delta p_\phi &\cong q\int_{t_1}^{t_2} v_z B_\rho \,\mathrm{d}t = q\int_{z_1}^{z_2} B_\rho \,\mathrm{d}z = -q\frac{\rho_0}{2}\int_{z_1}^{z_2}\left.\frac{\partial B_z}{\partial z}\right|_{\rho=0}\mathrm{d}z \\ &= -q\frac{\rho_0}{2}\int_{z_1}^{z_2}\left.\frac{\mathrm{d}B_z}{\mathrm{d}z}\right|_{\rho=0}\mathrm{d}z = -q\frac{\rho_0}{2}[B_z(z_2) - B_z(z_1)] = -q\frac{\rho_0}{2}B_0.\end{aligned} \tag{2.18}$$

Here we have explicitly assumed that the particle velocity does not change significantly in direction or magnitude while passing through the magnetostatic fringe field region enclosed in the interval (z_1, z_2). Equation (2.18) is sometimes known as Busch's theorem.

It is important to note that the total transverse momentum[2] "kick" imparted to a particle entering the solenoid, $\Delta p_\phi \cong q\rho_0 B_0/2$, gives rise to subsequent transverse motion with radius of curvature $R = \Delta p_\phi/qB_0 \cong \rho_0/2$. Therefore, a charged particle with no initial transverse motion displays helical motion inside of the solenoid, with radius of curvature such that the particle orbit passes through the axis ($\rho = 0$). This somewhat surprising result is illustrated in Fig. 2.8, where both the particle trajectory in this case, and the proper coordinate system for further analysis of the problem, are shown.

[2]Note that here we are examining the transverse momentum "kick" in the angular direction. This should not be confused with the angular momentum, which has a specific meaning in the context of Hamiltonian dynamics—it is the momentum which is canonical with the azimuthal coordinate ϕ (see Ex. 2.4), and of course has units of momentum times length.

If the particle begins its trajectory offset in x ($x = x_0$), but not in y, and with no transverse momentum before the magnetic field region, the angle that this particle's trajectory makes with respect to the x-axis is the Larmor angle θ_L. As can be seen from Fig. 2.8, the Larmor angle is simply related to the cyclotron angle of the motion (centered on $x = x_0/2$) by $2\theta_L - \theta_c = 0$, or $\theta_L = \theta_c/2$. Since the cyclotron angle grows linearly in time, so does the Larmor angle, and thus we can define the Larmor frequency,

$$\omega_L \equiv \frac{d\theta_L}{dt} = \frac{\omega_c}{2} = \frac{qB_0}{2\gamma m_0}. \tag{2.19}$$

The Larmor frame can now be introduced—it is simply the frame that rotates about the z-axis with the Larmor frequency, as illustrated in Fig. 2.8.

In the Larmor frame, the motion in x_L is simple harmonic with the Larmor frequency,

$$x_L = x_0 \cos(\theta_L) = x_0 \cos(\omega_L t). \tag{2.20}$$

The central advantage of introducing the Larmor frame is that the motion is simple harmonic for both of the transverse dimensions in this frame. To see this, we note that the coordinates in the Larmor frame can be written in compact form as

$$\begin{pmatrix} x_L \\ y_L \end{pmatrix} = \begin{pmatrix} \cos(\omega_L t) & \sin(\omega_L t) \\ -\sin(\omega_L t) & \cos(\omega_L t) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{2.21}$$

Thus, we have, using Eqs (2.2) and (2.21),

$$\ddot{x}_L + \omega_L^2 x_L = 0 \quad \text{and} \quad \ddot{y}_L + \omega_L^2 y_L = 0. \tag{2.22}$$

These equations describe simple harmonic oscillations in both x_L and y_L as observed in the Larmor frame. Thus, the motion in a solenoid can be described as a rotation about the origin that proceeds at the Larmor frequency, with simple harmonic (betatron) oscillations occurring in the rotating frame, also proceeding with angular frequency ω_L.

The last point to be made in discussing solenoid focusing concerns casting of the problem in standard accelerator physics form, with the distance z down the axis of the solenoid used as the independent variable. In the paraxial approximation, we take $v_z \cong v$, and thus

$$x_L'' + k_L^2 x_L = 0, \tag{2.23}$$

$$y_L'' + k_L^2 y_L = 0, \quad \text{with } k_L = \frac{\omega_L}{v_z} = \frac{qB_0}{2p_z} \cong \frac{qB_0}{2p}. \tag{2.24}$$

The inverse of the wave number associated with the betatron harmonic motion in the Larmor frame is approximately twice the radius of curvature $R = p/qB_0$ of a charged particle of the same momentum traveling normal to the solenoid field.

2.4 Motion in a uniform electric field

The substitution of a uniform electric field $\vec{E} = E_0\hat{z}$ for the uniform magnetic field changes our point of view quite dramatically, as the energy of the particle is

no longer constant, but increases due to longitudinal acceleration. The equations of motion in this case, with $\vec{B} = 0$, read

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = qE_0, \qquad \frac{\mathrm{d}\vec{p}_\perp}{\mathrm{d}t} = 0. \tag{2.25}$$

Since the momentum transverse to the electric field is conserved we may, for the moment, ignore its effects and concentrate on the one-dimensional problem of the motion in z.

The most straightforward way to accomplish this is to note that the electrostatic field $\vec{E} = E_0\hat{z}$ may be derived from a potential, and thus the potential energy is

$$q\phi_\mathrm{e} = -qE_0 z. \tag{2.26}$$

The Hamiltonian, which in this case is the total energy (see Section 1.4), is

$$H = \gamma m_0 c^2 + q\phi_\mathrm{e} = \gamma m_0 c^2 - qE_0 z. \tag{2.27}$$

Because the Hamiltonian is independent of t, it is a constant of the motion, and it can be evaluated at given initial conditions, $H|_{z=0}$. The particle's mechanical (rest plus kinetic) energy is therefore given by

$$\gamma m_0 c^2 = H|_{z=0} + qE_0 z, \tag{2.28}$$

or in normalized form,

$$\gamma(z) = \frac{H|_{z=0}}{m_0 c^2} + \gamma' z, \tag{2.29}$$

where $\gamma' \equiv qE_0/m_0c^2$, and the initial conditions-derived constant $H|_{z=0} = \gamma|_{z=0} m_0 c^2$. Note that the energy increases linearly in z, as $\mathrm{d}U = \vec{v} \cdot \mathrm{d}\vec{p}$, or for one-dimensional motion

$$\frac{\mathrm{d}U}{\mathrm{d}z} = \frac{\mathrm{d}p}{\mathrm{d}t} = qE_0. \tag{2.30}$$

This simple relation helps clarify some jargon encountered in the particle beam physics field, that the amplitude of the accelerating force (change in momentum per unit time) is referred to as an acceleration gradient, or spatial energy gradient, commonly quoted in units of MeV/m. The normalized acceleration gradient γ' defines a scale length for acceleration $L_\mathrm{acc} = \gamma'^{-1}$, over which the particle gains one unit of rest energy.

In this simple scenario, other relevant dynamical quantities can be derived from knowledge of $\gamma(z)$,

$$p(z) = \beta\gamma m_0 c = \sqrt{\gamma^2(z) - 1}\, m_0 c, \tag{2.31}$$

and

$$v(z) = \beta c = \frac{p(z)c^2}{U(z)} = c\sqrt{1 - \frac{1}{\gamma^2(z)}}. \tag{2.32}$$

The deduction of the velocity from γ is of prime importance because it allows the transformation of independent variable from t, which enters naturally into the equations of motion, to z.

In the analysis given in this section, we have effectively adopted z as the independent variable. It is also interesting to explore acceleration from the point of view of explicit time dependence. In this case, we can write the first of Eq. (2.25) as

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = m_0 c \frac{\mathrm{d}(\beta_z \gamma)}{\mathrm{d}t} = m_0 c[\gamma + \beta_z^2 \gamma^3] \frac{\mathrm{d}\beta_z}{\mathrm{d}t} \cong \gamma^3 m_0 c \frac{\mathrm{d}\beta_z}{\mathrm{d}t} = qE_0, \quad (2.33)$$

or

$$\frac{\mathrm{d}v_z}{\mathrm{d}t} = \frac{qE_0}{\gamma^3 m_0}, \quad (2.34)$$

where we have assumed in Eq. (2.33) that $\beta_z \cong \beta$, the motion is predominantly along the z direction. Equation (2.34) seems to indicate that the effective mass of a relativistic particle in accelerating parallel to its velocity vector is $\gamma^3 m_0$. This longitudinal mass effect will be revisited when we discuss charged particles' longitudinal oscillations in accelerators.

Let us now examine a phase plane plot of uniform acceleration, as illustrated in Fig. 2.9. The curves of constant H are hyperbolae, which look locally parabolic near $p_z = 0$, the region of the non-relativistic limit. Asymptotically in large $p_z/m_0 c$, the ultra-relativistic region, the curves approach straight lines.

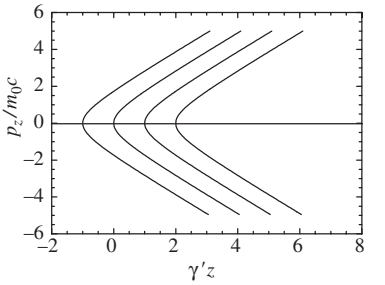


Fig. 2.9 Longitudinal phase plane $(\gamma' z, p_z/m_0 c)$ trajectories for uniform acceleration.

In Section 2.3, we found that edge effects—the effects of moving from a field-free region to one with field present—were very important in understanding the transverse motion of a charged particle in a longitudinal magnetic field. This is also true for the case of entry into a uniform electric field. The analysis proceeds just as in Eqs (2.16)–(2.18), and begins with an expansion near the axis of Eq. (1.17),

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\rho E_\rho = -\frac{\partial E_z}{\partial z}. \quad (2.35)$$

Integration of Eq. (2.35) outward from the axis yields

$$E_\rho \cong -\frac{\rho}{2}\left.\frac{\partial E_z}{\partial z}\right|_{\rho=0}, \quad (2.36)$$

which can be further used to find the radial momentum kick imparted to the particle as in passes through the fringing field region where E_ρ is non-zero,

$$\begin{aligned}\Delta p_\rho &= q\int_{t_1}^{t_2} E_\rho \,\mathrm{d}t = \frac{q}{v}\int_{z_1}^{z_2} E_\rho \,\mathrm{d}z = -\frac{\rho q}{2v}\int_{z_1}^{z_2}\left.\frac{\partial E_z}{\partial z}\right|_{\rho=0}\mathrm{d}z \\ &= -\frac{\rho q}{2v}\int_{z_1}^{z_2}\left.\frac{\mathrm{d}E_z}{\mathrm{d}z}\right|_{\rho=0}\mathrm{d}z = -\frac{\rho q}{2v}[E_z(z_2) - E_z(z_1)] = -\frac{\rho q}{2v}E_0. \quad (2.37)\end{aligned}$$

We have assumed in Eq. (2.37) that the particle moves from a field-free region to one with uniform longitudinal electric field E_0. For an accelerating field ($qE_0 > 0$), this effect is obviously focusing, with the momentum kick tending to push the particle towards the axis. When a particle leaves the uniform accelerating field region, however, the lines of force fringe outward instead of inward and correct evaluation of Eq. (2.37) in this case yields a defocusing momentum kick.

2.5 Motion in quadrupole electric and magnetic fields

In our discussion of beam optics thus far we have seen that one can use solenoidal magnetic fields for focusing particles during linear transport, and path length focusing in circular accelerators. These forms of focusing have introduced some basic concepts, such as harmonic betatron oscillations about the design orbit, but do not include the most widely used magnetic focusing scheme, that which employs transverse quadrupole fields.

Since we are using a coordinate system with one preferred longitudinal axis (z), we use cylindrical coordinates (ρ, ϕ, z) to evaluate the form of the scalar potentials from which either static transverse electric or magnetic[3] fields can be derived. Let ψ be such a potential, which, in the limit of a device long compared to its transverse dimensions, approximately obeys the two-dimensional Laplace equation

$$\vec{\nabla}_\perp^2 \psi = 0. \tag{2.38}$$

The solutions to this equation are of a form that is well behaved on axis ($\rho = 0$)

$$\psi = \sum_{n=1}^{\infty} a_n \rho^n \cos(n\phi) + b_n \rho^n \sin(n\phi), \tag{2.39}$$

Let us discuss the first few of the n multipole forms of the solution. For $n = 1$, we have

$$\psi_1 = a_1 \rho \cos(\phi) + b_1 \rho \sin(\phi) = a_1 x + b_1 y \tag{2.40}$$

and the fields are of the dipole form. Let us now assume we are talking about magnetic fields, so $\vec{B} = -\vec{\nabla}\psi$, and

$$\vec{B}_1 = -\vec{\nabla}\psi_1 = -a_1 \hat{x} - b_1 \hat{y}. \tag{2.41}$$

The form given in Eq. (2.41) is referred to as a dipole field because it can be formed with a magnet possessing only two poles. Such magnets will be discussed further in Chapter 6. They are the devices that yield the uniform fields that we have analyzed in Sections 2.1–2.3.

For $n = 2$, we have

$$\psi_2 = a_2 \rho^2 \cos(2\phi) + b_2 \rho^2 \sin(2\phi) = a_2(x^2 - y^2) + 2b_2 xy \tag{2.42}$$

and the associated magnetic field is

$$\vec{B}_2 = 2a_2(-x\hat{x} + y\hat{y}) - 2b_2(y\hat{x} + x\hat{y}), \tag{2.43}$$

If the potential coefficient a_2 is non-vanishing, there is a force on a charged particle traveling in the z direction, directed in the x dimension and which is proportional to y, and vice versa. This type of force is called skew quadrupole, and gives rise to generally undesirable coupling between the x and y phase planes. The coefficient b_2 indicates the presence of normal quadrupole fields, from which one obtains a force in the x dimension proportional to x, and a force in y that is proportional to y,

$$\vec{F}_\perp = q v_z \vec{z} \times \vec{B}_2 = 2q v_z b_2 (y\hat{y} - x\hat{x}). \tag{2.44}$$

Assuming qb_2 is positive, this force is focusing in the x dimension and defocusing in y. Obviously if qb_2 is negative, these focusing/defocusing roles are

[3] If there is no current density $\vec{J} = 0$ in the region of interest, the magnetic field can be derived from a scalar potential. The case where the current density is non-vanishing is treated in Chapter 6.

reversed. Since the coefficient b_2 measures the gradient in magnetic field away from the axis, one often gives its strength in terms of this gradient $b_2 = -\partial_x B_x/\partial y \equiv B'/2$. The term "quadrupole" can be understood by looking at Fig. 2.10, which shows a design for a quadrupole built at UCLA, with its four poles excited by current-carrying coils in alternating polarity. The design principles of this magnet are discussed in Chapter 6. Let us content ourselves at the present to point out that the iron poles, whose surfaces form magnetic equipotentials, are hyperbolae, as suggested by Eq. (2.42).

If we assume paraxial motion near the z-axis, we may write the transverse equations of motion for a particle of charge q and momentum p_0 as

$$x'' = \frac{F_x}{\gamma m_0 v_0^2} = -\frac{qB'}{p_0}x, \tag{2.45}$$

and

$$y' = \frac{F_y}{\gamma m_0 v_0^2} = \frac{qB'}{p_0}y. \tag{2.46}$$

These equations are written in standard oscillator form as

$$x'' + \kappa_0^2 x = 0 \tag{2.47}$$

and

$$y'' - \kappa_0^2 y = 0. \tag{2.48}$$

The square wave number $\kappa_0^2 \equiv qB'/p_0$ is sometimes known as the focusing strength K. It can be simply calculated by using the handy shortcut $qB'/p_0 = B'/BR$ where $BR(\mathrm{Tm}) = p_0(\mathrm{MeV}/c)/299.8$ (cf. Eq. 2.8) is known as the magnetic rigidity of the particle.

Assuming $\kappa_0^2 > 0$, one has simple harmonic oscillations in x,

$$x = x_0 \cos[\kappa_0(z - z_0)] + \frac{x_0'}{\kappa_0}\sin[\kappa_0(z - z_0)], \tag{2.49}$$

and the motion in y is hyperbolic,

$$y = y_0 \cosh[\kappa_0(z - z_0)] + \frac{x_0'}{\kappa_0}\sinh[\kappa_0(z - z_0)]. \tag{2.50}$$

If $\kappa_0^2 < 0$, the motion is simple harmonic (oscillatory) in y, and hyperbolic (unbounded) in x. Focusing with quadrupoles alone can only be accomplished in one transverse direction at a time. Ways of circumventing this apparent limitation in achieving transverse stability, by use of alternating gradient focusing, are discussed in the Chapter 3.

As noted in Chapter 1, it is much easier to build transverse field magnets than transverse electric field-supporting electrode arrays that impart equal force upon charged particles traveling nearly the speed of light. Therefore, the transverse electric field quadrupole is found mainly in very low energy applications such as the electron microscope. Note that an electric quadrupole that produces decoupled forces in analogy with the magnetic forces of Eq. (2.41) has an electrode array with hyperbolic surfaces rotated by 45° from that shown in Fig. 2.10. In this case we may say that from the viewpoint of electric forces, if

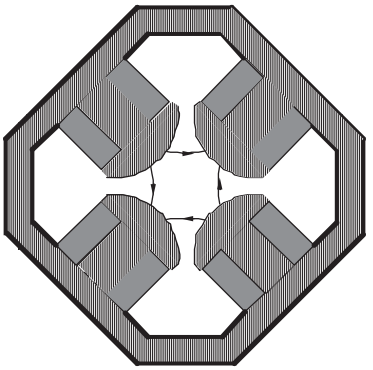


Fig. 2.10 Quadrupole magnet showing iron pole tips and yoke (hatching), current windings (gray), and orientation of magnetic field lines.

the same form of the solution for the electrostatic potential ϕ_e is used as was taken for ψ (Eq. 2.3), then a_2 is the coefficient of normal electric quadrupole, and b_2 is the coefficient of the skewed electric quadrupole.

The paraxial ray equations in the case of an electric quadrupole are identical to those given in Eqs (2.47) and (2.48), with $\kappa_0^2 \equiv qE'/p_0v_0$. When the motion is focusing in x it is defocusing in y, as before. The differences between the two cases lie in the deviations from the assumptions of the paraxial approximation. In the case of magnetic forces, the energy and total momentum are constant, but the longitudinal momentum p_z, and thus the longitudinal velocity v_z may change, which in turn forces a small modification of Eqs (2.47) and (2.48). With electric forces the situation is slightly different—the energy is not constant, as work may be performed on a particle with transverse motion, but the longitudinal momentum p_z is now constant. In this case the longitudinal velocity v_z may also change non-negligibly when the paraxial approximation is not valid.

2.6 Motion in parallel, uniform electric and magnetic fields

In the case of a uniform electric field, $\vec{E} = E_0\hat{z}$, and a parallel uniform magnetic field, $\vec{B} = B_0\hat{z}$, the electric field provides uniform acceleration while the magnetic field provides solenoidal focusing. This example scenario shows the combination of acceleration and focusing well, nicely illustrating the phenomenon of adiabatic damping of transverse oscillations. The equations of motion in this field configuration are

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = qE_0 \qquad \frac{\mathrm{d}\vec{p}_\perp}{\mathrm{d}t} = q(\vec{v}_\perp \times \vec{B}) = \frac{qB_0}{\gamma m_0}(\vec{p}_\perp \times \hat{z}). \tag{2.51}$$

Equation (2.51) are coupled by the presence of γ in the second equation, and so the method of solution offered previously for the transverse motion must be re-examined.

On the other hand, as the amplitude of $\vec{p}_\perp$ is invariant, the energy may still be trivially found as a function of z,

$$U(z) = \gamma(z)mc^2 = \sqrt{(p_\perp^2 + p_z^2|_{z=0})c^2 + (m_0c^2)^2} + qE_0z. \tag{2.52}$$

In the paraxial limit, one may ignore the transverse momentum contribution to the energy, and Eq. (2.52) gives

$$U(z) = \gamma(z)mc^2 = \sqrt{p_z^2|_{z=0}c^2 + (m_0c^2)^2} + qE_0z. \tag{2.53}$$

After the energy and the longitudinal momentum are determined by Eqs (2.53) and (2.31), one can begin to write the transverse equations of motion in the Larmor frame (rotating with local frequency $\omega_L(z) = qB_0/2\gamma(z)m_0$) through the definition

$$x_L' \equiv \frac{p_{x_L}}{p_z}, \tag{2.54}$$

with an analogous expression in y_L. Since p_{x_L} is a Cartesian projection of the constant amplitude $\vec{p}_\perp$, we expect the maximum angle in x_L (and y_L) to be

secularly—meaning on a time scale longer than the relevant (Larmor) oscillation period—damped by the acceleration, as $x'_L \propto p_z^{-1}$.

Concentrating on the motion in x_L, we differentiate Eq. (2.54) to obtain

$$x''_\mathrm{L} + \frac{p_{x_\mathrm{L}} p'_z}{p_z^2} - \frac{p'_{x_\mathrm{L}}}{p_z} = 0, \tag{2.55}$$

or using the paraxial approximation $p_z \cong p$,

$$x''_\mathrm{L} + \frac{(\beta\gamma)'}{\beta\gamma} x'_\mathrm{L} + \left(\frac{qB_0}{2\beta\gamma m_0 c}\right)^2 x_\mathrm{L} = 0. \tag{2.56}$$

If we look only at highly relativistic motion, $\beta \cong 1$, we can further approximate Eq. (2.56),

$$x''_\mathrm{L} + \frac{\gamma'}{\gamma} x'_L + \left(\frac{b\gamma'}{2\gamma}\right)^2 x_\mathrm{L} = 0. \tag{2.57}$$

The solution to this homogeneous equation is of the form

$$x_\mathrm{L}(z) = x_{\mathrm{L},0} \cos\left[\frac{b}{2} \ln\left(\frac{\gamma(z)}{\gamma_0}\right)\right] + \frac{2\gamma_0}{b\gamma'} x'_{\mathrm{L},0} \sin\left[\frac{b}{2} \ln\left(\frac{\gamma(z)}{\gamma_0}\right)\right], \tag{2.58}$$

where $b = B_0 c/E_0$ and the initial offset, angle, and Lorentz factor are $x_{\mathrm{L},0}, x'_{\mathrm{L},0}$, and γ_0, respectively. Thus the solution appears as a harmonic oscillator with a logarithmically (as opposed to linearly) increasing argument in the oscillatory trigonometric functions. We now compare this system to that of the simple harmonic osciallator.

Simple harmonic motion is associated with two invariants: the angular frequency ω and the value of the Hamiltonian (the total oscillator energy). As discussed in Sections 1.3 and 2.3, these quantities are related through

$$H = \frac{1}{2m}[p_x^2 + m^2\omega^2 x^2] = J_x \omega, \tag{2.59}$$

where the action $J_x = \oint p_x \, \mathrm{d}x/2\pi = A/2\pi = x_\mathrm{max} p_{x,\mathrm{max}}/2$, is the area in the phase plane over 2π, enclosed by the elliptical trajectory of the oscillator (see Fig. 1.4).

