PHYSICS FOR MATHEMATICIANS

MECHANICS I

PHYSICS FOR MATHEMATICIANS

MECHANICS I

MICHAEL SPIVAK

PUBLISH OR PERISH, INC.

Publish or Perish, Inc.
www.mathpop.com

ISBN 0-914098-32-2
EAN 978-0-914098-32-4

Printed in the United States of America

For my

Aunt Frieda

PREFACE

The purpose of this book, or possibly series of books, is indicated precisely by the title *Physics for Mathematicians*. It is only necessary for me to explain what I mean by a mathematician, and what I mean by physics.

By a mathematician I mean some one who has been trained in modern mathematics and been inculcated with its general outlook. No specific mathematical knowledge is expected, but for the purposes of this book on mechanics the material in *A Comprehensive Introduction to Differential Geometry*, Volumes 1 and 2, will generally be regarded as a prerequisite, not simply because I wrote this book, but because many of the concepts of mechanics are, in fact, best expressed in terms of basic differential geometric concepts. This will always be referred to as DG, rather than Spivak [2], which is how it appears in the bibliography.

And by physics I mean ... well, physics, what physicists mean by physics, i.e., the actual study of physical objects, even wheels, weights, ropes and pulleys (rather than the study of symplectic structures on cotangent bundles, for example). In addition to presenting the advanced physics, which mathematicians find so easy, I also want to explore the workings of elementary physics, and the mysterious maneuvers—which physicists seem to find so natural—by which one reduces a complicated physical problem to a simple mathematical question, which I have always found so hard to fathom.

As these remarks probably reveal, basically I have written this work in order to learn the subject myself, in a form that I find comprehensible. And readers familiar with some of my previous books probably realize that this has pretty much been the reason for those works also. I have been fortunate in being able to make a livelihood of sorts in this way, by indulging my desire to learn things in my own peculiar fashion while providing others with an account of the adventure. Perhaps this travelogue of an innocent abroad in a very different field will also turn out to be a book that mathematicians will enjoy (though physicists probably will not).

I am greatly indebted to many people and institutions for their help with this project. Richard Palais was, as always, an ever helpful and enthusiastic supporter of the project. Besides his help with mathematical questions, some discussions with him helped me enormously in understanding and formulating certain basic principles, though I hasten to add that he is not responsible for any heretical ideas that might appear here. Ted Shifin likewise provided unstinting help, as well as probing questions. Larry Jackal gently steered me away from some stupid mistakes and over-simplifications, and Mitch Baker heroically undertook a thorough examination of the first draft of Part I, resulting in many

corrections and improvements. John Milnor greatly contributed to my understanding of one vexing topic, and among other helpful people whom I have pestered I should mention Eisso Atzema, Robert Bryant, James Casey, Carmen Chicone, Poul Hjorth, Yildirim Hurmuzlu, Hermann Karcher, Tom Lehrer (as I couldn't plagiarize), David Nadler, Anders Persson, John Polking, and Olivier Thill.

I am grateful to the mathematics department at Rice University for affording me privileges allowing me to use the Rice University Library, and to the helpful librarians there, in particular, Erin McAfee and the science reference librarian Debra Kolah. Through the efforts of Martin Guest of Tokyo Metropolitan University and Yoshiaki Maeda of Keio University, I was able to give a series of lectures at Keio University on the material of Part I, the first time I had the opportunity to present some of this material to a live audience. A written version of the lectures was made available on the web, which providentially allowed me to be contacted by a fellow mathematician, Bruce Pourciau, who had studied many of the same questions that I puzzled through concerning Newton's work; many of his papers, listed in the Bibliography, provide additional details for the discussions of Chapters 1 and 2.

Looking 

ahead

Like every self-indulgent author, I like the think that you will try to solve all the problems ... or at least glance at them! Some problems will be used or referred to later on in the text, or sometimes in a future problem, and this crystal ball will alert you to look at them. The number in the crystal ball is the page where the problem is first used or mentioned. For example, page 41 is the first problem with a crystal ball.

I should also point out that the problems are provided mainly to help in understanding basic points, or to mention additional topics, rather than to provide proficiency in solving physics problems, and their number decreases rather rapidly after part I.

Michael Spivak
puborperish@gmail.com

CONTENTS

PART I

THE FOUNDATIONS OF MECHANICS

PROLOGUE

πᾶ βῶ καὶ κινῶ τὴν γῆν
a place to stand on ! and I will move the Earth
— Archimedes

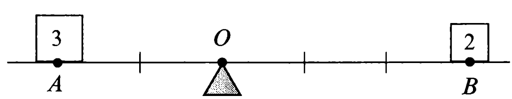
Archimedes' statement about the lever has come down to us in several forms, of which only the one appearing here is customarily translated with an exclamation point. This punctuation might even seem unnecessarily dramatic for modern tastes, because the lever no longer incites much wonder in us, so familiar are we with its principles. Yet who can forget the amazement of a child balancing an adult on a see-saw, simply by being placed at the right position. How could this be? Where did all that extra force come from!?

The only wonder nowadays is that a physics student is unlikely to produce a satisfactory answer to this question. Perhaps we will be offered a few mumblings about moments, force times distance, laws of the lever ... perhaps even the "principle of virtual work". But we probably won't get an answer that seems to explain where that extra force comes from; and it is highly unlikely that we will get an answer that begins by establishing principles about rigid bodies, even though the rigidity of the lever is an absolute necessity for it to work.