An analogue to the action in trace space can be defined as $J_{x,\mathrm{trace}} \equiv x_\mathrm{max} x'_\mathrm{max}/2$. For the system under study we obtain the angle in x_L from Eq. (2.58) as,

$$x'_\mathrm{L}(z) = -x_{\mathrm{L},0} \frac{b\gamma'}{2\gamma} \sin\left[\frac{b}{2} \ln\left(\frac{\gamma(z)}{\gamma_0}\right)\right] + \frac{\gamma_0}{\gamma} x'_{\mathrm{L},0} \cos\left[\frac{b}{2} \ln\left(\frac{\gamma(z)}{\gamma_0}\right)\right]. \tag{2.60}$$

Setting $x'_{\mathrm{L},0} = 0$ for simplicity, one can see that

$$J_\mathrm{trace} \equiv \frac{x_{\mathrm{L,max}} x'_{\mathrm{L,max}}}{2} = x_{\mathrm{L},0}^2 \frac{\gamma'}{\gamma} \frac{b}{2}, \tag{2.61}$$

and the action in trace space damps as γ^{-1}. According to Eq. (2.60), it is clear that this apparent damping is due to the diminishing of x'_L with acceleration. This diminishing should have the functional dependence on momentum of $(\beta\gamma)^{-1}$. We have only approximately obtained this value by assuming the limit $\beta = 1$

for the purposes of solving the paraxial equation of motion, Eq. (2.56), exactly. It should be equally clear that the true phase plane action variable is constant in this case.

As mentioned in Section 1.6, the role of the action in the general theory of oscillators is that they are adiabatic invariants. This means that if the oscillator parameters are changed on a time scale that is slow compared to the oscillation period, then the action is conserved. The true phase plane action is an adiabatic invariant and is conserved, not damped, under slow acceleration. On the other hand, the apparent action in trace space is damped. Because of the guaranteed adiabatic invariance of the phase plane action under slow acceleration and concomitant changing of the oscillation frequency, the damping of the trace space action is termed adiabatic.

Use of the term adiabatic damping to describe this phenomenon is thus actually a bit of a misnomer. It is especially so in the context of the system described in this section, where the transverse motion under acceleration is not required to be an adiabatic process at all. It was only required to be one in which the transverse momentum is constant and the longitudinal momentum grows.

2.7 Motion in crossed uniform electric and magnetic fields*

Let us now consider the case where there are crossed uniform electric and magnetic fields, which we choose as $\vec{E} = E_0\hat{y}$, and $\vec{B} = B_0\hat{z}$. In this case, there is a particular solution to Eq. (1.11) obtained when the Lorentz force vanishes, given by

$$\vec{E} + \vec{v} \times \vec{B} = 0. \tag{2.62}$$

The velocity that is consistent with Eq. (2.62) is termed the drift velocity, and is found by taking the cross-product of Eq. (2.62) with $\vec{B}$, to give

$$\vec{v}_\mathrm{d} = \frac{\vec{E} \times \vec{B}}{\vec{B}^2} = \frac{E_0}{B_0}\hat{x}. \tag{2.63}$$

The drift velocity is normal to both the magnetic and electric field directions.

The general motion in this case can be deduced by transforming the Lorentz frame traveling with the drift velocity using Eq. (1.58) to obtain

$$\begin{aligned} \vec{E}'_\perp &= \gamma_\mathrm{d}(\vec{E}_\perp + \vec{v}_\mathrm{d} \times \vec{B}_\perp) = 0 \\ \vec{B}' &= \gamma_\mathrm{d}(\vec{B}_\perp - \frac{1}{c^2}\vec{v}_\mathrm{d} \times \vec{E}_\perp) = \gamma_\mathrm{d} B_0 \left(1 - \frac{\vec{v}_\mathrm{d}^2}{c^2}\right)\hat{z} = \frac{B_0}{\gamma_\mathrm{d}}\hat{z}. \end{aligned} \tag{2.64}$$

The field is purely magnetic and uniform in the new frame. This means that any motion in this frame is simply that of the helical cyclotron form discussed in Section 2.1. Looking down the axis (z) of the magnetic field, a direction in which there is no net force, the motion appears cycloidal, in other words cyclic motion with a secular velocity (the drift velocity) superimposed, as shown in Figure 2.11.

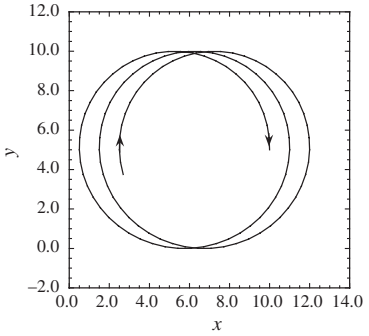


Fig. 2.11 Cycloid motion in crossed electric and magnetic fields: $\vec{E} = E_0\hat{y}, \vec{B} = B_0\hat{z}$.

The drift velocity and associated cycloid motion is not as important of an effect in accelerators as those encountered in previous sections. Such behavior

is notably found in an analysis of intense beam rotation in solenoids due to the longitudinal magnetic field and the radial self-electric field of the beam (see Ex. 2.14). The drift velocity is, however, quite important in plasma physics, and in separation devices—a particle with the correct velocity passes through a crossed electric and magnetic field region, regardless of charge. This is an important technique in experimental physics, and can be used for velocity identification of a particle. The cross-field configuration, in tandem with traversal of a region of uniform magnetic field, in which the momentum of the particle is determined, allows the mass and energy of the particle to also be determined.

2.8 Motion in a periodic magnetic field*

As a final scenario for this chapter, let us consider the case of a spatially periodic (in the z-direction) static magnetic field. There are two possible configurations of this undulator magnetic field, a planar or helical polarization, names that refer to analogous electromagnetic wave polarizations. In this section we will discuss the planar undulator, leaving analysis of the helical configuration for the exercises. Also, we will emphasize here the bending-plane motion of the particle, of the undulator, with the out-of-plane motion examined further in Chapter 3.

The periodic, vertically polarized magnetic field of interest here can be described mathematically by

$$\vec{B} = B_0 \sin(k_u z)\hat{y} \tag{2.65}$$

where $k_u = 2\pi/\lambda_u$ is the wave number associated with the undulator wavelength λ_u of the field. This expression is valid only in the symmetry plane of the field, which we take to be $y = 0$. For the planar configuration, there must be an additional longitudinal field component, as illustrated in Fig. 2.12, and analyzed in the Exercise 2.15.

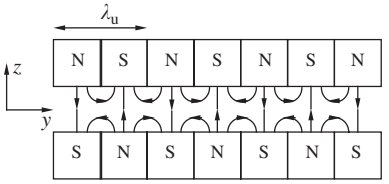


Fig. 2.12 Schematic representation of the magnetic field lines in a planar undulator configuration.

As can be seen by the results of Exercise 2.15, Eq. (2.65) is approximately correct if the vertical offset from the symmetry plane is much smaller than an undulator wavelength, $k_u y \ll 1$. Let us now examine the electron beam dynamics in the undulator field under this additional constraint, which can be considered an extension of the paraxial approximation. We will formulate the problem in Hamiltonian style. To begin, we write the canonical momenta

$$\begin{aligned} p_{c,x} &= \beta_x \gamma m_0 c + qA_x = \beta_x \gamma m_0 c - q\frac{B_0}{k_u}\cos(k_u z)\cosh(k_u y), \\ p_{c,y} &= \beta_y \gamma m_0 c, \\ p_{c,z} &= \beta_z \gamma m_0 c, \end{aligned} \tag{2.66}$$

from which we can derive the relativistically correct Hamiltonian,

$$H = \sqrt{\left(p_{c,x} + q\frac{B_0}{k_u}\cos(k_u z)\cosh(k_u y)\right)^2 c^2 + p_{c,y}^2 c^2 + p_{c,y}^2 c^2 + (m_0 c^2)^2}, \tag{2.67}$$

where the electrostatic potential ϕ_e vanishes. Viewing z as the independent variable, we write the new form of the Hamiltonian function as

$$G = -p_{z,c} = \sqrt{H^2 - (p_{c,y} - qA_y)^2c^2 - (p_{c,x} - qA_x)^2c^2 - (m_0c^2)^2}$$
$$= \sqrt{U^2 - (p_{c,y} - qA_y)^2c^2 - \left(p_{c,x} + q\frac{B_0}{k_u}\cos(k_u z)\cosh(k_u y)\right)^2 c^2 - (m_0c^2)^2}, \tag{2.68}$$

where we have substituted the numerical energy U for the old Hamiltonian functional energy H. This new Hamiltonian is independent of x and t, and thus the canonical x component of the momentum $p_{c,x}$ is a constant of the motion, as is the total mechanical energy U (or equivalently, the total mechanical momentum p_0), which is always the case in magnetostatic systems.

The first integrals of the Hamiltonian system (the momenta) are thus as follows:

$$p_{c,x} = \text{constant} = p_{x0} \Rightarrow \beta_x\,\gamma m_0 c = -qA_x + p_{x0} = q\frac{B_0}{k_u}\cos(k_u z) + p_{x0}, \tag{2.69}$$

$$p_{c,y} = \beta_y\,\gamma m_0 c = 0 \text{ (by assumption, taking } y = 0), \tag{2.70}$$

and

$$p_{c,z} = \sqrt{p_0^2 - \left(q\frac{B_0}{k_u}\cos(k_u z) + p_{x0}\right)^2}. \tag{2.71}$$

This analysis has been done formally, with the aid of the Hamiltonian, in order to make more advanced analysis of the motion in undulators possible in following chapters.

The most important step remaining in the analysis is to find the second integral in x, which is obtained by integrating the paraxial equation

$$x' = -\frac{\partial G}{\partial p_{x,c}} \cong \frac{qB_0}{p_0 k_u}\cos(k_u z) + \frac{p_{x0}}{p_0} = \frac{qB_0}{p_0 k_u}\cos(k_u z) + x_0', \tag{2.72}$$

to give, with initial (evaluated before entry into the undulator field) horizontal offset and angle (x_0, x_0'),

$$x \cong x_0 + x_0' z + \frac{qB_0}{p_0 k_u^2}\sin(k_u z). \tag{2.73}$$

It can be seen from Eq. (2.73) that the amplitude of the undulating portion of the transverse motion is $qB_0/p_0k_u^2$, and from Eq. (2.72) that the maximum angle associated with the undulating part of the motion is qB_0/p_0k_u. This angle is typically much smaller than unity (the bends are paraxial) for magnets refered as undulators. Furthermore, we note that Eq. (2.73) shows that there is no restoring, or focusing force in the x-direction associated with this configuration of magnetic field—an initial error $x_0' \neq 0$ is not corrected, and leads eventually to a trajectory with large horizontal offset x.

The conservation of total momentum combined with Eq. (2.69) gives Eq. (2.71), which indicates that the longitudinal momentum, and therefore the longitudinal velocity, must diminish in the undulator, approximately as

$$p_z = \sqrt{p_0^2 - \left(q\frac{B_0}{k_\mathrm{u}}\cos(k_\mathrm{u}z)\right)^2} \cong p_0\left[1 - \frac{1}{2}\left(\frac{qB_0}{k_\mathrm{u}p_0}\cos(k_\mathrm{u}z)\right)^2\right], \quad (2.74)$$

for paraxial bends. Averaging Eq. (2.74) over a period of the motion, we have simply

$$\langle p_z\rangle \cong p_0\left[1 - \left(\frac{qB_0}{2k_\mathrm{u}p_0}\right)^2\right]. \quad (2.75)$$

This "slowing" of the particle in its z-motion is an important effect in free-electron lasers, which are discussed in Chapter 8.

2.9 Summary and suggested reading

This chapter has been concerned with introducing a number of model problems, based on the relativistic motion of charged particles in static electric and magnetic field configurations. These configurations have included:

1. Uniform (dipole) magnetic fields, in which we deduce many aspects of the motion in both circular accelerators and focusing solenoids from the general case of helical motion. In both scenarios, we found variants of simple-harmonic betatron oscillations.
2. Uniform electric fields, in which acceleration of charged particles to relativistic energies are introduced.
3. Quadrupole magnetic and electric fields, where the motion is simple harmonic in one transverse dimension, and divergent in the other.
4. Superpositions of uniform electric and magnetic fields, which produce damped oscillations when the fields are parallel, and drift motion when the fields are crossed.
5. The periodic magnetic dipole field, or magnetic undulator. This device is shown to produce transverse undulating motion, which forms the basis of the free-electron laser.

These model problems also allowed us to introduce some rudimentary examples of analyses that are based on both relativistic and Hamiltonian formalisms. These analyses will help form the basis of more complex investigations of charged particle motion in the coming chapters.

Many texts in electromagnetism also introduce aspects of charged particle motion in electric and magnetic fields. Readers wishing to review the theory of electrostatic and magnetostatic fields may also wish to review such texts. The following texts may be recommended as a supplement to this chapter:

1. R. Wangsness, *Electromagnetic Fields* (Wiley, 1986). There are many examples of dynamics calculations in this book, including useful appendices.
2. J.R. Reitz, F.J. Milford, and R.W. Christy, *Foundations of Electromagnetic Theory* (Addison-Wesley, 1993). Another fine undergraduate electromagnetism text with examples.

3. J.D. Jackson, *Classical Electrodynamics* (Wiley, 1975). A wealth of assigned problems are relevant to our discussions here.
4. L.D. Landau and E.M. Lifschitz, *The Classical Theory of Fields* (Addison-Wesley, 1971).
5. S. Humphries, Jr., *Principles of Charged Particle Acceleration* (Wiley, 1986). This text also examines many model problems as a path to understanding accelerator behavior.
6. J.D. Lawson, *The Physics of Charged Particle Reams* (Clarendon Press, 1977). There are many clear explanations of basic problems in charged particle motion.
7. F.F. Chen, *Introduction to Plasma Physics and Controlled Fusion: Plasma Physics* (Plenum, 1984). The field of plasma physics has a wide variety of physical effects it addresses, and some differing dynamics problems are addressed in this excellent introduction.
8. J.B. Marion and S.T. Thornton, *Classical Dynamics of Particles and Systems* (Harcourt Brace & Company, 1995).
9. H. Goldstein, C. Poole and J. Safko, *Classical Mechanics* (Addison-Wesley, 2002). Dynamics in electromagnetic fields using modern mechanics at a sophisticated level.

Exercises

(2.1) The mass increase implied by Eq. (2.4) can be obtained by using the relativistically correct Lagrangian (Eq. 1.60) to write the problem in cylindrical coordinates,

$$L(\vec{x},\dot{\vec{x}}) = -\frac{m_0c^2}{\gamma} + q\vec{A}\cdot\dot{\vec{x}}.$$

(a) Show that the cylindrically symmetric vector potential for a uniform magnetic field is $\vec{A} = (B_0\rho/2)\hat{\phi}$.
(b) Write the Lagrangian in cylindrical coordinates (ρ,ϕ,z).
(c) Derive the Lagrange–Euler equations to show that the transverse motion (centered at the z-axis) is circular, with appropriate angular frequency and radius of curvature.

(2.2) In the betatron, the design orbit of the electrons is circular, with constant radius R, but the vertical magnetic field at the design orbit increases in time. This is described by the relation

$$p_0(t) = eB_0(t)R \quad \text{(i)}$$

which indicates that the design momentum must be increased in proportion to the strength of the guiding field. The acceleration for electrons traveling in a strictly circular orbit is given by

$$\frac{\mathrm{d}p_0}{\mathrm{d}t} = -eE_\phi, \quad \text{(ii)}$$

where E_ϕ is the azimuthal electric field, tangential to the electron path.

The betatron condition occurs when these requirements are consistent with each other. In order for this to be true, the field normal to the bend plane (B_z in cylindrical coordinates) must not be constant, but it should be a function of ρ, as shown in a representative way in Fig. 2.13. To derive the betatron condition, employ the following steps:

(a) Differentiate the expression (i) with respect to time, and

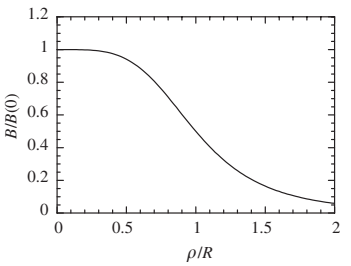


Fig. 2.13 Example of magnetic field profile as a function of radius in a betatron.

(b) Calculate the induced electric field from Stokes theorem applied to

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}.$$

The betatron condition is often given as $B_0(t) = \bar{B}(t)/2$, where $\bar{B} \equiv (\int \vec{B} \cdot d\vec{A})/(\pi R^2)$ is the average normal magnetic field over the disk whose boundary is the design orbit. Is this the most general form of the betatron condition? Note that the betatron condition implies that the field is as shown in Fig. 2.13, having larger amplitude for $\rho < R$. This point is returned to in the next chapter.

(2.3) The cyclotron spoken of in Chapter 1 is schematically shown in Fig. 2.14. It consists of two opposing D-shaped iron magnet pole pairs (so-called Ds), giving rise to a roughly constant magnetic field B_0 normal to the diagram inside of the Ds. This field bends the particles inside of the D-regions in semi-circular trajectories. The Ds also are excited by a time-dependent voltage source of constant angular frequency $\omega_{\rm rf}$ such that $V(t) = V_0 \sin(\omega_{\rm rf} t)$, where rf stands for radio-frequency. Thus, when the particles cross the gap between the Ds they can be accelerated by the electric field between the two different potential regions, so that their energy is maximally incremented by $\Delta U = qV_0$ per crossing. This occurs only when the motion is synchronous with voltage wave-form, requiring $\omega_{\rm c} \cong \omega_{\rm rf}$. Note that the particles have larger radius of curvature as they gain energy and are eventually ejected from the machine.

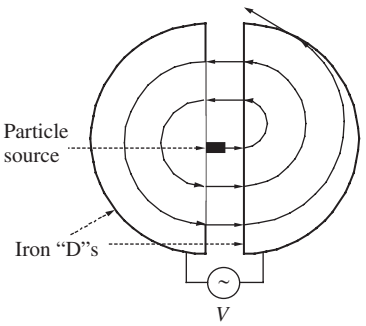


Fig. 2.14 Schematic diagram of magnetic field regions, voltage application, and design trajectory in a cyclotron. The particles are embedded in an approximately uniform magnetic field inside of the Ds, and accelerated by an electric field between the Ds.

(a) Assuming the particles are non-relativistic protons ($\gamma \cong 1$) and the magnetic field $B_0 = 1.5\,\mathrm{T}$, what is the rf frequency needed for the voltage supply? You may ignore the gap distance in this calculation.
(b) Assume the initial proton kinetic energy from the source is 150 keV, $V_0 = 150\,\mathrm{kV}$, and the radius of curvature of the outside of the Ds is 0.75 m. Approximately how many passes through the system does a proton circulate for before it is ejected?
(c) What is the value of γ of the proton at the machine exit? Can you estimate the total phase slippage of a proton with respect to the applied voltage during acceleration due to relativistic changes in the cyclotron frequency?

(2.4) The Tevatron at Fermi National Accelerator Laboratory is a circular proton (colliding with counter-propagating antiprotons) accelerator that uses high field superconducting dipole magnets. It is approximately 1 km in radius, and accelerates protons to 900 GeV in total energy.

(a) What is the average magnetic field needed to keep 900 GeV protons on the design orbit?
(b) What is the circulation (cyclotron) frequency in this accelerator? The protons are injected into the Tevatron at 150 GeV. What is the circulation frequency at this energy? What is the circulation frequency at 900 GeV?
(c) If electrons were to be stored in the Tevatron for collision with the 900 GeV protons instead of antiprotons, what would their energy be? Be careful with round-off errors due to calculators in this problem.

(2.5) Busch's theorem can also be derived by use of the Hamiltonian formalism.

(a) Using the results of Exercise 2.1, write the Hamiltonian inside of a solenoid in terms of cylindrical coordinates. Hint: Be careful in handling the vector potential component of the angular momentum.
(b) Show that the canonical angular momentum is conserved in this case.
(c) Assuming the canonical angular momentum is zero (why?), what is the value of the mechanical momentum at radial offset ρ? Compare this to Eq. (2.18).

(2.6) If the transverse momentum is not small compared to the total momentum, we may not assume $p \cong p_z$ and the betatron wave number in the solenoid is not as given by Eq. (2.24).

(a) Derive a general expression for the wave number as a function of initial offset radius at the entrance to the solenoid. Hint: Normalizing the initial offset radius to p/qB_0 may make the answer more physically meaningful.
(b) At a certain radius, the particle cannot penetrate the interior of the solenoid. This is because all of its longitudinal momentum is given up to angular

motion in the fringe field and the particle is reflected. For an electron with 100 keV kinetic energy, and $B_0 = 1$ T, what is this radius?

(2.7) By differentiation of Eq. (2.21) twice, derive Eq. (2.22).

(2.8) The high-energy linear collider at Stanford contains a linear accelerator of length $L = 3$ km, which uses electric fields in radio-frequency cavities to accelerate electrons (and positrons) to 50 GeV. The average longitudinal electric force performing work on the electrons in the accelerator is 16.7 MeV/m. The proper time τ, as observed in the frame of the accelerating electron, is related to the time measured in the laboratory t by $\mathrm{d}t = \gamma \cdot \mathrm{d}\tau$.

(a) Show that the time for the electrons to go from rest to final energy is

$$t = \frac{L}{c}\left[\sqrt{\frac{\gamma_\mathrm{f}+1}{\gamma_\mathrm{f}-1}}\right] \approx \frac{L}{c}$$

(b) Show that the proper time for the electrons to go from rest to final energy is

$$\tau = \frac{L}{c(\gamma_\mathrm{f}-1)}\cosh^{-1}(\gamma_f).$$

(c) What is the effective length of the linear accelerator as seen by the electrons at final energy? How long would it take to traverse this length at the final velocity of the electrons?

(2.9) Instead of writing the motion in terms of the variable v_z, consider the motion as a function of γ.

(a) Show that

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{\sqrt{\gamma^2-1}}{\gamma}\frac{qE_0}{mc}.$$

(b) Integrate this equation and compare to the results of directly integrating the first of Eq (2.25).

(2.10) For magnetic focusing channels that use quadrupoles, the motion is the same for particles of like charge but different mass if the momentum of each species is the same. This can be seen by examination of Eqs (2.42) and (2.43). Now consider the motion of particles with equal charge but unequal mass in electric quadrupoles. What is the condition on dynamical quantities of the different species so that they can be focused identically in the electric quadrupole channel? Show, in the relevant non-relativistic limit, that this condition reduces to the equality of kinetic energies in the two species.

(2.11) For magnetic focusing channels, show that a particle traveling in a certain direction in z has the same equation of motion as its antiparticle of the same momentum traveling in the opposite direction in z. This is the basic optics principle behind particle/antiparticle colliding beam rings. How do the paraxial optics change for these species if electric quadrupoles are used?

(2.12) In deriving Eq. (2.58) as a solution to Eq. (2.57), it is helpful to initially transform the independent variable to $u = \ln(\gamma(z))$, with inverse transformation $\gamma = \exp(u)$. This is known as a Cauchy transformation.

(a) Show that this transformation allows Eq. 2.58 to be written as

$$\frac{\mathrm{d}^2 x_\mathrm{L}}{\mathrm{d}u^2} + \left(\frac{b}{2}\right)^2 x_\mathrm{L} = 0.$$

(b) Using the equation derived in part (a), construct Eq. (2.58).

(2.13) Can you solve the paraxial equation of motion (analogous to Eq. 2.55) for the case when the electric field is constant, but the magnetic field is zero? You may start with $x' \equiv p_x/p_z$ and $p_x =$ constant, then integrate.

(2.14) For cylindrically symmetric, intense beams propagating in equilibrium in solenoids there is an electric self-field in the radial direction that produces both an outward radial force and an azimuthal rotation of frequency ω_r. This frequency may or may not be equal to the Larmor frequency, and the related rotation may be viewed as an example of the $\vec{E} \times \vec{B}$ drift discussed previously.

(a) Show, for a uniform density (n_b) beam that this radial space-charge electric force on a particle of charge q and rest mass m_0 is

$$F_\mathrm{e} = qE_\rho = \frac{q^2 n_\mathrm{b}}{2\varepsilon_0}\rho \equiv m_0\frac{\omega_\mathrm{p0}^2}{2}\rho,$$

where we have introduced the non-relativistic plasma frequency $\omega_\mathrm{p0} = \sqrt{q^2 n_\mathrm{b}/\varepsilon_0 m_0}$. Find the net (beam-induced) force, including the magnetic force, on the particle, and write it in terms of the electric force and the particle's axial velocity v_0.

(b) When the beam rotates, there are two radial forces that are introduced. One is due to centripetal acceleration, and the other is the radial component of the solenoidal force. Equilibrium (net radial force equal to zero) can be established only when the beam rotates in a certain direction because both the centripetal forces and the space-charge forces are outward. Show that the condition on ω which produces equilibrium can be written as

$$\omega_\mathrm{r}^2 + \frac{\omega_\mathrm{p}^2}{2} = \omega_\mathrm{r}\omega_\mathrm{c},$$

where the relativistically correct plasma frequency is given by $\omega_\mathrm{p}^2 = \omega_\mathrm{p0}^2/\gamma^3$.

(c) Illustrate this equilibrium condition by plotting ω_p/ω_c as a function of ω_r/ω_c. What rotation frequency ω_r maximizes the equilibrium density of the beam?

(2.15) The magnetic field given in Eq. (2.65) can be derived from a vector potential,

$$\vec{A} = -(B_0/k_u)\cos(k_u z)\hat{x},$$

in the $y = 0$ plane. Unfortunately, this vector potential does not satisfy the Laplace equation in the current-free region, $\nabla^2\vec{A} = 0$, as required by the Maxwell relations, and so it cannot be valid.

(a) Show that the following appropriate potential (uniform in x, symmetric in y)

$$\vec{A} = -(B_0/k_u)\cos(k_u z)\cosh(k_u y)\hat{x}$$

obeys the Laplace equation.

(b) Find the field components associated with this vector potential.

(2.16) A helically polarized magnetic undulator has an on-axis field profile given by

$$\vec{B}(z) = \frac{B_0}{\sqrt{2}}[\sin(k_u z)\hat{x} + \cos(k_u z)\hat{y}].$$

It is typically constructed as shown schematically in the Fig. 2.15, with two main counter-rotating helical windings (bifilar helical undulator).

(a) The vector potential which gives rise to this field can be written in cylindrical coordinates as

$$A_\phi = \frac{B_0}{k_u}[I_0(k_u\rho) + I_2(k_u\rho)]\cos(\phi - k_u z)$$

and

$$A_\rho = -\frac{B_0}{k_u}[I_0(k_u\rho) - I_2(k_u\rho)]\sin(\phi - k_u z).$$

Here I_0 and I_2 are modified Bessel functions.

Show that these components satisfy the Laplace equation, $\nabla^2\vec{A} = 0$.

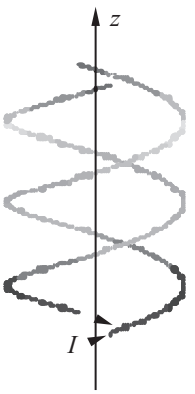


Fig. 2.15 Schematic of a bifilar helical undulator.

(b) Find the field components for this vector potential and verify that they give the correct field profile on axis.

(c) Write the Hamiltonian and equations of motion, using z as the independent variable, for the helical undulator. Show that in the limit $k_u\rho \ll 1$, there exists an orbit that is a perfect helix, in which the mechanical angular momentum is a constant. Note this is not true for the helix in the solenoid, where the canonical angular momentum is constant (the Hamiltonian is independent of azimuth).

3 Linear transverse motion

3.1 Weak focusing in circular accelerators 52
3.2 Matrix analysis of periodic focusing systems 54
3.3 Visualization of motion in periodic focusing systems 60
3.4 Second-order (ponderomotive) focusing 63
3.5 Matrix description of motion in bending systems 67
3.6 Evolution of the momentum dispersion function 70
3.7 Longitudinal motion and momentum compaction 73
3.8 Linear transformations in six-dimensional phase space* 75
3.9 Summary and suggested reading 76
Exercises 77

The phenomenon of stable transverse motion of charged particles near a design trajectory has been introduced in Chapter 2. Simple harmonic betatron oscillations were encountered in the examples of the focusing solenoid, the quadrupole, and in the circular accelerator. The general physical and mathematical tools for describing betatron oscillations are developed in this chapter. The discussion begins, appropriately enough, within the historical context of the betatron, where the ideas of path length focusing and quadrupole focusing are used to provide simultaneous stability in both transverse dimensions. After introduction of this type of first-order (in the field strength), or weak, focusing we will move to the discussion of second-order, or strong, focusing. We end this chapter with a discussion of the first-order effects in transverse motion due to the dispersion of particles having total momenta deviating from the design value.

3.1 Weak focusing in circular accelerators

The motion of a charged particle in a uniform magnetic field has been shown to exhibit effective focusing in the bend (horizontal, x) plane. In a description based on examination of trajectories along the circular design orbit, this effective focusing is ascribed to differential path lengths taken by particles on differing betatron orbits. Such path length focusing is effective in stabilizing the horizontal motion, but not the vertical motion, which is completely unconstrained in a uniform magnetic field. On the other hand, in our initial discussion of acceleration in the betatron (see Ex. 2.2), we have seen that the vertical component of the magnetic field is not uniform, but diminishes with distance away from the axis.

At the design orbit, which defines a curvilinear coordinate system, it is most natural to adopt the lowest order (in $x \equiv \rho - R$) approximation of the variation in the vertical component of the field as,

$$B_y(x) = B_0 + B'x + \cdots . \tag{3.1}$$

The field gradient is negative, $B' < 0$, in the case of the betatron. Equation (3.1) implies, along with Ampere's law, that a horizontal component of the field exists,

$$B_x(x) = \int_0^y \frac{\partial B_y}{\partial x} \mathrm{d}\tilde{y} \cong B'y. \tag{3.2}$$

In this coordinate system, it is apparent that, near the design orbit, the magnetic field appears as a superposition of vertically oriented dipole and vertically focusing (horizontally defocusing) quadrupole fields. Since the field is not

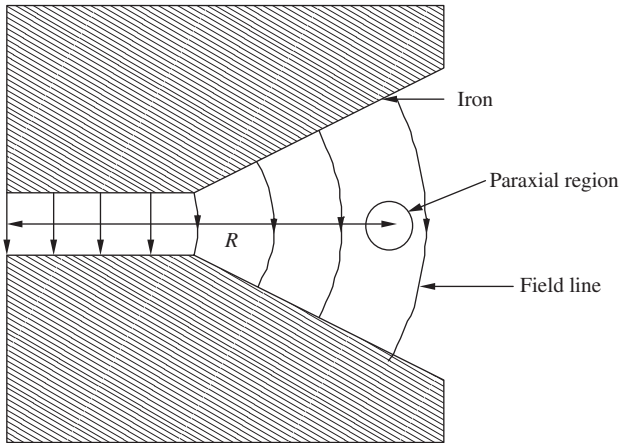


Fig. 3.1 Configuration of magnetic field in betatron, showing superposition of dipole and vertically focusing quadrupole components in vicinity of the design orbit.

designed to be a pure multipole in this case, but a superposition of two multipoles, it is termed a combined-function magnetic field (illustrated in Fig. 3.1) configuration.

Following the derivation of the horizontal equation of motion in the uniform magnetic field leading to Eq. (2.12), and using the equations of motion in a quadrupole (Eqs (2.42)–(2.43)), we can write the following equations describing paraxial motion in this combined function field:

$$x'' + \left[\left(\frac{1}{R}\right)^2 + \frac{B'}{B_0 R}\right] x = 0, \tag{3.3}$$

and

$$y'' - \frac{B'}{B_0 R} y = 0, \tag{3.4}$$

where we indicate differentiation with respect to the independent variable s as $[\,]' = \mathrm{d}/\mathrm{d}s$. If we normalize the focusing strengths in these equations to R^{-2}, we can write the resultant expressions in standard form,

$$x'' + \left(\frac{1}{R}\right)^2 [1 - n]x = 0, \tag{3.5}$$

and

$$y'' + \frac{n}{R^2} y = 0, \tag{3.6}$$

where the field index is given by $n \equiv -B'R/B_0$.

Equations (3.5) and (3.6) are often termed the Kerst–Serber equations, as they were deduced by D.W. Kerst and F. Serber during initial development of the betatron. From these relations, the condition for simultaneous stability in horizontal and vertical motion (both focusing strengths are positive) is simply given in terms of the field index,

$$0 < n < 1. \tag{3.7}$$

The simultaneous stability in x and y is clearly achieved by partially, but not fully, removing the natural path length focusing in the horizontal dimension through the introduction of a defocusing quadrupole field component. This

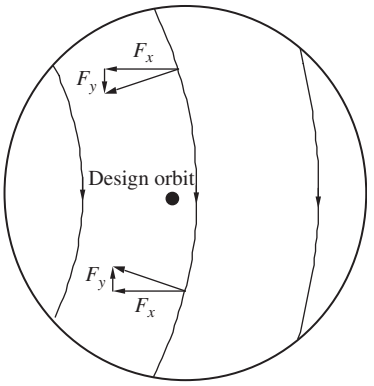


Fig. 3.2 Close-up view of paraxial region shown in Fig. 3.1, illustrating the vertically focusing components of the magnetic field at vertically offset positions.

component yields a force that is focusing in the vertical dimension, as can be seen in Fig. 3.2.

The Kerst–Serber equations are often written, especially in the context of simple magnetic optics systems, with the azimuthal angle $\theta = s/R$ around the design orbit as the independent variable,

$$\frac{\mathrm{d}^2x}{\mathrm{d}\theta^2} + [1-n]x = 0 \quad \text{or} \quad \frac{\mathrm{d}^2x}{\mathrm{d}\theta^2} + \nu_x^2 x = 0, \tag{3.8}$$

and

$$\frac{\mathrm{d}^2y}{\mathrm{d}\theta^2} + ny = 0 \quad \text{or} \quad \frac{\mathrm{d}^2y}{\mathrm{d}\theta^2} + \nu_y^2 x = 0. \tag{3.9}$$

The forms of the Kerst–Serber equations displayed in Eqs (3.8) and (3.9) are illustrative, because they give a direct normalization of the betatron oscillation frequency in terms of the circulation frequency. The tunes (normalized frequencies), $\nu_x = \sqrt{1-n}$ and $\nu_y = \sqrt{n}$, are the number of betatron oscillations per revolution in the horizontal and vertical dimensions, respectively. Assuming stability in both dimensions, the tunes are restricted to be smaller than unity. This restriction on the oscillation frequency is the source of the term "weak focusing." The weakness of the focusing becomes apparent when one attempts to scale the circular accelerator to larger design momenta and concomitant larger radii of curvature. For a given angular error of beam particle x' launched from the design orbit, the maximum offset found in the device is $x_\mathrm{m} = Rx'/\nu_x$. Thus, the size of the beam that is contained in the machine scales with the radius of curvature. This scaling has serious implications in the design of the magnets and vacuum systems of the device. If the beam size, and thus the large clear aperture between magnetic poles becomes too large, the magnet becomes difficult, if not impossible, to build. Given a way to introduce stable oscillations with much higher tunes $\nu_{x,y} \gg 1$, the size of the beam could be greatly reduced. The discovery of such a method based on quadrupole magnets, termed strong focusing, has, therefore, allowed the development of very large radius of curvature circular accelerators. We will next introduce the tools needed to analyze strong focusing.

3.2 Matrix analysis of periodic focusing systems

The betatron oscillations of a particle in a weak focusing system can be intuitively understood, as they are merely examples of simple harmonic motion. This motion is characterized by having an oscillator (focusing) strength, κ_0^2, which is of second-order in field strength and/or the field gradient (parameterized by n in Section 3.1). Specifically, in a weak focusing circular accelerator, the path length focusing, when combined with a quadrupole gradient, allows simultaneous stability in x and y, yielding a focusing strength that is of second-order in the magnetic field amplitude $B_0, \kappa_0^2 \propto B_0^2$. In a simple quadrupole, however, the motion is not simultaneously stable in both transverse dimensions. Even though quadrupoles have a focusing strength that is of first-order in field gradient, strong focusing systems—based on arrays of quadrupoles and not dependent on combination with path length focusing—are also second-order focusing systems. We also note that systems based on solenoids have a focusing

strength that is also of second-order in the field amplitude. We may infer from these examples that most focusing systems that are simultaneously stable in both transverse directions are of second order in field gradient or amplitude. In fact, this is characteristic of all focusing schemes that have no charge or current present within the beam channel.[1]

The purpose of this section, as well as in Section 3.3, is to introduce some analysis methods associated with strong focusing, in cases where the restoring force is periodic. We indicate this periodicity in the equation of motion by writing

$$x'' + \kappa_x^2(z)x = 0 \quad \text{with} \quad \kappa_x^2(z + L_\mathrm{p}) = \kappa_x^2(z). \tag{3.10}$$

The assumed period length, L_p, must be, in the context of a circular accelerator, no larger than the circumference C of the accelerator. Most large accelerators are made up of several (M_p) identical modules and, therefore, have a periodicity of $L_\mathrm{p} = C/M_\mathrm{p}$. Also, in linear accelerators and transport lines, this periodic focusing (often termed alternating gradient focusing) is used by typically employing a very simple array of quadrupole magnets with differing sign field gradients. In fact, because the act of bending the design orbit introduces an asymmetry between the focusing in the bend (x) direction and non-bend (y) direction (as seen in Section 3.1), focusing in rectilinear—as opposed to curvilinear—transport is inherently simpler. For this reason, and to make the discussion as general as possible, we adopt z as the independent variable for the description of the motion in this section as well as in Sections 3.3 and 3.4.

There are two cases of interest that can be readily analyzed: (a) the focusing is piece-wise constant, as in magnetic quadrupoles, and (b) the focusing is a sinusoidally varying function in z. The first case is quite straightforward, as for a piece-wise constant value of the focusing strength, $\kappa_x^2(z) = \kappa_0^2$ in Eq. (3.10), the problem is reduced to that of the simple harmonic oscillator. The only generalization involved is that the focusing changes its oscillator characteristics in discrete steps at a finite number of points in z. Thus, the solution to the entire problem will involve "stitching together" a number of simple harmonic oscillator solutions. This is a straightforward process based on a matrix description of the dynamics.

The second case requires a perturbative analytical approach, which cannot be used for all conceivable physical parameters, so it is by nature a bit more problematic. In this section, we only discuss the first case and leave the second case for Section 3.4. The problem of the sinusoidally varying oscillator strength is encountered in more esoteric situations in charged particle dynamics (e.g. focusing of particles due to radio-frequency waves, ponderomotive forces in laser fields), but is of high conceptual value, since it illuminates the physical basis of the matrix results.

We now consider a piece-wise constant, periodic focusing described by a focusing strength (κ_0^2) that may be positive, negative, or vanishing. Physically, these cases may correspond, for example, to a focusing quadrupole, a defocusing quadrupole, and a force-free drift, respectively. For $\kappa_0^2 > 0$, Eq. (3.10) is a differential equation that describes a simple harmonic oscillator, where the solution can be written, with the constants of integration determined in terms of an initial state vector (a vector defined in trace space),

$$\vec{x}(z_0) \equiv \begin{pmatrix} x \\ x' \end{pmatrix}_{z=z_0} = \begin{pmatrix} x_i \\ x_i' \end{pmatrix},$$

[1] An electric focusing system based on a uniform column of charged particles (opposite in sign to the beam particles) lying along the beam axis produces a lens with cylindrically symmetric electric focusing fields (see Ex. 2.14), and is, therefore, stable in both transverse directions. This type of focusing (ion focusing) has a strength that is of first-order in the resultant electric field. A similar device based on magnetic fields can be created by using a current carrying plasma column (plasma lens).

as

$$x(z) = x_i \cos(\kappa_0(z - z_0)) + \frac{x_i'}{\kappa_0} \sin(\kappa_0(z - z_0)). \tag{3.11}$$

By differentiating Eq. (3.11), we also obtain the angle

$$x'(z) = -\kappa_0 x_i \sin(\kappa_0(z - z_0)) + x_i' \cos(\kappa_0(z - z_0)). \tag{3.12}$$

Equations (3.11) and (3.12) can be conveniently represented by a matrix expression, illustrating the relationship between the initial state vector $\vec{x}(z_0)$ and final state vector $\vec{x}(z)$ as

$$\vec{x}(z) = \mathbf{M}_{\mathrm{F}} \cdot \vec{x}(z_0) \quad \text{with } \mathbf{M}_{\mathrm{F}} = \begin{bmatrix} \cos(\kappa_0(z - z_0)) & \frac{1}{\kappa_0} \sin(\kappa_0(z - z_0)) \\ -\kappa_0 \sin(\kappa_0(z - z_0)) & \cos(\kappa_0(z - z_0)) \end{bmatrix}. \tag{3.13}$$

For the transformation of the vector $\vec{x}$ through a focusing section (lens) of length l, the matrix describing the transformation is simply

$$\mathbf{M}_{\mathrm{F}} = \begin{bmatrix} \cos(\kappa_0 l) & \frac{1}{\kappa_0} \sin(\kappa_0 l) \\ -\kappa_0 \sin(\kappa_0 l) & \cos(\kappa_0 l) \end{bmatrix}. \tag{3.14}$$

In the case of a defocusing lens, we have $\kappa_0^2 < 0$ in Eq. (3.10), and the solution to this equation takes the form,

$$x(z) = x_i \cosh(|\kappa_0|(z - z_0)) + \frac{x_i'}{|\kappa_0|} \sinh(|\kappa_0|(z - z_0)), \tag{3.15}$$

with

$$x'(z) = |\kappa_0| x_i \sinh(|\kappa_0|(z - z_0)) + x_i' \cosh(|\kappa_0|(z - z_0)). \tag{3.16}$$

Thus, the transformation matrix describing the propagation of the vector $\vec{x}$ through a defocusing lens of length l is written, in analogy with Eq. (3.14), as

$$\mathbf{M}_{\mathrm{D}} = \begin{bmatrix} \cosh(|\kappa_0| l) & \frac{1}{|\kappa_0|} \sinh(|\kappa_0| l) \\ |\kappa_0| \sinh(|\kappa_0| l) & \cosh(|\kappa_0| l) \end{bmatrix}. \tag{3.17}$$

Two interesting cases can be obtained by taking the limits of Eqs (3.14) and (3.17). The first is the force-free drift, in which case it is obvious that we are taking the limit that the force disappears, $\kappa_0 \to 0$. As a result, both Eqs (3.14) and (3.17) yield the same limit,

$$\mathbf{M}_{\mathrm{O}} = \begin{bmatrix} 1 & L_{\mathrm{d}} \\ 0 & 1 \end{bmatrix}, \tag{3.18}$$

where L_{d} is the length of the drift space. In a drift, the position x changes while the angle x' does not.

The other case of interest is the so-called thin-lens limit, where $\kappa_0^2 l$ is kept finite and constant, while $l \to 0$ in Eqs (3.14) and (3.17). As a result, we have

$$\mathbf{M}_{\mathrm{F(D)}} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}, \tag{3.19}$$

where the lens focal length is defined as $f \equiv (\kappa_0^2 l)^{-1}$, and is positive for focusing lenses and negative for defocusing lenses. In the thin-lens limit, the change in position x is negligible and only the angle x' is transformed.

Two properties of the formalism developed so far should be noted. First, all of the transformation matrices—the focusing, defocusing, drift, and thin-lens matrices—have determinant equal to 1. This is a general property of linear transformations, which are precisely those that can be written as matrix transformations (see Ex. 3.8). The property of unit determinant transformation matrix is also a manifestation of Liouville's theorem, as is discussed later in this section.

The second notable property of the matrix formalism is that the full solution of the motion through a number of focusing elements (regions of constant κ_0^2) can be written in terms of the component element matrices' product. To clarify this, consider the example of a periodic system composed of a thin focusing lens, followed by a drift, a thin defocusing lens, and a final drift. In this case, the full transformation through one period of the system is written as

$$\vec{x}(L_\mathrm{p}+z_0) = \vec{x}(2L_\mathrm{d}+2l+z_0) = \mathbf{M}_\mathrm{O}\cdot\mathbf{M}_\mathrm{D}\cdot\mathbf{M}_\mathrm{O}\cdot\mathbf{M}_\mathrm{F}\cdot\vec{x}(0) \equiv \mathbf{M}_\mathrm{T}\cdot\vec{x}(z_0). \quad (3.20)$$

The total transformation matrix $\mathbf{M}_\mathrm{T}$, being the product of matrices all of unit determinant, also has the property $\det(\mathbf{M}_\mathrm{T}) = 1$. Note that the matrix product given in Eq. (3.20) is written in reverse order from that in which the component matrices are physically encountered in the beam line. Confusion on the ordering of matrices is the most common mistake made in the matrix analysis of beam dynamics!

For the example of Eq. (3.20), the total transformation matrix can be explicity written as

$$\mathbf{M}_\mathrm{T} = \begin{bmatrix} \frac{\partial x}{\partial x_i} & \frac{\partial x}{\partial x_i'} \\ \frac{\partial x'}{\partial x_i} & \frac{\partial x'}{\partial x_i'} \end{bmatrix} = \begin{bmatrix} 1 - \frac{L_\mathrm{d}}{f} - \left(\frac{L_\mathrm{d}}{f}\right)^2 & 2L_\mathrm{d} + \frac{L_\mathrm{d}^2}{f} \\ -\frac{L_\mathrm{d}}{f^2} & \frac{L_\mathrm{d}}{f} + 1 \end{bmatrix}. \quad (3.21)$$

Note that in the first display of the matrix elements in Eq. (3.21), we indicate their general meaning as the first partial derivatives of the final conditions with respect to the initial conditions. The partial derivative form of the matrix shows explicitly that it can also be interpreted as a generalized linear transformation of coordinates in trace space. The determinant of this matrix is known, in the context of coordinate transformations, as the Jacobian of the transformation; it is the ratio of differential area elements of the final and original coordinates, $\det \mathbf{M}_\mathrm{T} = \mathrm{d}x\,\mathrm{d}x'/\mathrm{d}x_i\,\mathrm{d}x_i'$. The fact that this determinant is unity in all linear trace space transformations indicates that they are area preserving, as anticipated by application of Liouville's theorem to trace space.

A few salient aspects of the transformation can be deduced by inspection of Eq. (3.21). The matrix element $M_{\mathrm{T}11}$ is the spatial magnification $(\partial x/\partial x_i)$, and it can be seen in this case to be strictly less than one in our example. Further, if this matrix element is zero, this indicates a parallel-to-point transformation of the trajectory. The "focusing" matrix element $M_{\mathrm{T}21}$ $(\partial x'/\partial x_i)$ can be compared to that found in Eq. (3.19), to deduce an equivalent thin-lens $f_\mathrm{thin} = f^2/L_\mathrm{d}$ that is always positive (focusing). The focusing indicated by $M_{\mathrm{T}21}$ is of second-order in the lens strength f because the two first-order contributions, due to the focusing and defocusing lenses, respectively, cancel. The second-order focusing naturally disappears when $L_\mathrm{d} \to 0$, as it must when two thin-lenses of equal and opposite strength are directly joined. If the two lenses were of unequal strength, a first-order effect would appear, but it would be focusing

in x and defocusing in y, or vice versa. Thus, in order that there be equal net focusing in both transverse dimensions x and y, the focusing and defocusing lens strengths in a two-lens system should be equal. When the matrix element $M_{\mathrm{T}12}$ vanishes, this indicates the point-to-point imaging condition, where the final offset is independent of the initial angle. Finally, $M_{\mathrm{T}22}$ is the magnification in x'; when it vanishes, this indicates a point-to-parallel transformation.

Thus far, we have not explicitly required that the matrix transformations discussed correspond to periodic systems. In many scenarios in charged particle optics (e.g. circular accelerators and storage rings) and in light optics (e.g. laser resonators, cf. Section 8.3), the particles are stored in a periodic focusing system like that described by Eq. (3.20) for many traversals of the system. In a large colliding beam storage ring, this may mean billions of turns around the machine. It is, therefore, of primary interest to find out what the effect of negotiating this system many times is on the particle motion—in particular, whether or not the motion is linearly stable. Assurance of the stability of particle motion under forces that are linear in displacement from the design orbit (as in Eq. (3.10)) is a necessary, but not sufficient, condition for absolutely stable motion. Nonlinear forces may also cause unstable orbits to appear at large amplitude in an otherwise linearly stable system.

In order to begin this analysis, we first note that the transformation corresponding to n passes through the system can be written as

$$\vec{x}(NL_{\mathrm{p}} + z_0) = \mathbf{M}_{\mathrm{T}}^n \cdot \vec{x}(z_0). \tag{3.22}$$

This expression is simplified if we note that the transformation of the vector can be written in terms of eigenvectors (akin to the familiar normal mode vectors of coupled oscillator systems) as

$$\vec{x}(L_{\mathrm{p}} + z_0) = \mathbf{M}_{\mathrm{T}} \cdot \vec{x}(z_0) = a_1\lambda_1\vec{d}_1 + a_2\lambda_2\vec{d}_2. \tag{3.23}$$

In Eq. (3.23), the eigenvectors $\vec{d}_j$ have the defining property that the matrix transformation only changes them by a constant factor, $\mathbf{M}_{\mathrm{T}} \cdot \vec{d}_j = \lambda_j\vec{d}_j$. The coefficients defined in Eq. (3.23) are the projections $a_j = \vec{x}(z_0) \cdot \vec{d}_j$ of the initial conditions. Using the eigenvectors, we can recast Eq. (3.22) as

$$\vec{x}(NL_{\mathrm{p}} + z_0) = \mathbf{M}_{\mathrm{T}}^n \cdot \vec{x}(z_0) = a_1\lambda_1^n\vec{d}_1 + a_2\lambda_2^n\vec{d}_2. \tag{3.24}$$

We will see below the physical meaning of these eigenvectors. Before discussing them, we first note, from Eq. (3.24), that the eigenvalues of the transformation must be complex numbers of unit magnitude, or the motion will be exponential—meaning either unbounded (positive exponent), or decaying (negative exponent). In either of these cases, the motion is termed unstable.

To find the eigenvalues, we write the transformation of an eigenvector through one period as

$$(\mathbf{M}_{\mathrm{T}} - \lambda_j\mathbf{I}) \cdot \vec{d}_j = 0, \tag{3.25}$$

where $\mathbf{I}$ is the identity matrix. Requiring the determinant of the matrix operating on the eigenvector vanish, we have

$$\lambda_j^2 - (\mathbf{M}_{\mathrm{T}11} + \mathbf{M}_{\mathrm{T}22})\lambda_j + (\mathbf{M}_{\mathrm{T}11}\mathbf{M}_{\mathrm{T}22} - \mathbf{M}_{\mathrm{T}12}\mathbf{M}_{\mathrm{T}21}) = 0, \tag{3.26}$$

or, using the fact that $\det(\mathbf{M}_{\mathrm{T}}) = 1$,

$$\lambda_j^2 - (\mathbf{M}_{\mathrm{T}11} + \mathbf{M}_{\mathrm{T}22})\lambda_j + 1 = 0. \tag{3.27}$$

Recognizing that the eigenvalue is of unit magnitude, $|\lambda_j| = 1$, when the motion is stable, we now choose to write it as $\lambda_j = \exp(\pm \mathrm{i}\mu)$. Here μ, when it is real, is referred to as the phase advance per period. With this choice, the solution to Eq. (3.27) becomes

$$\exp(\pm \mathrm{i}\mu) = \cos(\mu) \pm \mathrm{i}\sin(\mu) = \frac{\mathrm{Tr}(\mathbf{M}_{\mathrm{T}})}{2} \pm \mathrm{i}\sqrt{1 - \left(\frac{\mathrm{Tr}(\mathbf{M}_{\mathrm{T}})}{2}\right)^2}, \tag{3.28}$$

and we have employed in Eq. (3.28) the definition of the trace of the transformation matrix, $\mathrm{Tr}(\mathbf{M}_{\mathrm{T}}) \equiv M_{\mathrm{T}11} + M_{\mathrm{T}22}$. It is clear that the absolute value $|\exp(\pm \mathrm{i}\mu)| = 1$ if $|\mathrm{Tr}(\mathbf{M}_{\mathrm{T}})| \leq 2$, and, furthermore, the value of μ is real under the same condition. Thus, the condition for stable motion, in short, is

$$|\mathrm{Tr}(\mathbf{M}_{\mathrm{T}})| = |\lambda_1 + \lambda_2| \leq 2. \tag{3.29}$$

For the case of the focus-drift–defocus-drift, or FODO, array (in the context of periodic systems, the term FODO lattice is applied) discussed above, this stability criterion becomes

$$\frac{L_{\mathrm{d}}}{f} \leq 2, \tag{3.30}$$

which provides an upper limit on the defined phase advance. This limit is given in the FODO system by

$$\cos(\mu) = \frac{\mathrm{Tr}(\mathbf{M}_{\mathrm{T}})}{2} = 1 - \frac{1}{2}\left(\frac{L_{\mathrm{d}}}{f}\right)^2. \tag{3.31}$$

Note that the maximum stable phase advance per period is $\mu = \pi$. Also, it may be remarked that since the eigenvalues of stable motion are complex, the eigenvectors are generally complex.

The meaning of the phase advance per period and the associated eigenvectors can be illustrated with a few examples of ray tracing, as in Figs 3.3 and 3.4. In these ray plots, the conventional representation of focusing and defocusing lenses as thin convex and concave forms, respectively, is used. The first example, shown in Fig. 3.3, is the case $\mu = \pi/2$ ($\lambda_j = \pm \mathrm{i}$), where (the real component of) an eigenvector trajectory (1) with pure offset initial conditions (midway through

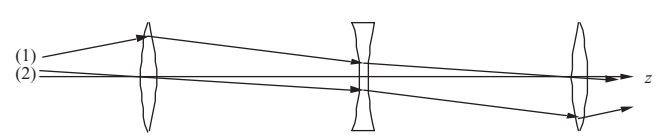


Fig. 3.3 The real components of the eigenvector rays for the case $\mu = \pi/2$ ($L/f = \sqrt{2}$) in a FODO periodic array. Ray (1) is the cosine-like trajectory and ray (2) is the sine-like trajectory (beginning propagation in the center of the first lens).

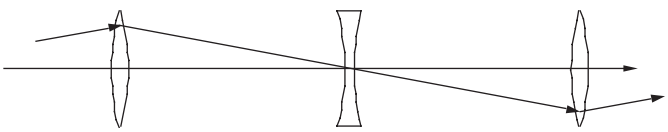


Fig. 3.4 The real component of the eigenvector ray for the case $\mu = \pi$ ($L/f = 2$) in a FODO periodic array, showing limits of stability. The eigenvalues are degenerate (both equal to -1) in this case.

the first focusing lens), $(x, x') = (x_i, 0)$, is converted to a trajectory with pure angle and no offset in final (conditions midway through the final focusing lens, $(x, x') = (0, x'_f)$). This behavior is reminiscent of the cosine function over the first 90° of phase advance. Note that trajectory (2) is the converse of trajectory (1) since it begins with only an angular deviation $(x, x') = (0, x'_i)$, and ends with only an offset $(x, x') = (x_f, 0)$, just as a sine function does over its initial 90°. This example illustrates why the initial conditions of type (1) are said to generate the cosine-like solution to the motion and the initial conditions of type (2) give rise to the sine-like solution.

This point of view suggests yet another alternative notation for the transformation matrix, indicating the matrix elements as the coefficients of the cosine- and sine-like orbits (C and S), and their derivatives (C' and S'),

$$\mathbf{M}_T = \begin{bmatrix} C & S \\ C' & S' \end{bmatrix}. \tag{3.32}$$

The cosine-like component begins in Fig. 3.3 midway through the first focusing lens with a parallel, unit-offset ray, and the sine-like component begins at this point with an on-axis ray having unit angle(!). These are just mathematical definitions; one need not be concerned with a unit angle being a violation of the paraxial ray approximation, as, in practice, the coefficient (initial angle) that is multiplied by S will always be small compared to unity.

The case of $\mu = \pi$, where $\mathrm{Tr}(\mathbf{M}_T) = -2$ and is therefore directly at the limit of stability, is shown in Fig. 3.4. It is revealed that there is only one non-trivial eigenvector corresponding to the degenerate eigenvalue $\lambda = -1$. With only one eigenvector, there can only be one type of stable trajectory, that which enters and exits the focusing lens with equal and opposite angles, and passes through the axis at the defocusing lens position. One can see that, as this situation is approached, $\mu \to \pi$, and the particles tend to undergo very large excursions at the focusing lens position, while having very small offsets near the defocusing lens. In a beam composed of many particles, this implies that the beam will be much larger near the focusing lens than it is near the defocusing lens.

It should be noted that the trace of the matrix describing a period of the motion is independent of the choice of initial position z_0—the proof of this assertion is left for Exercise 3.5. Thus, the eigenvalues and phase advance μ are independent of this choice. The eigenvectors, in contrast, are dependent on choice of z_0.

3.3 Visualization of motion in periodic focusing systems

In a periodic focusing system such as the FODO lattice introduced in the previous section, the motion in x, when plotted continuously at every point in z (see Fig. 3.5) shows a simple harmonic oscillator-like behavior, with some notable local "errors" in the trajectory. In fact, these errors can be seen to be due to a fast, yet small-amplitude, oscillation about a slower, secular simple harmonic motion. The period of the fast oscillatory motion is clearly identical to that of the FODO lattice, due to the alternating sign of the focusing and defocusing forces in a FODO period, while the secular oscillation period is much longer.

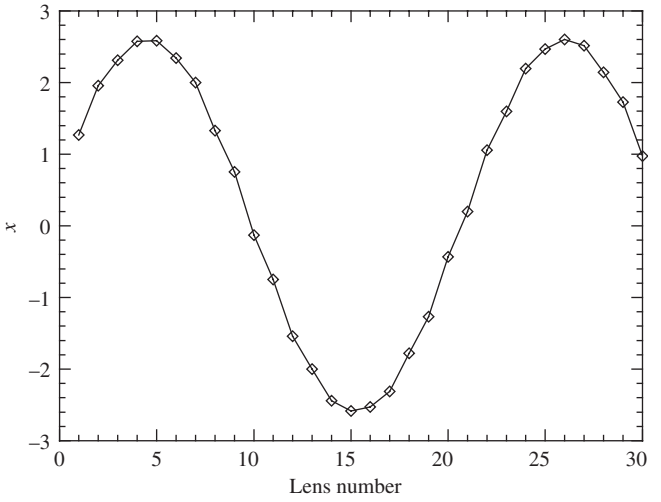


Fig. 3.5 Motion of a particle in a FODO channel with $\mu = 33°$. Lenses are at positions marked with diamond symbols. Note the deviation from simple harmonic motion occurring with the FODO period.

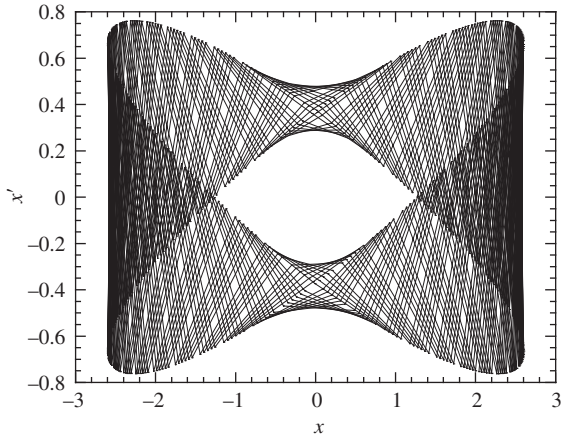


Fig. 3.6 Motion of a particle in a FODO channel of $\mu = 33°$, plotted in trace space. The fast deviations from simple harmonic motion occurring with the FODO period have a large angular spread.

As we shall see below, the fast motion is due to first-order (in the applied field amplitude) forces, while the slower motion may be ascribed to second- (or higher-order) forces that become apparent when averaging over a period of the fast oscillation. The fast motion, despite its small spatial amplitude, will also be seen to have relatively large angles associated with it.

The continuous plotting of the motion in a periodic focusing system is especially troublesome if one follows the motion in trace space, as shown in Fig. 3.6. In this plot, the fast errors in the trajectory have large angular oscillations, and the trace space plot fills in a distorted annular region, yielding unclear information about the nature of the trajectory. On the other hand, if one only plots the trace space point of a trajectory once per FODO period, a method of graphical representation known as a Poincaré plot, then the motion is regular. It is in fact an ellipse in trace space, as illustrated by Fig. 3.7, and is thus reassuringly reminiscent of a simple harmonic oscillator. This "stroboscopic" method of plotting allows the secular motion to be displayed without the interference of the fast oscillations so dramatically illustrated in Fig. 3.6.

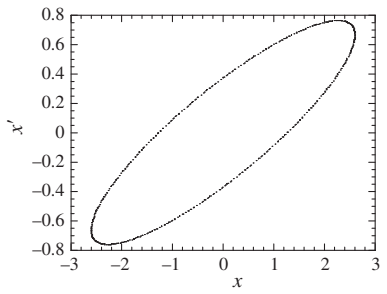


Fig. 3.7 Poincaré plot of the motion of a particle in a FODO channel of $\mu = 33°$, shown previously in Fig. 3.6, but here plotted only at the end of every FODO array.

By examining the eigenvalues and eigenvectors of the matrices for a period of the system, it is clear why the secular motion of a particle in a periodically focusing system is very nearly simple harmonic. The eigenvalues are $\lambda_j = \exp(\pm i\mu)$, and so both indicate motion with a single spatial frequency $k_{\text{sec}} = \mu/L_{\text{p}}$, where L_{p} is the length of one period. This motion is therefore simple harmonic about the real components of the eigenvector directions, which are the major and minor axes of the ellipse shown in Fig. 3.7. Since the real components of the eigenvectors are not in general parallel to the (x, x') axes, the simple harmonic motion does not necessarily yield an ellipse aligned to these axes, but one that is aligned to the eigenvector axes.

This discussion brings up the question of which quantities are independent of where in the period one chooses to interrogate the motion while performing a matrix analysis, or generating a Poincaré plot. As discussed above, and in Exercise 3.5, the trace of a matrix is independent of where in the period z_0 one chooses to begin the transport description. Thus the eigenvalues of the motion and the phase advance per period are independent of the choice of z_0. This is intuitively obvious from Fig. 3.5—the average, secular motion frequency cannot depend on which point in the fast periodic motion one begins the analysis. On the other hand, the eigenvectors depend on the choice of z_0, and so the orientation of the ellipse typified by that shown in Fig. 3.7 is also dependent on choice of z_0. The area inside of the ellipse, being related to the action, and therefore the energy of the slow, secular oscillation (see Chapter 1), is independent of z_0. This area and its physical meaning will be discussed further in this chapter, as well as in Chapter 5.

If one is concerned primarily with the approximate secular focusing effects of a periodic lattice, they can be taken into account by use of the smooth approximation, in which only the average focusing effect is used in the equation of motion,

$$x'' + k_{\text{sec}}^2 x = 0. \tag{3.33}$$

This approximation will be used in several upcoming analyses of oscillatory forces in beam physics phenomena. A direct analytical method for deriving the smooth (secular), focusing strength k_{sec}^2 will be presented in Section 3.4. For now, we restate that the average focusing strength employed in Eq. (3.33) can be simply deduced from $k_{\text{sec}} = \mu/L$. In the case of an alternating, focus–defocus (no drift) lattice (see Ex. 3.6), the phase advance is given exactly by

$$\cos(\mu) = \cos\left(\frac{\kappa_0 L_{\text{p}}}{2}\right)\cosh\left(\frac{\kappa_0 L_{\text{p}}}{2}\right). \tag{3.34}$$

For small phase advance per period, $\mu \ll 1$, Eq. (3.34) can be approximated as

$$1 - \frac{\mu^2}{2} \cong \left(1 - \left(\frac{\kappa_0 L_{\text{p}}}{2}\right)^2\right)\left(1 + \left(\frac{\kappa_0 L_{\text{p}}}{2}\right)^2\right) \quad \text{or} \quad \mu \cong \frac{1}{4\sqrt{2}}(\kappa_0 L_{\text{p}})^2, \tag{3.35}$$

and the average focusing strength is given by

$$k_{\text{sec}}^2 \cong \frac{1}{32} k_0^4 L_p^2. \tag{3.36}$$

Assuming magnetic quadrupole focusing, Eq. (3.36) indicates that the focusing strength is proportional to the applied field squared, $k_{\text{sec}}^2 \propto B'^2$. This is similar to

the situation we have found in solenoid focusing, and it again points to a general result mentioned above—all linear focusing that is equal in both transverse dimensions, and occurs in a macroscopic charge and current-free region, has an average strength that is second-order in the applied field strength. The reasons for this result in the context of quadrupole focusing will be more apparent after the discussion in Section 3.4.

3.4 Second-order (ponderomotive) focusing

As we noted at the beginning of Section 3.3, periodic focusing lattices can be analyzed in a number of ways. In many situations encountered in charged particle optics the forces are piece-wise constant, and the matrix methods of Section 3.3 provide an exact and powerful description of the motion. On the other hand, there are situations where the periodic focusing is not piece-wise constant, and another method is appropriate, one based on a Fourier decomposition of the periodic forcing function $\kappa^2(z)$. The basis of this analysis is an examination of purely harmonic, or sinusoidally varying, focusing. To be specific, we consider an equation of motion of the form

$$x'' + \kappa_0^2 \sin(k_p z) x = 0, \tag{3.37}$$

where we have chosen the phase of the focusing function for ease of further analysis,[2] and defined the wave number associated with the focusing strength period, $k_p = 2\pi/L_p$.

[2]This form indicates that we will always be choosing the phase of the focusing function so that it is odd with respect to the origin. Since we are interested only in finding the averaged second-order focusing in this analysis, no loss of generality in the discussion is suffered due to this assumption.

Equation (3.37) is classified in mathematics texts as a form of Hill's equation (an oscillator with periodic "focusing" coefficient), and also more specifically a form of the Mathieu equation (an oscillator with a component of sinusoidally periodic focusing). The exact solutions to the Mathieu equation have been studied in detail, but these solutions are not terribly useful or illuminating. For our purposes, therefore, we would like to use approximate perturbative methods from which we can learn some lessons about the physics of this system. This type of equation of motion is often found in electrodynamics, most commonly in the case of a charged particle oscillating in a spatially non-uniform electromagnetic wave, e.g. a free-electron near the focus of a powerful laser. We will return to this example later.

The discussion of Section 3.3, where we saw that the secular simple harmonic motion in a periodic lattice has a small, fast "error" trajectory overlaid upon it, motivates our choice of analysis method. The approximation we will employ here assumes that the motion can be broken down into two components, one which contains the small amplitude fast oscillatory motion (the perturbed part of the motion), and the other that contains the slowly varying or secular, large amplitude variations in the trajectory. This is written explicitly in one transverse direction as

$$x = x_{osc} + x_{sec}. \tag{3.38}$$

By this assumption, we expect the case described by Eq. (3.37), to display motion that looks qualitatively similar like the example given in Fig. 3.8.

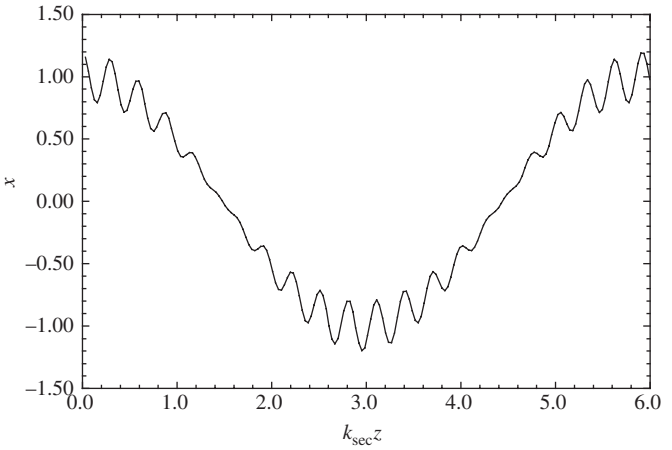


Fig. 3.8 An example of fast, small amplitude oscillatory motion superimposed upon slow, large amplitude, secular oscillation.

The oscillatory component is analyzed by making the approximation that the offset $x = x_{\rm sec} \cong$ constant ($x''_{\rm sec} \ll x''_{\rm osc}$) over an oscillation in the second term on the right hand side of Eq. (3.37),

$$x''_{\rm osc} + \kappa_0^2 \sin(k_{\rm p} z) x_{\rm sec} = 0, \tag{3.39}$$

where we have used the assumed $|x_{\rm osc}| \ll |x_{\rm sec}|$.

Since we will eventually add the secular solution $x_{\rm sec}$ to the perturbed component $x_{\rm osc}$, we do not have to examine the full solution to Eq. (3.39) at this point, only the inhomogeneous solution. We can add the homogeneous components along with $x_{\rm sec}$ later, when we apply the initial conditions to the full solution.

The inhomogeneous component of the solution to Eq. (3.39) is

$$x_{\rm osc} = \sin(k_{\rm p} z) \frac{\kappa_0^2}{k_{\rm p}^2} x_{\rm sec}. \tag{3.40}$$

This solution for the oscillatory portion of the motion is accurate if the assumptions leading to it are valid, requiring that

$$\frac{|x_{\rm osc}|}{|x_{\rm sec}|} \ll 1 \quad \text{or} \quad \frac{\kappa_0^2}{k_p^2} \ll 1. \tag{3.41}$$

In short, if the fast oscillation amplitude is small compared to the secular amplitude, then we expect the approximate solution to Eq. (3.39) to be accurate. This will be true if $\kappa_0^2/k_{\rm p}^2 \ll 1$, implying that the focusing is weak, in the sense that no significant part of a secular oscillation can occur during the focusing strength period $2\pi/k_{\rm p}$. This condition, of course, also implies that the secular phase advance per period is small, $\mu \ll 1$.

With Eq. (3.40) in hand, it is now possible to substitute it into Eq. (3.39) to obtain

$$x'' = -\kappa_0^2 \sin(k_{\rm p} z) x \cong -\kappa_0^2 \sin(k_{\rm p} z) \left[1 + \sin(k_{\rm p} z) \frac{\kappa_0^2}{k_{\rm p}^2} \right] x_{\rm sec}. \tag{3.42}$$

The next step in the analysis is to convert Eq. (3.42) into an averaged expression that will give us the behavior of the secular component of the motion, $x_{\rm sec} \equiv \langle x \rangle$, where the indicated average is over a period $L_{\rm p}$. We now obtain, by averaging Eq. (3.42) over one period of the fast oscillation,

$$x''_{\rm sec} \equiv \langle x'' \rangle = -\langle \kappa^2 \rangle x_{\rm sec} \cong -\frac{\kappa_0^2}{2}\frac{\kappa_0^2}{k_{\rm p}^2} x_{\rm sec} \tag{3.43}$$

or, in standard simple harmonic oscillator form,

$$x''_{\rm sec} + \frac{\kappa_0^4}{2k_{\rm p}^2} x_{\rm sec} = 0. \tag{3.44}$$

Equation (3.44) predicts simple oscillations having a spatial frequency (wave number) $k_{\rm sec} = \sqrt{\langle \kappa^2 \rangle} = \kappa_0^2/\sqrt{2}k_{\rm p}$. The derivation of these oscillations can be viewed as an algorithm for extracting the smooth approximation picture equivalent to Eq. (3.37). Note that one of the assumptions used in generating Eq. (3.39) was that $x''_{\rm sec} \ll x''_{\rm osc}$, which we now see implies $\kappa_0^4/2k_{\rm p}^2 \ll k_{\rm p}^2$—essentially the same criterion as given by Eq. (3.41). Equation (3.44) is derived through an averaging technique that is reminiscent of the stroboscopic visualization analysis given in the Section 3.3; the slow oscillations in an alternating periodic focusing system become easily apparent if one examines the system only once per period, á la Poincaré.

The full solution to Eq. (3.37) is, assuming that our analysis is valid for the parameters of interest,

$$x = \left[1 + \sin(k_{\rm p} z)\frac{\kappa_0^2}{k_{\rm p}^2}\right]\left[A\cos\left(\frac{\kappa_0^2}{\sqrt{2}k_{\rm p}} z\right) + B\sin\left(\frac{\kappa_0^2}{\sqrt{2}k_{\rm p}} z\right)\right], \tag{3.45}$$

where A and B are constants of integration. These constants, as mentioned above, can be used to ensure that the initial conditions are properly taken into account.

Note that, unlike our discussion in Section 3.3, we have no predictions concerning the values of the focusing amplitude that will produce instability, as Eq. (3.44) always indicates stable, simple harmonic motion. This is because Eq. (3.44) loses validity in precisely the parameter range where the system would be found unstable, $\kappa_0^2 \approx k_{\rm p}^2$. In this range, one observes that the oscillations behave as in the thin/thick lens alternating focusing system (see Ex. 3.6) in an unstable regime—x may change wildly in amplitude during a single focusing period.

For cases where our approximations are valid, one can trivially predict the phase advance per period μ. This quantity can be seen to be merely the argument of the sinusoidal functions in the secular component of x,

$$\mu \cong \kappa_0^2 L_{\rm p}/\sqrt{2}k_{\rm p} = \sqrt{2}\pi(\kappa_0^2/k_{\rm p}^2) = \kappa_0^2 L_{\rm p}^2/\sqrt{8}\pi, \tag{3.46}$$

which is nearly the same as that given by Eq. (3.36). It should be noted again that this approximation is useful only for small values of μ since the validity of the analysis implies $\kappa_0^2 \ll k_{\rm p}^2$. In fact, if we examine the treatment of the thin-lens FODO system discussed in Section 3.3, we see that the stability limit is reached when $\mu = \pi$. Thus, the estimate given in Eq. (3.46) is valid when we are far below the stability limit. If one is near the stability limit, then the actual

value of μ can be obtained by integrating Eq. (3.37) numerically over a period. One algorithm to follow involves starting from a maximum or minimum in the focusing ($k_p z_0 = 0, \pi$), which gives the correct symmetry in focusing with the initial conditions $(x, x')|_{z=z_0} = (1, 0)$ to yield the cosine-like orbit. The phase advance is then found numerically, through $\mu = \cos^{-1}(x(z_0 + L_p))$.

Once one has developed this approximate analysis of a sinusoidal focusing function, a much more general result follows immediately—the secular oscillatory behavior of an arbitrary periodic focusing function $\kappa^2(z)$ can be predicted. Using of a Fourier series representation, one can write (for odd symmetry functions $\kappa^2(z)$)

$$\kappa_x^2(z) = \kappa_0^2 \sum_{n=0}^{\infty} a_n \sin(n k_p z), \tag{3.47}$$

where the Fourier coefficients

$$a_n = \frac{2}{L_p \kappa_0^2} \int_0^{L_p} \kappa_x^2(z) \sin(n k_p z)\, \mathrm{d}z. \tag{3.48}$$

Employing the same derivation methods as those given in Eqs (3.38)–(3.43) yields an averaged expression similar to Eq. (3.44),

$$x''_{\mathrm{sec}} + \frac{\kappa_0^4}{2k_p^2} \left[\sum_{n=1}^{\infty} \frac{a_n^2}{n^2} \right] x_{\mathrm{sec}} = 0. \tag{3.49}$$

This expression can be used to more accurately compare to the results of the matrix analysis introduced in Section 3.3. For the case of a thick-lens focusing described in Exercise 3.6, we have

$$\kappa^2(z) = \begin{cases} \kappa_0^2, & 0 \le z < L_p/2, \\ -\kappa_0^2, & L_p/2 \le z < L_p, \end{cases} \tag{3.50}$$

and the Fourier coefficients are

$$a_n = \begin{cases} 4/(\pi n), & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases} \tag{3.51}$$

The secular focusing in this case is simply described by

$$x''_{\mathrm{sec}} + \frac{8\kappa_0^4}{\pi^2 k_p^2} \left[\sum_{n=1, n \text{ odd}}^{\infty} \frac{1}{n^4} \right] x_{\mathrm{sec}} = 0. \tag{3.52}$$

The sum in Eq. (3.52) is very close to unity, $\sum_{n=1,n \text{ odd}}^{\infty} n^{-4} = 1.015$. Therefore, the focusing is almost exclusively due to the fundamental harmonic $n = 1$, which gives an average focusing strength $\langle \kappa^2 \rangle = 8\kappa_0^4/\pi^2 k_p^2$. The predicted

approximate phase advance per period, including all harmonics, is

$$\mu \cong \sqrt{\langle \kappa^2 \rangle} L_\mathrm{p} = 4\sqrt{2.03}\kappa_0^2/k_\mathrm{p}^2 = 5.70 \cdot \kappa_0^2/k_\mathrm{p}^2. \tag{3.53}$$

The exact phase advance per period can be found by matrix techniques (Eq. (3.34)) and is

$$\mu = \cos^{-1}(\cos(\kappa_0 \pi/k_\mathrm{p}) \cosh(\kappa_0 \pi/k_\mathrm{p})) \cong \pi^2 \kappa_0^2/\sqrt{3}k_\mathrm{p}^2 \cong 5.70 \cdot \kappa_0^2/k_\mathrm{p}^2, \tag{3.54}$$

which is in complete agreement with Eq. (3.53). We have a curious ancillary result from this exercise, a fundamental mathematical artifact obtained by comparing Eqs (3.53) and (3.54): the series summation $\sum_{n=1,n\ \mathrm{odd}}^{\infty} n^{-4} = \pi^4/96$.

This type of system, with a sinusoidally time-dependent focusing force, or more generally, a sinusoidally time-dependent force having a linear gradient in the direction of the force, is found in many situations in physics. A simple example is that of a spatially localized electromagnetic wave (e.g. a laser beam). In this case, a charged particle oscillates under the influence of the light wave, but the force is slightly smaller during the half-cycle that the particle spends in the weaker field region. By this mechanism, the particle is pushed secularly towards regions of smaller wave intensity or energy density. This effect is known as a ponderomotive force, and is always characterized by having a secular force (the focusing strength in this case) that is second-order in the applied field amplitude.

We will also see that the results of this section are directly applicable to analyzing the focusing experienced by charged particles accelerating in radio-frequency linear accelerators. In this case, the transverse electromagnetic forces are oscillatory and quickly varying and give rise to a second-order secular, or ponderomotive, focusing effect. It is also interesting to note that a similar second-order ponderomotive effect occurs in the longitudinal dynamics of the electrons in a free-electron laser.

3.5 Matrix description of motion in bending systems

The discussions in Sections 3.3 and 3.4, in which the focusing optics are assumed to be periodic functions of the independent variable, have clearly been motivated by desire to understand circular accelerators, where the periodicity is enforced every turn around the device. Most of the tools needed to understand the motion in circular systems, which of necessity contain magnets that bend the particle trajectories, have been introduced by now, but a few remain. Here we discuss two of these tools: the method of analyzing entrance and exit effects in magnets, and a more detailed discussion of the dispersion function introduced in Chapter 2. Note that we change our notation convention slightly in this section, with the independent variable indicated by s and not z, as we have already done once before for our weak focusing analysis.

We must first examine the coordinate conventions that we are to use before we develop the description of the motion in bends. When one encounters a bend

magnet, the most common convention dictates that positive x is in the direction away from the center of curvature. If one has started the calculation of the optics with the opposite convention, it is necessary to flip the coordinates by use of the negative identity matrix, $\vec{x} = -\mathbf{I} \cdot \vec{x} = -\vec{x}$ before proceeding.[3] Then the matrix of a finite-length magnet can be broken up into three separate matrices

[3]Note we have already indicated in Chapter 2 that this transformation is by convention not applied to the vertical (y) direction simultaneously, and that an initially right-handed coordinate system will then be left-handed.

$$\mathbf{M}_{\text{mag}} = \mathbf{M}_{\text{exit}}\mathbf{M}_{\text{bend}}\mathbf{M}_{\text{entrance}}, \tag{3.55}$$

where the bend matrix in a combined-function magnet can be deduced simply from Eqs (3.3) and (3.4) to be

$$\begin{aligned}\mathbf{M}_{\text{bend}} &= \begin{bmatrix} \cos(\kappa_{\text{b}} l) & \frac{1}{\kappa_{\text{b}}}\sin(\kappa_{\text{b}} l) \\ -\kappa_{\text{b}}\sin(\kappa_{\text{b}} l) & \cos(\kappa_{\text{b}} l) \end{bmatrix} \\ &= \begin{bmatrix} \cos(\sqrt{1-n}\,\theta_{\text{b}}) & \frac{R}{\sqrt{1-n}}\sin(\sqrt{1-n}\,\theta_{\text{b}}) \\ -\frac{\sqrt{1-n}}{R}\sin(\sqrt{1-n}\,\theta_{\text{b}}) & \cos(\sqrt{1-n}\,\theta_{\text{b}}) \end{bmatrix}. \end{aligned} \tag{3.56}$$

Here the focusing wave number is given by $\kappa_x = \kappa_{\text{b}} = \sqrt{1-n}/R$, and the bend angle is $\theta_{\text{b}} = l/R$.

The edge matrices can be constructed by careful consideration of fringe field and differential path length effects. These effects will be treated here in the thin-lens approximation, meaning that the momentum kick imparted to the particle occurs in such a short distance that the particle does not change its transverse position appreciably during the transit of the edge regions. The horizontal kick can be understood purely in terms of differential path length, as illustrated in Fig. 3.9, which gives a more detailed picture of the magnet entrance edge region shown in Fig. 3.10. The offset trajectory integrates a different total magnetic force, which is linearly dependent on the amount of horizontal offset x, than the design trajectory. This integrated kick is written, following the discussion leading to Eq. (2.13), as the amount of deflecting momentum impulse (force integral) encountered due to the differing path length inside of the magnet edge at a given offset x,

$$\Delta p_{x,\text{edge}} = -q\int_{\text{edge}} B_0\left(1 + \frac{B'x}{B_0 R}\right)\mathrm{d}s \cong qB_0\tan(\theta_{\text{E}})x. \tag{3.57}$$

In Eq. (3.57), we have only included terms linear in x (due to the field index), dropping quadratic terms in this small quantity. The additional path length for the offset particle is seen to be $(\theta_{\text{E}})x$, where the angle θ_{E}, defined as the angle from the (outward) edge normal to the incoming trajectory, is taken to be positive when it points toward the center of curvature in the magnet. A positive θ_{E} indicates that, for a positive offset x, there is a "missing" field region, and the effect of the edge is defocusing. For a negative θ_{E}, there is "additional" field encountered, and the edge has a focusing effect. For the exit angles, the same convention apply, and the angular kick encountered at an edge is given by

$$\Delta x'_{\text{edge}} = \frac{\tan(\theta_{\text{E}})}{R}x. \tag{3.58}$$

The thin-lens matrix associated with an edge is thus found to be

$$\mathbf{M}_{\text{edge}} = \begin{bmatrix} 1 & 0 \\ \frac{\tan(\theta_{\text{E}})}{R} & 1 \end{bmatrix}. \tag{3.59}$$

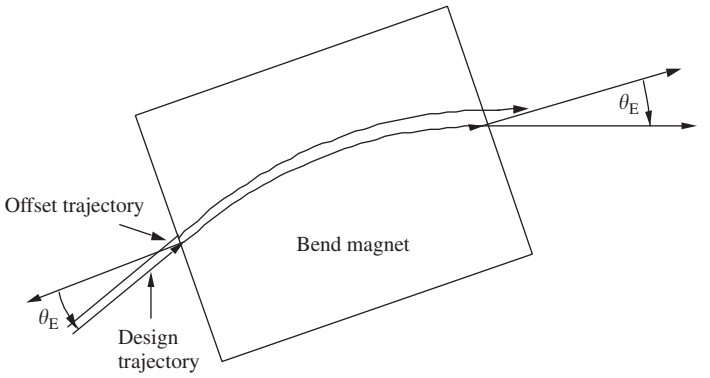


Fig. 3.9 Geometry for defining edge angles and considering the horizontal focusing effects of entrance and exit angles.

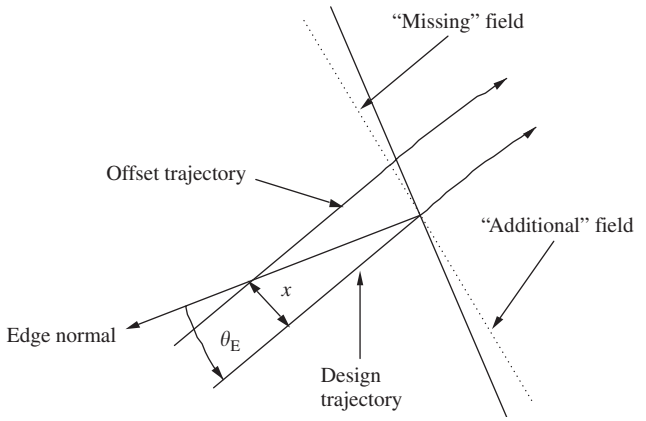


Fig. 3.10 Close-up picture of magnet entrance region showing differential path length encountered by offset trajectory. In this case, the offset trajectory sees less field region, and is defocused by the edge.

The focal length of the edge is $f = -R/\tan(\theta_E)$ and is, by convention, positive in the focusing case, $\theta_E < 0$.

As an example of this focusing, the magnet shown in Fig. 3.9 is a rectangular wedge magnet. For a field index equal to zero (flat-field magnet), it has the following total matrix transformation in x:

$$\begin{aligned}\mathbf{M}_{T,x} &= \mathbf{M}_{\text{edge}}\mathbf{M}_{\text{bend}}\mathbf{M}_{\text{edge}} \\ &= \begin{bmatrix} 1 & 0 \\ \frac{\tan(\theta_E)}{R} & 1 \end{bmatrix} \begin{bmatrix} \cos(\theta_b) & R\sin(\theta_b) \\ -\frac{1}{R}\sin(\theta_b) & \cos(\theta_b) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{\tan(\theta_E)}{R} & 1 \end{bmatrix}.\end{aligned} \tag{3.60}$$

For the symmetric trajectory shown in Fig. 3.9, the edge angle $\theta_E = \theta_b/2$, and the transformation is written as

$$\mathbf{M}_{T,x} = \begin{bmatrix} 1 & R\sin(\theta_b) \\ 0 & 1 \end{bmatrix}. \tag{3.61}$$

Note that this matrix takes the form of a simple drift, and there is no focusing or magnification associated with it.

In the vertical dimension, the edge transformation can be deduced by noting that, in the fringe field, the vertical component of the field goes from zero to the maximum value as one traces along the edge normal (see Fig. 3.10). Terming

this coordinate ξ, we have $\partial B_y/\partial\xi = \partial B_\xi/\partial y \neq 0$ in the fringe region, and thus the field lines have a component in the normal direction, $B_\xi = \int_0^y (\partial B_\xi/\partial y)\,\mathrm{d}y \cong (\partial B_y/\partial\xi)_{y=0}\, y$. If the particle enters the fringe field at a non-zero edge angle, θ_E, there is a component of B_ξ normal to the trajectory, and we have a vertical force, $F_y = qv_0 \sin(\theta_\mathrm{E}) B_\xi$. This force is gives rise to an integrated angular kick,

$$\begin{aligned}\Delta y'_{\mathrm{edge}} &= -\frac{q}{p_0}\int_{\mathrm{edge}} [B_\xi \sin(\theta_\mathrm{E})]\,\mathrm{d}s = -\frac{q}{p_0}\int_{\mathrm{edge}} \left(\frac{\partial B_y}{\partial\xi}\right)_{y=0} y\frac{\sin(\theta_\mathrm{E})}{\cos(\theta_\mathrm{E})}\,\mathrm{d}z \\ &= -\frac{qB_0}{p_0}\tan(\theta_\mathrm{E})y \cong -\frac{\tan(\theta_\mathrm{E})}{R}y. \end{aligned} \tag{3.62}$$

The vertical kick is equal in magnitude and opposite in sign to the horizontal edge kick.

For the rectangular, flat-field wedge magnet, the focusing lies entirely in the vertical dimension. The total vertical transformation matrix is written as

$$\begin{aligned}\mathbf{M}_{\mathrm{T},y} &= \begin{bmatrix} 1 & 0 \\ -\frac{\tan(\theta_\mathrm{E})}{R} & 1 \end{bmatrix}\begin{bmatrix} 1 & R\theta_\mathrm{b} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -\frac{\tan(\theta_\mathrm{E})}{R} & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1-\theta_\mathrm{b}\tan\left(\frac{\theta_\mathrm{b}}{2}\right) & R\theta_\mathrm{b} \\ -\frac{\tan\left(\frac{\theta_\mathrm{b}}{2}\right)}{R}\left[2-\theta_\mathrm{b}\tan\left(\frac{\theta_\mathrm{b}}{2}\right)\right] & 1-\theta_\mathrm{b}\tan\left(\frac{\theta_\mathrm{b}}{2}\right) \end{bmatrix}. \end{aligned} \tag{3.63}$$

This magnet is focusing and has an effective focal length given by

$$\frac{1}{f_y} = \tan\frac{\left(\frac{\theta_\mathrm{b}}{2}\right)}{R}\left[2-\theta_\mathrm{b}\tan\left(\frac{\theta_\mathrm{b}}{2}\right)\right]. \tag{3.64}$$

The rectangular wedge magnet displays the opposite characteristics of the zero-edge angle case, which is termed a **sector magnet**. In this case, the edge matrices are ignorable, and $\mathbf{M}_{\mathrm{T},x} = \mathbf{M}_{\mathrm{bend}}$. The effective horizontal focal length is $f_x = R/\sin(\theta_\mathrm{b})$, which for small angle $\theta_\mathrm{b} \ll 1$ gives the same value as Eq. (3.64). The vertical focusing is a small, but perhaps non-negligible effect in this case. It arises when the charged particle deflects noticeably in the fringe-field region of the magnet (i.e. when the magnet gap is large). The calculation of this effect is left as an exercise to the reader.

3.6 Evolution of the momentum dispersion function

Now that we have essentially concluded our introductory discussion of betatron motion, where we have been concerned with trajectories of particles possessing the design momentum, but having **offset** and **angular** errors, we can proceed to the discussion of the effects of **momentum errors**. To do this, we must re-examine the momentum dispersion function, which was introduced in Section 2.2 and defined by the differential relation $\eta_x = \partial x/\partial[\delta p/p_0]$. This function allows us to write the horizontal (transverse, in bend-plane) first-order (in betatron and momentum error amplitude) offset as (see Eq. (2.14)) $x = x_\beta + \eta_x(\delta p/p_0)$, with the betatron offset x_β being governed by the analyses of Sections 3.1–3.5. If one needs to analyze a beam optics system with bends that do not lie all in one plane, it is also necessary to introduce a vertical momentum dispersion function, so that $y = y_\beta + \eta_y(\delta p/p_0)$.

The momentum dispersion function, when multiplied by the relative momentum error $\delta p/p_0$, is most usefully thought of as the trajectory of a particle with a (unit) relative momentum error alone, and no betatron error. Therefore, the dispersion is a normalized trajectory and, to the extent that the momentum error is small, will behave nearly as a paraxial betatron trajectory when subjected only to focusing forces—it will approximately obey Eq. (3.10) in such a case. Introducing a bending force on the design orbit produces a qualitatively different effect, however, because the radii of curvature associated with the differing momenta are different. Since the differential acceleration of these orbits is constant, the horizontal dispersion function is also driven by curvature term $x''_{\delta p/p_0} = \eta''_x \cdot (\delta p/p_0) = R(p_0)^{-1} - R(p_0 + \delta p/p_0)^{-1}$, giving a new term in the dispersion evolution equation, $\eta''_x \cong R(p_0)^{-1} \equiv R_0^{-1}$. With the quadrupole and path length focusing terms included as well, we have a differential equation governing the evolution of the horizontal dispersion,

$$\eta''_x + \left(\frac{1}{R_0^2} + \frac{qB'}{p_0}\right)\eta_x = \frac{1}{R_0},$$

or more simply

$$\eta''_x + \kappa_b^2 \eta_x = \frac{1}{R_0}. \tag{3.65}$$

Here, $R_0 = R(p_0)$ is the design radius of curvature and $\kappa_x = \kappa_b = \sqrt{(1-n)}/R_0$ in a bend. The left-hand side of the equation (the homogeneous portion of Eq. (3.65)) has the same form of solutions as given by Eqs (3.11) and (3.15). Equation (3.65) has an inhomogeneous component in a bend and, therefore, has a particular solution in a bend magnet, $\eta_{x,\text{part}} = 1/\kappa_b^2 R_0$. The full solution is, assuming a (possibly combined-function) magnet with net interior horizontal focusing ($\kappa_b^2 > 0$, or $n < 1$), formally

$$\eta_x = A\cos(\kappa_b s) + B\sin(\kappa_b s) + \frac{1}{\kappa_b^2 R_0}, \tag{3.66}$$

where the magnet entrance is taken to be $s = 0$. The full solution can be constructed by matching boundary conditions at the entrance of the bend magnet. Continuity of the horizontal dispersion and its derivative at the entrance to the magnet requires

$$\eta_x(s) = \frac{1}{\kappa_b^2 R_0} + \left[\eta_x(0) - \frac{1}{\kappa_b^2 R_0}\right]\cos(\kappa_b s) + \frac{\eta'_x(0)}{\kappa_b}\sin(\kappa_b s), \tag{3.67}$$

with

$$\eta'_x(s) = \left[\frac{1}{\kappa_b R_0} - \kappa_b \eta_x(0)\right]\sin(\kappa_b s) + \eta'_x(0)\cos(\kappa_b s). \tag{3.68}$$

This linear transformation is written in matrix form by using a 3×3 matrix based on the betatron matrix operating on a dispersion state vector with a dummy

third entry, that is,

$$\begin{pmatrix} \eta_x(s) \\ \eta_x'(s) \\ 1 \end{pmatrix} = \begin{bmatrix} \cos(\kappa_b s) & \frac{1}{\kappa_b}\sin(\kappa_b s) & \frac{1-\cos(\kappa_b s)}{\kappa_b^2 R_0} \\ -\kappa_b \sin(\kappa_b s) & \cos(\kappa_b s) & \frac{\sin(\kappa_b s)}{\kappa_b R_0} \\ 0 & 0 & 1 \end{bmatrix} \times \begin{pmatrix} \eta_x(0) \\ \eta_x'(0) \\ 1 \end{pmatrix}. \tag{3.69}$$

For defocusing systems with $\kappa_b^2 < 0$, it is straightforward to show that

$$\begin{pmatrix} \eta_x(s) \\ \eta_x'(s) \\ 1 \end{pmatrix} = \begin{bmatrix} \cosh(|\kappa_b| s) & \frac{1}{\kappa_b}\sinh(|\kappa_b| s) & -\frac{(1-\cosh(|\kappa_b| s)}{|\kappa_b^2| R_0} \\ \sinh(|\kappa_b| sl) & \cosh(|\kappa_b| s) & \frac{\sinh(|\kappa_b| s)}{|\kappa_b| R_0} \\ 0 & 0 & 1 \end{bmatrix} \times \begin{pmatrix} \eta_x(0) \\ \eta_x'(0) \\ 1 \end{pmatrix}. \tag{3.70}$$

The upper left hand 2×2 block in the transformation matrices of Eqs (3.69) and (3.70) are simply the betatron transformation matrices, **M**.

When the beam is in a straight section, $R_0 \to \infty$, and the upper two elements of the third column vanish of the matrix—the third column and third row do not affect the horizontal dispersion or its derivative, as expected. It should be noted in this regard that, since edge matrices are of negligible length, there is no bending of the design orbit and the dispersion transformation matrices are constructed just as in a straight section. That is, the upper left hand 2×2 block in the transformation matrix is the thin-lens matrix, and the final column is (0,0,1).

As with betatron motion, convention dictates that positive η_x be defined in the direction away from the center of curvature. If one has started the calculation of the optics with the opposite convention (for instance, when the direction of a bend changes from one magnet to the next), it is again necessary to flip the sign of η_x and η_x' before proceeding with the matrix calculation of the dispersion and its derivative. In matrix form, this procedure is written

$$\begin{pmatrix} \eta_x \\ \eta_x' \\ 1 \end{pmatrix}_{\text{new}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \eta_x \\ \eta_x' \\ 1 \end{pmatrix}_{\text{old}}. \tag{3.71}$$

For cases where the bending field changes continuously in space, it is not possible to utilize the matrix-based approach to the solution of the dispersion evolution, and one must solve Eq. (3.65) in another manner. As an example of this type of scenario, we recall the case of the magnetic undulator introduced in Section 2.8. In this device, the undulating component of the motion was seen to be

$$x = \frac{qB_0}{p_0 k_u^2}\sin(k_u z), \tag{3.72}$$

from which we have the relation

$$\frac{\partial x}{\partial p_0} = -\frac{qB_0}{p_0^2 k_u^2}\sin(k_u z), \tag{3.73}$$

and

$$\eta_x = p_0 \frac{\partial x}{\partial p_0} = -\frac{qB_0}{p_0^2 k_\mathrm{u}^2} \sin(k_\mathrm{u} z) \equiv -\frac{a_\mathrm{u}}{\beta_0 \gamma_0 k_\mathrm{u}} \sin(k_\mathrm{u} z). \tag{3.74}$$

Here, we introduce the undulator parameter $a_\mathrm{u} \equiv qB_0/k_\mathrm{u} m_0 c$, which we will see (in Chapter 8) plays a central role in the operation of the free-electron laser.

3.7 Longitudinal motion and momentum compaction

While this chapter is nominally only concerned with first-order transverse motion, in this section we open the door to considering the effects of momentum errors on longitudinal motion by further discussing the effects of momentum dispersion. As we shall see below, momentum dispersion plays a critical role in the first-order theory of longitudinal motion and, thus, in the interest of completeness, we include a discussion of this dispersion-based effect. In particle transport systems, longitudinal motion in accelerators can be analyzed approximately using a Taylor series expansion of the motion in the relative longitudinal momentum error. The point of the present analysis is to generate a longitudinal equation analogous to the paraxial betatron equations of motion found in our treatment of transverse motion, for example, Eqs (3.5) and (3.6). In nonbending systems, this analogy is particularly close, because the paraxial betatron equations are also based on first-order expansion of the motion in transverse momentum errors. With bending systems, because of dispersion, the equations of motion are slightly more complicated.

Since the independent variable has been taken to be distance along the design trajectory (s or z), the canonical dependent coordinate in the longitudinal direction is time t, which we then measure in Hamiltonian analyses relative to the time of arrival of the design particle, $\tau = t - t_0$. The variable that plays the role of the "momentum," which is canonical to time "coordinate," is the particle's mechanical energy. However, since we will begin our discussion of longitudinal motion without employing Hamiltonians, and have invested much effort in the previous analyses based on momentum differences, we will continue in this somewhat less rigorous way. To make the present analysis connect more easily to Hamiltonian approaches, we will find it useful to introduce a parameterization of the time through a spatial variable, $\zeta = -v_0 \tau$. This is the distance that must be traveled at the design velocity by the design particle, to reach the position of the temporally advanced (or delayed) particle.

The time of flight of an off-momentum particle to a point along the design orbit is given by

$$\tau(p) = \frac{L(p)}{v(p)}, \tag{3.75}$$

where $L(p)$ is the distance traveled by the particle as a function of its momentum. The first-order logarithmic expansion of this expression is, assuming the paraxial approximation,

$$\frac{\delta\tau}{t_0} = \frac{\delta L}{L_0} - \frac{\delta v_z}{v_0} \cong \left[\alpha_\mathrm{c} - \frac{1}{\gamma_0^2}\right] \frac{\delta p}{p_0}. \tag{3.76}$$

Here we define the path length parameter $\alpha_c \equiv (\delta L/L_0)/(\delta p/p_0)$ and use $\delta v_z/v_0 \cong (1/\gamma_0^2)(\delta p/p_0)$. Equation (3.76) is used to define the momentum compaction, sometimes termed the time dispersion,

$$\eta_\tau \equiv \frac{\partial(\delta\tau/t_0)}{\partial(\delta p/p_0)} = \alpha_c - \frac{1}{\gamma_0^2}, \tag{3.77}$$

which is analogous to the horizontal dispersion discussed in Section 3.6. For a given section of particle transport, e.g. a complete revolution around a circular accelerator, there is a certain energy at which the time dispersion vanishes, and all particles pass through the system in the same amount of time. This energy, given by

$$\gamma_t = \alpha_c^{-1/2}, \tag{3.78}$$

is termed the transition energy. Below transition, particles of higher momentum pass through the system more quickly, which is the natural state of affairs in linear systems. Above transition energy they take—somewhat anti-intuitively—more time to pass through the system, since the added path length of a higher-momentum trajectory outweighs the added advantage in velocity, which becomes progressively smaller as particles become more relativisitic.

To lowest order in momentum error, the path length parameter is

$$\alpha_c = \frac{1}{s - s_0}\int_{s_0}^{s} \frac{\eta_x(\tilde{s})}{R(\tilde{s})}\,\mathrm{d}\tilde{s}. \tag{3.79}$$

Note that contributions to the integral in Eq. (3.79) vanish in straight sections, where $R \to \infty$. The path length parameter changes only in bend regions, where the linear dependence of the radius of curvature on the momentum produces a first-order difference in path length. Since the dispersion is "naturally" positive (see Ex. 3.13), one must work at making the dispersion negative and thus usually $\alpha_c > 0$. It is possible under some circumstances to make α_c vanish or be negative, in which case the transport is always "above transition," regardless of energy.

One particular magnetic system, the magnetic undulator, lends itself easily to path length parameter analysis. From Eq. (3.74), we can derive the paraxial (small bend limit) expression

$$\alpha_c = -\frac{1}{s - s_0}\int_{s_0}^{s} \frac{a_u \sin(k_u\tilde{z})}{k_u\beta_0\gamma_0 R(\tilde{s})}\,\mathrm{d}\tilde{s} \cong -\frac{1}{s - s_0}\int_{s_0}^{s} \frac{a_u^2 \sin^2(k_u\tilde{s})}{\beta_0^2\gamma_0^2}\,\mathrm{d}\tilde{s} = -\frac{a_u^2}{2\beta_0^2\gamma_0^2}, \tag{3.80}$$

and

$$\eta_\tau = -\frac{1}{\gamma_0^2}\left[1 + \frac{a_u^2}{2\beta_0^2}\right]. \tag{3.81}$$

This result should be compared with the one that can be directly deduced from the longitudinal momentum calculation of the paraxial undulator given by

Eq. (2.75). In this limit we can write, with $v_z = p_z c^2/\sqrt{p_0^2 c^2 + (m_0 c^2)^2}$,

$$\eta_\tau = -\frac{\partial v_z}{\partial p_z}\frac{p_0}{v_0} = -\left[\frac{1}{m\gamma_0^3} + \left(\frac{qB_0}{2k_u p_0}\right)^2 \left(\frac{1+\beta_0^2}{\beta_0^2(\gamma_0 mc)^3}\right)\right] m\gamma_0$$
$$\cong -\frac{1}{\gamma_0^2}\left[1 + \frac{a_u^2}{2}\right], \tag{3.82}$$

in good agreement with Eq. (3.81).

3.8 Linear transformations in six-dimensional phase space*

Previously in this chapter, we have introduced analyses of the motion based on linear dependences of the dynamics on errors in the momenta in one phase plane (actually, one trace space) at a time. In practice, as one needs to keep track of the motion in the entire six-dimensional phase space at once when designing an actual accelerator, the 2×2 matrices that allow transformation of one two-entry trace space vector at a time is replaced by a 6×6 matrix that contains all three $((x, x')(y, y'), (\zeta, \zeta'))$ trace space planes, plus additional information, the possible coupling between trace spaces. This generalized transport matrix transforms the six-entry phase space vector $\vec{\Phi} \equiv (x, x', y, y', \zeta, \zeta')$, where for symmetry we have changed our notation somewhat in writing the normalized longitudinal momentum error, $\zeta' = \delta p_z/p_0 \cong \delta p_0/p_0$. The standard form of this transformation is

$$\vec{\Phi}(s) \equiv \mathbf{R}(s, s_0) \cdot \vec{\Phi}(s_0), \tag{3.83}$$

where the 6×6 $\mathbf{R}(s, s_0)$ matrix is specified to transform the $\vec{\Phi}$-vector from one position in the beamline s_0 to another s. The matrix $\mathbf{R}(s, s_0)$ is formally written

$$\mathbf{R}(s, s_0) = \begin{bmatrix}
\frac{\partial x_\mathrm{f}}{\partial x_\mathrm{i}} & \frac{\partial x_\mathrm{f}}{\partial x_\mathrm{i}'} & \frac{\partial x_\mathrm{f}}{\partial y_\mathrm{i}} & \frac{\partial x_\mathrm{f}}{\partial y_\mathrm{i}'} & \frac{\partial x_\mathrm{f}}{\partial \zeta_\mathrm{i}} & \frac{\partial x_\mathrm{f}}{\partial \zeta_\mathrm{i}'} \\
\frac{\partial x_\mathrm{f}'}{\partial x_\mathrm{i}} & \frac{\partial x_\mathrm{f}'}{\partial x_\mathrm{i}'} & \frac{\partial x_\mathrm{f}'}{\partial y_\mathrm{i}} & \frac{\partial x_\mathrm{f}'}{\partial y_\mathrm{i}'} & \frac{\partial x_\mathrm{f}'}{\partial \zeta_\mathrm{i}} & \frac{\partial x_\mathrm{f}'}{\partial \zeta_\mathrm{i}'} \\
\frac{\partial y_\mathrm{f}}{\partial x_\mathrm{i}} & \frac{\partial y_\mathrm{f}}{\partial x_\mathrm{i}'} & \frac{\partial y_\mathrm{f}}{\partial y_\mathrm{i}} & \frac{\partial y_\mathrm{f}}{\partial y_\mathrm{i}'} & \frac{\partial y_\mathrm{f}}{\partial \zeta_\mathrm{i}} & \frac{\partial y_\mathrm{f}}{\partial \zeta_{\mathrm{i}_0}'} \\
\frac{\partial y_\mathrm{f}'}{\partial x_\mathrm{i}} & \frac{\partial y_\mathrm{f}'}{\partial x_\mathrm{i}'} & \frac{\partial y_\mathrm{f}'}{\partial y_\mathrm{i}} & \frac{\partial y_f'}{\partial y_\mathrm{i}'} & \frac{\partial y_\mathrm{f}'}{\partial \zeta_\mathrm{i}} & \frac{\partial y_\mathrm{f}'}{\partial \zeta_\mathrm{i}'} \\
\frac{\partial \zeta_\mathrm{f}}{\partial x_\mathrm{i}} & \frac{\partial \zeta_\mathrm{f}}{\partial x_\mathrm{i}'} & \frac{\partial \zeta_\mathrm{f}}{\partial y_\mathrm{i}} & \frac{\partial \zeta_\mathrm{f}}{\partial y_\mathrm{i}'} & \frac{\partial \zeta_\mathrm{f}}{\partial \zeta_\mathrm{i}} & \frac{\partial \zeta_\mathrm{f}}{\partial \zeta_\mathrm{i}'} \\
\frac{\partial \zeta_\mathrm{f}'}{\partial x_\mathrm{i}} & \frac{\partial \zeta_\mathrm{f}'}{\partial x_\mathrm{i}'} & \frac{\partial \zeta_\mathrm{f}'}{\partial y_\mathrm{i}} & \frac{\partial \zeta_\mathrm{f}'}{\partial y_\mathrm{i}'} & \frac{\partial \zeta_\mathrm{f}'}{\partial \zeta_\mathrm{i}} & \frac{\partial \zeta_\mathrm{f}'}{\partial \zeta_\mathrm{i}'}
\end{bmatrix}, \tag{3.84}$$

where the subscript i indicates the initial value of the vector element at s_0, and the subscript f indicates its final value at s. Equations (3.83) and (3.84) explicitly give the final state phase space vector components in terms of their first-order dependences on the initial phase space vector components, e.g.

$$x_\mathrm{f} = \frac{\partial x_\mathrm{f}}{\partial x_\mathrm{i}} x_\mathrm{i} + \frac{\partial x_\mathrm{f}}{\partial x_\mathrm{i}'} x_\mathrm{i}' + \frac{\partial x_\mathrm{f}}{\partial y_\mathrm{i}} y_\mathrm{i} + \frac{\partial x_\mathrm{f}}{\partial y_\mathrm{i}'} y_\mathrm{i}' + \frac{\partial x_\mathrm{f}}{\partial \zeta_\mathrm{i}} \zeta_\mathrm{i} + \frac{\partial x_\mathrm{f}}{\partial \zeta_\mathrm{i}'} \zeta_\mathrm{i}'. \tag{3.85}$$

With this form in mind, it is possible to identify the components of $\mathbf{R}(s, s_0)$ in terms of familiar quantities. The upper left diagonal 2×2 block in $\mathbf{R}(s, s_0)$

is the horizontal betatron transformation matrix; the middle diagonal 2×2 block is the vertical betatron transformation matrix; the lower right diagonal 2×2 block is the longitudinal linear transformation matrix (which will be more familiar when we study acceleration in Chapter 4). Other matrix elements outside of these blocks have also been previously discussed—for example R_{16} and R_{26} are η_x and η'_x, respectively, and R_{56} is $-\eta_\tau \Delta s$. Matrix elements linking the initial and final x and y phase planes would be non-zero if the planes are coupled, as happens in the solenoid, where the Larmor rotation completely mixes the upper left 4×4 diagonal block (cf. Ex. 3.20).

The 6×6 transport matrix $\mathbf{R}(s, s_0)$ has much in common with its 2×2 diagonal blocks. For instance, in the absence of acceleration, it has unit determinant, which is a manifestation of Liouville's theorem concerning phase space density. The 6×6 transport matrix transformation of the $\vec{\Phi}$-vector is used in most of the computer codes employed for accelerator beam dynamics calculations.

3.9 Summary and suggested reading

This chapter began with a discussion that deepened the introductory remarks made on betatron oscillations in Chapter 2, introducing the notion of weak focusing in circular accelerators. In such a scheme, one can have simultaneous stability in both vertical and horizontal transverse dimensions.

In order to circumvent the natural scaling of weak focusing, strong focusing based on periodic arrays of focusing and defocusing (alternating gradient) quadrupoles has been introduced. The piece-wise periodic focusing arising from these arrays is analyzed by powerful matrix methods. These methods, along with the stroboscopic Poincaré trace space-mapping visualization tool, have been explored in detail in this chapter. An alternative approach to understanding periodic focusing, which employs a perturbative analytical technique, has allowed us to identify alternating gradient focusing as a type of ponderomotive focusing.

Periodic focusing naturally arises in the context of circular machines, so we, of necessity, examined strong focusing effects, using matrix techniques, in bending systems. The bending of the design trajectory also introduces a new class of paraxial trajectory error—momentum dispersion. This phenomenon was also treated in this chapter by use of matrix methods.

Momentum dispersion couples the longitudinal phase plane with the transverse phase plane, by causing trajectory errors in the bend plane. Likewise, these trajectory errors can also affect the longitudinal motion of a particle, by changing its path length through a section of beamline. We have analyzed the competition between this effect and the change in the velocity of the particle on the time-of-flight through the system. The results we have obtained will be needed in Chapters 4 and 5 when we discuss longitudinal motion in circular accelerators.

Trace (or phase) planes are, in general, coupled by effects like dispersion and rotation (e.g. in solenoids). Thus, we ended this chapter by introducing the general six-dimensional square matrix description of the six-dimensional phase space dynamics of a charged particle.

The material in this chapter contains the core concepts of charged particle transverse optics. As such, many other texts in accelerator physics also treat

the subjects we have introduced here. There is a wide variety of approaches to discussing linear transverse motion and many levels at which the discussion is given in other books. The following texts may be recommended as a supplement to this chapter:

1. P. Dahl, *Introduction to Electron and Ion Optics* (Academic Press, 1973). A primer on optics.
2. D.C. Carey, *The Optics of Charged Particle Beams* (Harwood Academic Publishers, 1987). A complete treatment of transverse charged particle motion, including an excellent treatment of higher-order (nonlinear) forces.
3. D. Edwards and M. Syphers, *An Introduction to the Physics of High Energy Accelerators* (Wiley, 1993).
4. H. Wiedemann, *Particle Accelerator Physics I: Basic Principles and Linear Beam Dynamics* (Springer-Verlag, 1993). The rigor of the presentation in this text should be helpful in clarifying issues arising from our less formal presentation.
5. S.Y. Lee, *Accelerator Physics* (World Scientific, 1996).

A number of texts may be used as an introduction to advanced topics in transverse beam dynamics:

6. M. Reiser, *Theory and Design of Charged Particle Beams*. This is an encyclopedic reference on transverse motion in beams, with and without collective effects.
7. M. Berz, *Modern Map Methods in Particle Beam Physics*. Rigorous treatment of modern analytical methods for study of linear and nonlinear beam dynamics. Written at the graduate-to-professional level.
8. H. Wiedemann, *Particle Accelerator Physics II: Nonlinear and Higher-order Beam Dynamics* (Springer-Verlag, 1999). Advanced topics in nonlinear dynamics, using the first volume in the series as a basis. Written at the graduate-to-professional level.

Exercises

(3.1) As will be discussed in detail in a following chapter, the surfaces of the iron in a ferromagnet are roughly magnetic equipotentials. Given this fact, what mathematical form should a combined-function magnet surface like that shown in Fig. 3.1 have to support a combination of dipole and quadrupole fields?

(3.2) Assuming a mechanism for equipartion of energy, and thus temperature equilibrium, between phase planes, so that the root-mean-square (rms) angle $x'_{\rm rms} = \sqrt{\langle x'^2 \rangle} = \sqrt{kT/m_0c^2}/\beta\gamma = y'_{\rm rms}$, what field index n should one use to guarantee that the beam sizes in a betatron obey $x_{\rm rms} = 2y_{\rm rms}$? You can use $x_{\rm rms} = Rx'_{\rm rms}/\nu_x, y_{\rm rms} = Ry'_{\rm rms}/\nu_y$.

(3.3) Explain, in terms of the transverse velocity components of the motion, why the oscillator strength associated with betatron oscillations in solenoids is of second-order, even though the force on the particle is first order in the field amplitude.

(3.4) Consider a thin-lens system in which there is a repetitive application of a focusing lens with focal length f, each separated by a drift of length L.

(a) What is the total transformation matrix corresponding to one period of the system in this case?
(b) What is the relationship between f and L that guarantees linear stability of the transformation?

(3.5) To prove that the trace of a matrix $\mathbf{M}_1$ is the same as that of another matrix $\mathbf{M}_2$, it is sufficient to show that the two matrices are related by a similarity transformation,

$$\mathbf{M}_1 = \mathbf{A}^{-1}\mathbf{M}_2\mathbf{A},$$

where $\mathbf{A}$ is a matrix of unit determinant. Show that any two matrices representing a periodic focusing system $\mathbf{M}_1$ and $\mathbf{M}_2$, corresponding to two different choices of z_0 ($z_{0,1}$ and $z_{0,2}$), are related by such a transformation. Hint: examine matrix $\mathbf{A}$, which is the transformation matrix ($\mathbf{M}$) from $z_{0,1}$ to $z_{0,2}$.

(3.6) Consider a charged particle transport system, created by two thick lenses of length l, with opposite focusing strengths κ_0^2 and $-\kappa_0^2$, so that the matrix mapping for the focusing and defocusing lenses are

$$\mathbf{M}_\mathrm{F} = \begin{bmatrix} \cos(\kappa_0 l) & (1/\kappa_0)\sin(\kappa_0 l) \\ -\kappa\sin(\kappa_0 l) & \cos(\kappa_0 l) \end{bmatrix}$$

and

$$\mathbf{M}_\mathrm{D} = \begin{bmatrix} \cosh(\kappa_0 l) & (1/\kappa_0)\sinh(\kappa_0 l) \\ \kappa_0\sinh(\kappa_0 l) & \cosh(\kappa_0 l) \end{bmatrix},$$

respectively. These two lenses are applied repetitively so that the transformation of the coordinate vector at the nth step $\vec{x} = \binom{x}{x'}$ is $\vec{x}_{n+1} = \mathbf{M}_\mathrm{F}\cdot\mathbf{M}_\mathrm{D}\cdot\vec{x}_n$.

(a) What is the phase advance per period of the oscillation as a function of $\kappa_0 l$?
(b) Like the case of a thin-lens-based system, when κ_0^2 becomes large enough the transformation is unstable. Unlike the thin-lens case, other regions of stability are encountered when κ_0^2 is raised even. Plot a trajectory from the second stability region (the one encountered after the first unstable region). Hint: It is easy to construct this plot from the piece-wise solutions to the motion if you start with conditions $\vec{x} = \binom{x_i}{0}$ at the middle of one of the lenses. Can you tell from this trajectory why the second stability region exists for the alternating thick lens, but not the alternating thin-lens system?

(3.7) Consider a thin-lens system in which there is a repetitive application of a focusing lens with focal length f_1 and defocusing lens with focal length $-f_2$, each separated by a drift of length L. Let us examine what happens if $f_1 \neq f_2$:

(a) What is the total transformation matrix corresponding to one period of the system?
(b) Sketch out the region of stability in the parameters f_1 and f_2. This is best accomplished by drawing the borders of stability on a two-dimensional plot where the axes are f_1/L and f_2/L.

(3.8) The Wronskian determinant (or simply, the Wronskian) of a linear second-order differential equation

$$x'' + v(z)x' + w(z)x = 0$$

can be formed as the product of the two linearly independent solutions of the equation x_1 and x_2, and their first derivatives,

$$W(z) \equiv x_1(z)x_2'(z) - x_1'(z)x_2(z).$$

(a) Derive the differential equation governing the Wronskian,

$$W'(z) + v(z)W(z) = 0.$$

For $v(z) = 0$, this clearly implies that the Wronskian is constant. This constant is determined in the (simple harmonic oscillator) example of constant $w(z)$, in which case one has $W = 1$.
(b) Show that the solution to the Wronskian differential equation is

$$W(z) = W(0)\exp\left[-\int_0^z v(\tilde{z})\,\mathrm{d}\tilde{z}\right].$$

For the damped oscillator equation describing betatron oscillations during acceleration in a solenoid (Eq. (2.56)), show that the Wronskian damps along with the trace space area, that is, $W(z) = \beta\gamma(0)/\beta\gamma(z)$.

(3.9) Construct a Poincaré plot of a FODO lattice with a phase advance per period equal to 50°. Verify that the plot produces an ellipse aligned to the (x, x') axes if one begins the matrix construction in the middle of either the focusing or defocusing lenses. This requires that the thin-lens matrix of one of the lenses be split into two equal thin-lens matrices with twice the focal length.

(3.10) Consider the smooth approximation applied to a FODO lattice. Find the value of k_sec^2 in the limit that $\mu \ll 1$. How does this result compare to that of Eq. (3.36)?

(3.11) Consider a charged particle transport system created by two repetitively applied thin lenses with opposite focal lengths f and $-f$, separated by a distance L_d.

(a) Using matrix analysis, obtain the phase advance per period μ. In order to compare this result to part (b), expand Eq. (3.31) for small μ.
(b) Now using the harmonic analysis of Section 3.4, find the value of μ predicted for this system. In order to do this, you should take the periodic focusing to be given by

$$\kappa^2(z) = -\frac{1}{f}\delta\left(z - \frac{L_\mathrm{p}}{4}\right) + \frac{1}{f}\delta\left(z - \frac{3L_\mathrm{p}}{4}\right),$$

where $L_\mathrm{p} = 2L_\mathrm{d}$ is the period of the system. Compare to the result of part (a).

(3.12) Consider a so-called FOFO particle focusing system created by periodic application of a thick lens of length l with focusing strength κ_0^2, followed by a drift of length l.

(a) Using matrix methods, find the phase advance per period μ as a function of κ_0 and l.
(b) If $\kappa_0 l \ll 1$, find from the matrix result an approximate expression for the smooth approximation focusing strength k_{sec}^2 in this case. Hint: the approximate extraction of $\mu^2 \cong 4k_{\text{sec}}^2 l^2$ from the matrix calculation must include all terms up to fourth order in κ_0 and second-order in μ^2.
(c) Now consider the analysis of this system as the superposition of harmonic (sine) components in focusing strength. Write the Fourier decomposition of the focusing strength.
(d) What is the average secular focusing strength of this system? Hint: this system is a superposition of a uniform focusing system of strength $\kappa_0^2/2$, and the FD system analyzed at the end of Section 3.4, also with half-strength $\pm\kappa_0^2/2$.
(e) Compare the results of parts (b) and (d). (You should obtain near agreement by using the expansion of $\langle\kappa^2\rangle$ found in Eq. (3.49) up to fourth order in κ_0 and exact agreement if you sum the entire series in Eq. (3.49).)

(3.13) An ultra-relativistic charged particle accelerates uniformly under the influence of a longitudinal electric field $\vec{E} = E_0\hat{z}$. It is also confined by a sinusoidally varying gradient focusing quadrupole channel, mathematically stated as $B'(z) = B'_0 \sin(k_{\text{p}} z)$.

(a) Using the approximation $\beta \cong 1$ and assuming that the energy $\gamma m_0 c^2$ does not appreciably change over a period of the focusing, find an expression for the secular focusing strength.
(b) Write the transverse equation of motion *á la* Eq. (2.57) and solve.

(3.14) A common type of spectrometer magnet is shown in Fig. 3.11. The particles emitted from the point source are dispersed in momentum by the differences in radius of curvature. The momenta are well determined at the horizontal focus after the final drift B.

(a) What is the final edge angle as a function of the bend angle θ_{b}? Hint: refer to Fig. 2.6 to see the relation between θ_{b} and the pole edge orientation.
(b) Write the horizontal matrix transformation including the drifts A and B.
(c) Determine the length B that yields a point-to-point horizontal focus as a function of R, θ_{b}, and A. As one changes R, what curve do these focal points describe?

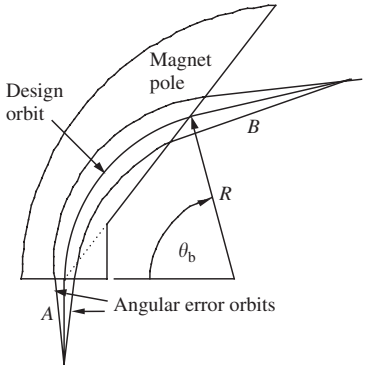


Fig. 3.11 Flat field spectrometer magnet, showing outline of pole, and focal properties.

(3.15) In the smooth approximation, one may assign the focusing in an entire strong focusing circular accelerator to be proportional to the tune, $k_{\text{sec},x} = \nu_x/R_0$. In this case, the dispersion may also have an average value—deduced from the particular solution to Eq. (3.65). Find this value for the Tevatron at Fermilab, in which $\nu_x = 19.4$, and the average radius of curvature is $R_0 = 1$ km.

(3.16) The flat-field spectrometer magnet discussed in Exercise 3.14 has interesting dispersive properties that are displayed schematically in Fig. 3.12.

(a) Show that, in the region after the magnet, the momentum dispersion is constant, $\eta'_x = 0$.
(b) Find the value of η_x in this region as function of θ_{b} and R_0.

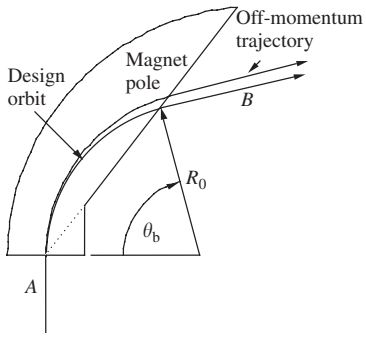


Fig. 3.12 Flat field spectrometer magnet of Fig. 3.11, showing dispersive properties.

(3.17) Evaluate α_{c} for the spectrometer magnet in Exercise 3.14.

(3.18) A transport line that translates the beam to the side while allowing the dispersion and its derivative to vanish at the second bend exit can be constructed by the deployment of magnets shown in Fig. 3.13. In this case it can be seen that the dispersion vanishes at the mid-point between the bend magnets.

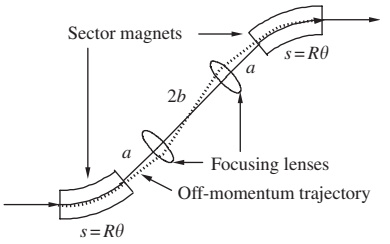


Fig. 3.13 Schematic of side-translating beamline with no residual dispersion at exit.

(a) Derive the focal length associated with the focusing lenses that gives this behavior of the dispersion as a function of a, b, R, and θ. Hint: it is best to normalize all lengths in the problem to the radius of curvature in the dipoles, R.
(b) Evaluate α_c for this transport system. Remember that even though the dispersion function changes sign in between the bend magnets, the convention of the sign of the dispersion in the second bend, with opposing sign radius of curvature, must also change. If you do not take this into account, α_c will vanish, which is clearly not true, as indicated by the above picture.

(3.19) A common magnet configuration used to rearrange particles longitudinally is the so-called chicane, as shown below in Fig. 3.14. It has flat-field dipoles of magnetic field into and out of the viewed plane, as indicated. Note that it is similar to one period of an undulator magnet, but with a large bend angle (not small compared to unity).

The chicane consists of four bend magnets of equal strength and size so that, for the design energy, R and θ are the same in each. The first bends out an angle θ, the second and third bend in by $-\theta$, and the fourth bends out again by θ.

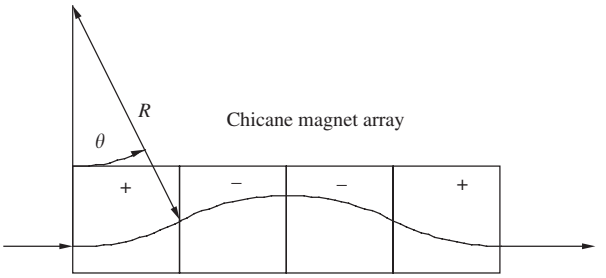


Fig. 3.14 Chicane magnet array, with design trajectory.

(a) Find the total matrix transformation in x and y. This exercise should illustrate that this array of magnets is equivalent to a drift in x and is focusing in y.
(b) Find the dispersion everywhere inside of the magnets. The quantities η_x and η_x' should disappear at the end of the fourth magnet.
(c) Find the momentum compaction $\alpha_c = 1/s - s_0 \times \int_{s_0}^{s} \eta_x(\tilde{s})/R(\tilde{s})\, d\tilde{s}$ in this device.
(d) For an electron initially trailing the 20 MeV design electron by 1 psec, chicane magnets with $R = 0.5$ m and $\theta = 30°$, what must its momentum offset δp be for it to catch up to the design particle at the chicane exit?

(3.20) Write the $\mathbf{R}(s, s_0)$ matrix describing full passage of a paraxial particle through a solenoid magnet. Hint: write the matrix as the product of three matrices—an entrance matrix that projects the initial conditions into the Larmor frame, a simple harmonic Larmor oscillation matrix (decoupled in x_L and y_L phase planes), and an exit matrix that projects the final Larmor (rotated by the Larmor angle in the x–y plane) conditions back into the x and y phase planes.

Acceleration and longitudinal motion

4

Now that we have surveyed the basic concepts of linear transverse motion in charged particle optics systems, it is time to turn our attention to the problem of acceleration. This chapter begins with an introduction to acceleration due to confined electromagnetic waves, as a way of introducing the physics of radio-frequency linear accelerators (rf linacs). Within this context, it is possible to study both the strong acceleration typical of electron linear accelerators, and the comparatively gentle acceleration found in ion linacs. Once the fundamental ways of analyzing linear acceleration and related longitudinal (along the direction of beam propagation) dynamics processes are discussed, these methods are extended to allow an understanding of longitudinal motion in the circular accelerator based on rf acceleration, the synchrotron. We end this chapter on a complementary note to the last sections of Chapter 3—we examine the possible effects of the acceleration process on transverse motion.

4.1 Acceleration in periodic electromagnetic structures 81
4.2 Linear acceleration in traveling waves 83
4.3 Violently accelerating systems 86
4.4 Gentle accelerating systems 87
4.5 Adiabatic capture 90
4.6 The moving bucket 93
4.7 Acceleration in circular machines: the synchrotron 95
4.8 First-order transverse effects of accelerating fields* 99
4.9 Second-order transverse focusing in electromagnetic accelerating fields* 102
4.10 Summary and suggested reading 106
Exercises 107

4.1 Acceleration in periodic electromagnetic structures

As will be discussed further in Chapter 7, in free space, the solutions to the electromagnetic wave equation are transversely polarized waves (the electric field is transverse to the propagation vector) that have phase velocity c, the speed of light. These properties are problematic from the viewpoint of charged particle acceleration, because in order for a charged particle to absorb energy from an applied electric force, the motion of the charged particle must have a component parallel to the electric field. This statement is quantified by the expression that gives the time-rate-of-change of the particle energy,

$$\frac{\mathrm{d}U}{\mathrm{d}t} = q(\vec{v} \cdot \vec{E}). \tag{4.1}$$

In all of the cases we consider in this chapter, the motion of the particle in an accelerating wave will be rectilinear in the z-direction, and thus the electric field must be rotated to have a longitudinal component in order for acceleration to occur. This can be accomplished by using a smooth-walled waveguide, in which case we note the existence of the familiar transverse magnetic (TM) modes.[1] These modes have a longitudinal electric field, but have phase velocity larger than c, and thus cannot remain phase synchronous with a charged particle whose velocity must always be less than c. This means that in order to allow a traveling wave to stay in a nearly constant phase relationship with an accelerating particle,

[1] If TEM modes are not familiar, please see the discussion of electromagnetic modes in waveguides given in Section 7.3.

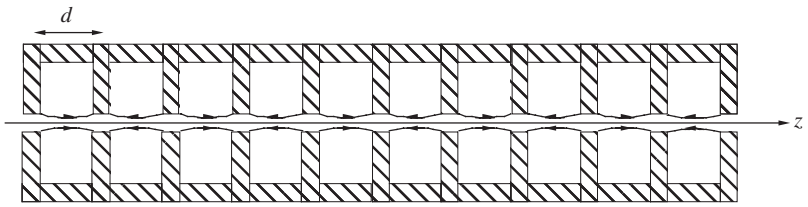


Fig. 4.1 Bisected view of a cylindrically symmetric, standing wave linear accelerator structure. The hatched portion is a conducting wall, which is typically made of copper, or a superconducting material. The electric field lines indicate the structure is operating in the π-mode, in which the longitudinal electric field changes sign every structure period.

the phase velocity of the wave must be slowed down. This is accomplished by loading the wave-guide with obstructions. An example of this loading is displayed in Fig. 4.1. Here, disks with irises cut into them about the axis of symmetry form the obstructions. The irises have two purposes: (1) to allow flow of electromagnetic energy along the z-axis, and (2) to allow the unimpeded passage of paraxial beam particles.

The linear accelerator structure[2] shown in Fig. 4.1 does not in fact display a traveling wave, but a standing wave, as can be seen from examining the longitudinal electric field pattern. The actual dependence of the solutions to the electromagnetic wave equation in such a structure is examined in detail below, with the present discussion limited to an idealized, on-axis representation of the pure harmonic (at the *rf frequency*) standing wave longitudinal electric field,

$$E_z(z,t) = 2E_0 \sin(k_z z)\cos(\omega t) = E_0[\sin(k_z z - \omega t) + \sin(k_z z + \omega t)]. \quad (4.2)$$

[2]Because the electromagnetic waves used in most linear accelerator structures are in the radio band, they are often termed radio-frequency linear accelerator structures, or more compactly, rf linacs.

It can be seen from this expression that the standing wave can be written as the superposition of two traveling waves. The forward wave component of the standing wave has argument, $k_z z - \omega t = k_z(z - v_\varphi t)$ where $v_\varphi \equiv \omega/k_z$. The backward wave component has argument $k_z z + \omega t$, and thus has equal and opposite phase velocity $v_\varphi = -\omega/k_z$. For a standing wave, both forward and backward wave components have equal amplitude.

It should be noted that the structure shown in Fig. 4.1 is periodic, with period length d. The two traveling wave components listed in Eq. (4.2) are a subset of all possible solutions of the wave equation with periodic boundary conditions. In fact, a theorem due to Floquet states that the spatial component of the solutions to the Helmholtz equation (the simplified wave equation obtained after substitution of a harmonic time dependence $\exp(-i\omega t)$),

$$\left[\vec{\nabla}^2 + \frac{\omega^2}{c^2}\right]\begin{Bmatrix}\vec{E}\\ \vec{B}\end{Bmatrix} = 0 \quad (4.3)$$

with spatial periodicity enforced by boundary conditions, can always be written in the form $E_i(z+d) = E_i(z)\exp(i\psi)$. Thus, a given solution is characterized by a phase shift per period ψ.

Simple Fourier decomposition of the on-axis solution then gives the useful form

$$E_z(z) = E_0 \,\mathrm{Im} \sum_{n=-\infty}^{\infty} a_n \exp\left[i\frac{(2\pi n + \psi)}{d}z\right]. \quad (4.4)$$

With this general form of the solution, the field can be viewed as the sum of many wave components, which are termed spatial harmonics, having different longitudinal wave numbers $k_{z,n} = (2\pi n + \psi)/d$, and thus different phase

velocities $v_{\varphi,n} = \omega/k_{z,n}$. A pure traveling wave solution has only one non-vanishing amplitude coefficient a_n, whereas a pure standing wave solution has two non-vanishing a_n. In general, in order to have the full solution obey the conducting boundary conditions, all of the components of the Floquet expansion in Eq. (4.4) must be considered. To settle on a normalization convention, will take $a_0 = 1$ in this text, so that the average accelerating field at the optimal phase in the wave will always be E_0.

The structure shown in Fig. 4.1 illustrates the concept of the phase shift per period well, as it clearly shows a field reversal every period, indicating $\psi = \pi$. This is the so π-called-mode, which is a common field configuration for standing wave accelerators. For a pure harmonic standing wave field, we have $a_{-1} = -1$ and $a_0 = 1$, with all other components vanishing.

If we now consider a charged particle traveling on-axis through an accelerator structure at approximately constant velocity $v_z \cong v_{\varphi,0} = \omega/k_{z,0}$, and integrate Eq. (4.1) through an integer number of periods M, we obtain an energy gain due to each spatial harmonic of the field,

$$\Delta U = \begin{cases} qE_0 Md \sin(\varphi), & n = 0 \\ 0, & n \neq 0. \end{cases} \tag{4.5}$$

In Eq. (4.5) we have introduced φ, the phase of the particle with respect to the $n = 0$ wave crest. The $n = 0$ component is termed the **fundamental spatial harmonic**. Because it travels in this case at approximately the same velocity as the particle (we have ignored the fact that the velocity may change slightly during the acceleration over the integration interval), it is also termed the **synchronous wave**. All other non-synchronous components do not contribute in this constant-velocity approximation to the **secular** (averaged over a period) acceleration occurring over lengths greater than a period of the structure.

What this discussion illustrates is that, for cases where the velocity does not change appreciably during a period of the structure, in calculating energy gain, all components of the electric field may be neglected except the synchronous wave. In fact, even in cases where the particle changes velocity dramatically during passage through a structure, one finds that consideration of a dominant synchronous component is enough to describe the long-term secular acceleration of the particle. Any notable effects that are due to backward or other non-synchronous components are therefore typically localized to a region smaller than a structure period.

4.2 Linear acceleration in traveling waves

We begin our analysis of acceleration in traveling wave structures by adopting a Hamiltonian approach. To construct the Hamiltonian, we note that the longitudinal electric field associated with a single traveling wave in an accelerating structure can be derived from a vector potential with only a longitudinal component,[3]

$$A_z(z - v_\phi t) = -\frac{E_0}{k_z v_\phi} \cos[k_z(z - v_\phi t)], \tag{4.6}$$

as

$$E_z(z - v_\phi t) = -\frac{\partial A_z}{\partial t} = E_0 \sin[k_z(z - v_\phi t)]. \tag{4.7}$$

[3] We have at this point decided on a phase convention in our description of the sinusoidal traveling wave. There are of course other possible conventions which are followed in other analyses (we could have used, e.g. a cosine), but since we need to discuss such a wide array of physical scenarios, and since stable motion will occur at different phases depending on which scenario is discussed, we will stay with this convention throughout our analysis.

The Hamiltonian associated with this vector potential can be written as (see Eq. 1.63)

$$H = \sqrt{\left(p_{z,\mathrm{c}} + \frac{qE_0}{k_z v_\phi}\cos[k_z(z - v_\phi t)]\right)^2 c^2 + (m_0c^2)^2}, \tag{4.8}$$

where we are, consistent with a paraxial ray approximation, only considering longitudinal motion. This Hamiltonian generates the correct canonical equations of motion,

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial H}{\partial p_{z,\mathrm{c}}} = \frac{p_z c^2}{\sqrt{p_z^2 c^2 + (m_0c^2)^2}} = v_z \tag{4.9}$$

and

$$\frac{\mathrm{d}p_{z,\mathrm{c}}}{\mathrm{d}t} = -\frac{\partial H}{\partial z} = \frac{p_z c^2 (qE_0/v_\phi)\sin[k_z(z - v_\varphi t)]}{\sqrt{p_z^2 c^2 + (m_0c^2)^2}} = \frac{qE_0 v_z}{v_\varphi}\sin[k_z(z - v_\varphi t)], \tag{4.10}$$

where we have used the relationship between the mechanical and canonical momentum, $p_z = p_{z,\mathrm{c}} - qA_z$. The equation of motion for the mechanical momentum is recovered from Eqs (4.6) and (4.10),

$$\frac{\mathrm{d}p_z}{\mathrm{d}t} = \frac{\mathrm{d}p_{z,\mathrm{c}}}{\mathrm{d}t} - q\frac{\mathrm{d}A_z}{\mathrm{d}t} = \frac{\mathrm{d}p_{z,\mathrm{c}}}{\mathrm{d}t} - q\left[\frac{\partial A_z}{\partial t} + v_z\frac{\partial A_z}{\partial z}\right] = qE_0\sin[k_z(z - v_\varphi t)], \tag{4.11}$$

where we have evaluated the total time derivative at the particle position using the sum of the partial and the convective derivatives, $\mathrm{d}/\mathrm{d}t = \partial/\partial t + v_z\partial/\partial z$.

The main problem with the form of the Hamiltonian given in Eq. (4.8) is that it is not a constant of the motion, as its partial time derivative does not vanish. In order to make phase plane plots of the longitudinal motion, we must convert the form of the Hamiltonian to one in which it is constant in time. This is done by use of a canonical transformation (a Galilean, not Lorentz transformation, as in Section 1.3) of coordinate[4]

$$\zeta = z - v_\varphi t. \tag{4.12}$$

[4]This is a Galilean transformation to the wave frame. It is only a mathematical transformation, with a precise interpretation in the context of classical mechanics theory. It is not to be confused with a Lorentz transformation—it is not a physical, only a mathematical, change of variable descriptions.

With this choice of new coordinate, the new momentum is set equal to the old $p_\zeta = p_z$, and the new Hamiltonian is obtained from the old Hamiltonian as

$$\begin{aligned}\tilde{H}(\zeta, p_{\zeta,\mathrm{c}}) &= H(\zeta, p_{\zeta,\mathrm{c}}) - v_\phi p_{\zeta,\mathrm{c}} \\ &= \sqrt{\left(p_{\zeta,\mathrm{c}} + \frac{qE_0}{k_z v_\varphi}\cos[k_z\zeta]\right)^2 c^2 + (m_0c^2)^2} - v_\varphi p_{\zeta,\mathrm{c}}.\end{aligned} \tag{4.13}$$

It is clear that the new Hamiltonian is in fact a constant of the motion, and can be used as such. With this choice of coordinate, the equations of motion derived

from the new Hamiltonian are thus

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{\partial \tilde{H}}{\partial p_{\zeta,\mathrm{c}}} = \frac{p_{\zeta,\mathrm{c}} + (qE_0/k_z v_\varphi)cos[k_z\zeta]}{\gamma m_0} - v_\varphi = \frac{p_\zeta}{\gamma m_0} - v_\varphi = v_z - v_\varphi, \tag{4.14}$$

and

$$\frac{\mathrm{d}p_{\zeta,\mathrm{c}}}{\mathrm{d}t} = -\frac{\partial \tilde{H}}{\partial \zeta} = \frac{qE_0 v_z}{v_\varphi}\sin[k_z(\zeta)], \tag{4.15}$$

or, writing Eq. (4.15) in terms of the mechanical momentum,

$$\frac{\mathrm{d}p_\zeta}{\mathrm{d}t} = qE_0 \sin[k_z\zeta]. \tag{4.16}$$

As can be seen from this short discussion, use of canonical variables is a bit trickier in this case than use of familiar mechanical variables. Because we have concluded the formal discussion of the Hamiltonian, however, we can now revert to the mechanical description, to write the constant of the motion functionally as

$$\tilde{H}(\zeta, p_\zeta) = \sqrt{p_\zeta^2 c^2 + (m_0 c^2)^2} - v_\varphi p_\zeta + \frac{qE_0}{k_z}\cos[k_z\zeta]. \tag{4.17}$$

This form of the Hamiltonian could even be used (with less than perfect rigor!) to generate equations of motion, even though it is nominally not written in terms of the canonical momentum. In fact, the mechanical momentum in this time-independent case can effectively be treated as if it were canonical, since the form of Eq. (4.17) is identical to that which would be derived from an electrostatic potential giving the same acceleration as Eq. (4.16).

Since we have already derived the correct equations of motion in Eqs (4.13)–(4.16), the more important use of Eq. (4.17) is that it can be used to visualize the motion of charged particles in the longitudinal phase plane (ζ, p_ζ). Before we move into discussion of specific examples of such motion, let us note that Eq. (4.17) can be written in normalized form as

$$\frac{\tilde{H}}{m_0 c^2} = \sqrt{(\beta_z\gamma)^2 + 1} - \beta_\phi\beta_z\gamma + \alpha_{\mathrm{rf}}\cos[k_z\zeta]. \tag{4.18}$$

In Eq. (4.18), the quantity

$$\alpha_{\mathrm{rf}} \equiv \frac{qE_0}{k_z m_0 c^2} = \frac{\gamma'_{\max}}{k_z} \tag{4.19}$$

is defined as the ratio of the maximum spatial rate of change of the normalized particle energy (Lorentz factor γ), to the maximum rate spatial rate of change of the particle's phase in the wave.

It will be seen below that the type of behavior one observes in particle acceleration by a traveling wave can be divided into two distinct regimes. The first occurs when $\alpha_{\mathrm{rf}} \ll 1$, and is typically encountered in heavy particle (proton or ion) linacs, in which the acceleration is very gentle. The second regime occurs when α_{rf} is of order unity, or above. This is the regime of violent acceleration, which occurs in electron linacs. The physics of violent acceleration is discussed in the next section.

4.3 Violently accelerating systems

The case of $\alpha_{\rm rf} \geq 1$ corresponds to an accelerating wave in which the particle can gain more than one unit of rest energy by remaining in synchronism with the wave for a radian or less of spatial propagation, $k_z \Delta z \leq 1$. This is a violent acceleration scenario, in the sense that a particle can be picked up from rest and accelerated to relativistic velocities in less than one wave cycle, thus being captured by the wave. As illustrated by the examples given in Exercise 4.2, this type of acceleration in practice can only be achieved by use of the lightest charged particle, the electron (or its antiparticle, the positron). In analyzing this system, we make use of the approximation that the phase velocity of the wave reaches its ultra-relativistic limit, $v_\varphi \to c$. We adopt this approximation precisely because the acceleration in this case is violent, and thus the charged particles are expected to asymptotically (actually, within a few rf wavelengths) attain ultra-relativistic velocities. If the v_φ is chosen to be noticeably less than c, the particles can accelerate past this phase velocity, and eventually outrun the wave to the point where they may enter a decelerating phase.

In the approximation $v_\varphi = c$, we may write the mechanical version of the Hamiltonian relation, Eq. (4.17), simply as

$$\tilde{H}(\zeta, p_\zeta) = m_0 c^2 [\gamma - \beta_z \gamma + \alpha_{\rm rf} \cos[k_z \zeta]] = m_0 c^2 \left[\sqrt{\frac{1-\beta_z}{1+\beta_z}} + \alpha_{\rm rf} \cos[k_z \zeta] \right]. \tag{4.20}$$

This form of the Hamiltonian allows us to both perform rudimentary analysis, and also to draw modified phase plane plots to illustrate the dynamics of the acceleration process. An example of such a plot is shown in Fig. (4.2), in which the momentum axis is parameterized by

$$\chi = \frac{m_0 c^2}{U - p_z c} = \sqrt{\frac{1+\beta_z}{1-\beta_z}}, \tag{4.21}$$

in order to have a positive quantity displayable in a semi-log plot. The vertical axis chosen to be logarithmic in order to show the large changes in momentum as the particle becomes relativistic.

Figure 4.2, which illustrates a number of curves of constant $\tilde{H}$, shows several interesting aspects of the longitudinal motion. The first is that for the

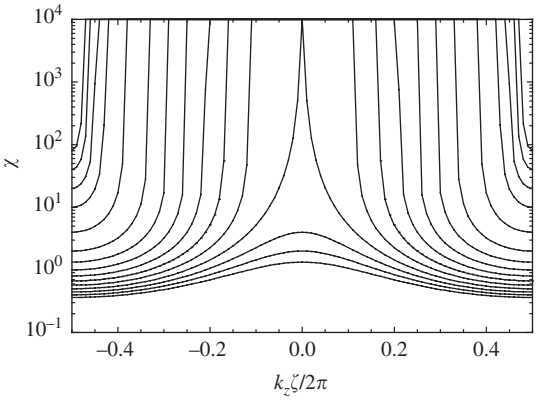


Fig. 4.2 Modified longitudinal phase plane plot, with $\alpha_{\rm rf} = 1$.

case shown ($\alpha_{\rm rf} = 1$), particles initially at rest ($\chi = 1$) can, for certain initial phases in the wave $\varphi_0 = k_z\zeta_0$, be accelerated to very high energy ($\chi \gg 1$). In addition, at high energy the phase the particle occupies in the wave is approximately stationary,

$$\tilde{H} = m_0c^2[\chi^{-1} + \alpha_{\rm rf}\cos[k_z\zeta]] \Rightarrow \alpha_{\rm rf}m_0c^2\cos[k_z\zeta] = \alpha_{\rm rf}m_0c^2\cos(\varphi_{\rm f}). \tag{4.22}$$

Since the Hamiltonian is a constant of the motion, as $\chi^{-1} \Rightarrow 0$ the phase of the particle in the wave asymptotically approaches $\varphi_{\rm f}$. There is no notable further slippage of the particle in the wave as it becomes ultra-relativistic, (that is as its velocity approaches the wave phase velocity, $v \Rightarrow c = v_\varphi$), thus giving the stationary final value of φ indicated by Eq. (4.22).

Note that in Fig. 4.2 all motion in φ is towards the left following the constant $\tilde{H}$ curves, as the particles are always traveling slower than the wave. At times this velocity difference is significant, and at other times it is imperceptible, in which case the direction of the curve is near vertical. Since the motion in φ is unidirectional, for positive φ the motion is seen to be accelerating, while for negative φ it is decelerating, as could have easily been deduced from Eq. (4.16).

One may ask what the minimum field amplitude is that allows the phenomenon of capture, where a particle starts from rest and is subsequently accelerated to very high energy. This field value is termed the trapping threshold, and gives an idea, for example, of when electrons that are field-emitted from the linac structure walls (so-called dark current) can actually be accelerated to high energy. An electron that is barely trapped by the accelerating wave will approach the fixed point at $(\varphi_{\rm f}, \chi) = (0, \infty)$ (here both the accelerating field and the particle's velocity relative to the wave vanish), so the final value of the Hamiltonian is

$$\tilde{H} = \alpha_{\rm rf}m_0c^2\cos(\varphi_{\rm f}) \Rightarrow \alpha_{\rm rf}m_0c^2. \tag{4.23}$$

The initial state is given by

$$\tilde{H} = m_0c^2[1 + \alpha_{\rm rf}\cos(\varphi_0)]. \tag{4.24}$$

We can examine the minimal conditions for trapping by assuming initial phase $\varphi_0 = \pi$, at the other vanishing phase of the field. Equating the right-hand sides of Eqs (4.23) and (4.24) (with $\varphi_0 = \pi$) gives a value for the field corresponding to $\alpha_{\rm rf} = 0.5$. The barely trapped particle is launched at $\varphi_0 = \pi$ and slips back to at final state at $\varphi_{\rm f} = 0$ for this value of $\alpha_{\rm rf}$.

4.4 Gentle accelerating systems

The longitudinal motion of a heavy charged particle (a proton, or heavier ion) in an accelerating wave is not similar to that discussed in the last section for any reasonable values of the electric field amplitude and wavelength (see Ex. 4.2(b)). For heavy particles, one always finds that $\alpha_{\rm rf} \ll 1$, and the resulting gentle accelerating motion is qualitatively different. This situation requires a different approach to the analysis than employed in the previous section. For gentle acceleration systems, the energy gain over a wavelength of the accelerator is

much less than the rest mass, and so we are not led to assume that the motion is ultra-relativistic, nor that the phase asymptotically approaches a constant value. In fact we will see that the motion in these systems is characterized by simple harmonic motion near the stable fixed point of the system. At this fixed point, the accelerating field vanishes, and the particle has the same velocity as that of the wave. In Sections 4.2 and 4.3, we have assumed that this phase velocity is constant. While we will begin the discussion of gentle acceleration under this assumption, we will see that it is necessary to allow this phase velocity to increase in order for significant, long-range gentle acceleration of particles to occur. This generalization will also cause our view of the phase plane fixed point to change somewhat.

We begin by rewriting the Hamiltonian of Eq. (4.17) by expanding it for small amplitude motion about the design momentum

$$p_0 = \gamma_\varphi m v_\varphi = \frac{m v_\varphi}{\sqrt{1-(v_\varphi/c)^2}} = \frac{m v_0}{\sqrt{1-(v_o/c)^2}} = \gamma_0 m v_0. \tag{4.25}$$

which is resonant with the phase velocity of the wave. Keeping terms up to second order in $\delta p = p_\zeta - p_0$, we have the expression

$$\begin{aligned}\tilde{H}(\zeta,\delta p_\zeta) &\cong \gamma_0 m_0 c^2 + v_0 \delta p + \frac{\delta p^2}{2\gamma_0^3 m_0} - v_0(p_0+\delta p) + \frac{qE_0}{k_z}\cos(k_z\zeta) \\ &= \frac{m_0 c^2}{\chi_0} + \frac{\delta p^2}{2\gamma_0^3 m_0} + \frac{qE_0}{k_z}\cos(k_z\zeta).\end{aligned} \tag{4.26}$$

The addition and subtraction of constants in the Hamiltonian have no effect on the form of the phase plane curves, or on the derived equations of motion. We therefore are free to reformulate Eq. (4.26) in more suggestive form,

$$\tilde{H}(\zeta,\delta p_\zeta) = (m_0 c^2)\left[\frac{\beta_0^2}{2\gamma_0 p_0^2}(\delta p^2) + \alpha_{\rm rf}[\cos(k_z\zeta)+1]\right]. \tag{4.27}$$

As a check on the derivation of Eq. (4.27), it is instructive to extract the equations of motion from it by differentiation

$$\dot{\zeta} = \frac{\partial \tilde{H}}{\partial(\delta p)} = \frac{m_0(\beta_0 c)^2}{\gamma_0 p_0^2}\delta p = \frac{\delta p}{\gamma_0^3 m_0}, \tag{4.28}$$

$$\delta\dot{p} = -\frac{\partial \tilde{H}}{\partial \zeta} = \alpha_{\rm rf} k_z m_0 c^2 \sin(k_z\zeta) = qE_0 \sin(k_z\zeta). \tag{4.29}$$

The effective longitudinal mass $\gamma_0^3 m_0$ first encountered in Eq. (2.34) is again displayed in Eq. (4.26).

It can be seen from Eq. (4.27) that the maximum fractional momentum change that can be imparted to a particle in this potential is of the order

$$\frac{\delta p_{\rm max}}{p_0} \cong \frac{\sqrt{2\alpha_{\rm rf}\gamma_0}}{\beta_0}.$$

This quantity is much smaller than one by design, however, as we are assuming that the particles are only moderately relativistic (γ_0 is not many orders of

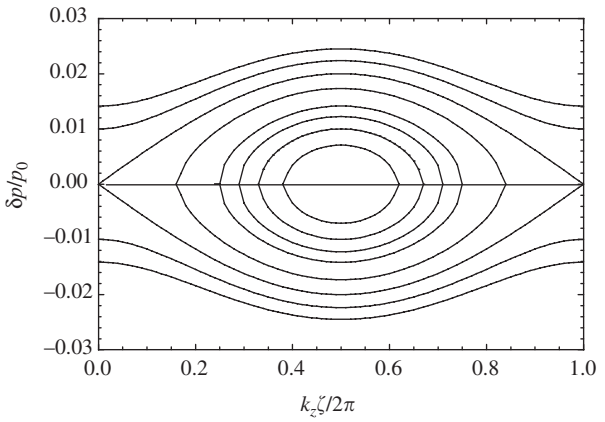


Fig. 4.3 Phase plane trajectories showing the stable region ("bucket") of vibrational motion, bounded by a separatrix, with unbounded phase motion (librational motion) outside. This case is for $\beta_0^2/\gamma_0 = 0.5$, and $\alpha_{\rm rf} = 10^{-4}$.

magnitude larger than unity), and also that $\alpha_{\rm rf} \ll 1$. Also, it can be seen that the Hamiltonian in Eq. (4.27) is of the form corresponding to a pendulum, where the stable phase—the minimum "potential energy"—of the pendulum is chosen as $\varphi_{\rm min} = \pi$ (as above, the particle phase is defined as $\varphi = k_z\zeta$).

The phase plane plot of the pendulum-like trajectories (rigorously, it is a trace space, as we are normalizing the momentum error to the design momentum) is displayed in Fig. 4.3 for a physically realistic case, with $\beta_0^2/\gamma_0 = 0.5$ and $\alpha_{\rm rf} = 10^{-4}$. The phase plane, which should look familiar to any student of Hamiltonian mechanics, is divided into two regions: one having stable, bounded orbits (vibrational motion) and one displaying unbounded trajectories (librational motion). The boundary between the two regions referred to as the separatrix. Note that the existence of an unstable (unbounded) region in phase space is due to the nonlinearity of the applied force, as anticipated by the discussion of Section 3.2.

In the stable region, or bucket, the small amplitude motion in the neighborhood of the fixed point at $(\varphi, \delta p) = (\pi, 0)$ is nearly simple harmonic, as is discussed further below. The motion inside of the bucket at larger amplitudes is nonlinear—the oscillations slow down as the potential becomes less well approximated by a parabola, and the longitudinal focusing becomes effectively weaker. As the amplitude approaches that of the separatrix, the period of the motion becomes infinite, since on the separatrix the particles are unable to traverse the unstable fixed points at $(\phi, \delta p) = (0, 0)$, and $(\phi, \delta p) = (2\pi, 0)$. Note that the motion along the constant $\tilde{H}$ curves in this case is towards positive ζ for $\delta p > 0$, and negative ζ for $\delta p < 0$. This is needed for the vibrational orbits to exist, of course, and also points to the fact that the librational orbits always proceed in one direction in ζ above the separatrix, and another direction in the region below.

The equation for the separatrix can be obtained by evaluating the value of the Hamiltonian at an unstable fixed point, for example,

$$\tilde{H}(0,0) = 2\alpha_{\rm rf} m_0 c^2. \tag{4.30}$$

Use of Eqs (4.27) and (4.30) together yields

$$\frac{\delta p_{\rm sep}}{p_0} = \pm\frac{1}{\beta_0}\sqrt{2\alpha_{\rm rf}\gamma_0[1-\cos(k_z\zeta)]} = \pm\sqrt{\frac{4\alpha_{\rm rf}\gamma_0}{\beta_0^2}}\sin\left(\frac{k_z\zeta}{2}\right). \tag{4.31}$$

Thus the peak momentum offset encountered in the bucket (at $k_z\zeta = \pi$) is simply

$$\frac{\delta p_{\max}}{p_0} = \pm\sqrt{\frac{4\alpha_{\rm rf}\gamma_0}{\beta_0^2}}. \tag{4.32}$$

The area of the stable phase plane $(\zeta, \delta p)$ region or bucket area, $A_{\rm b}$, can be found by integrating the area between the curves of the functions given by Eq. (4.31),

$$A_{\rm b} = 4p_0\sqrt{\frac{\alpha_{\rm rf}\gamma_0}{\beta_0^2}}\int_0^{2\pi/k_z}\sin\left(\frac{k_z\zeta}{2}\right)d\zeta = \frac{16p_0}{k_z}\sqrt{\frac{\alpha_{\rm rf}\gamma_0}{\beta_0^2}}. \tag{4.33}$$

These phase plane dynamics are quite unlike those of the transverse motion, which are stable to all amplitudes under linear transformations of the type discussed in Chapter 3. Because the longitudinal motion is mediated by a sinusoidal force (instead of one linearly proportional to the offset from a stable fixed point), it is inherently nonlinear. In fact, since the force is periodic, we observe unstable fixed points one half of a wavelength away from the stable fixed points. The existence of both types of fixed points implies that only a finite region of the phase plane about the stable fixed point has vibrational orbits.

Even though this large amplitude motion (with its nonlinear characteristics) is unfamiliar, the small amplitude motion about the stable fixed point is quite familiar. If we expand the Hamiltonian near this point, we have

$$\tilde{H}(\zeta, \delta p) \cong (m_0c^2)\left[\frac{\beta_0^2}{2\gamma_0 p_0^2}(\delta p^2) + \frac{\alpha_{\rm rf}(k_z\delta\zeta)^2}{2}\right], \tag{4.34}$$

where $\delta\zeta = \zeta - \pi/k_z$. This small amplitude Hamiltonian can be used to obtain the equations of motion for $\delta\zeta$ and δp, which can be combined to give a single simple harmonic oscillator equation, that is,

$$\delta\ddot{\zeta} + \frac{\alpha_{\rm rf}(k_zc)^2}{\gamma_0^3}\delta\zeta = 0. \tag{4.35}$$

Equation 4.35 gives solutions, termed synchrotron oscillations, that are harmonic with the synchrotron frequency

$$\omega_{\rm s} = k_zc\sqrt{\frac{\alpha_{\rm rf}}{\gamma_0^3}} = \sqrt{\frac{\alpha_{\rm rf}}{\gamma_0^3}}\frac{\omega}{\beta_0}. \tag{4.36}$$

As the ratio inside of the square-root sign on the right-hand side of Eq. (4.36) is much smaller than one, the synchrotron frequency is much smaller than the frequency of the wave ω. Note that we have used time as the independent variable in writing Eq. (4.35), as opposed to the distance along the design trajectory to allow simple comparison between the rf and synchrotron frequencies.

4.5 Adiabatic capture

The phase plane plot displayed in Fig. 4.3 only tells part of the story, as it maps the trajectories of particles in the phase plane given a constant value of $\tilde{H}$.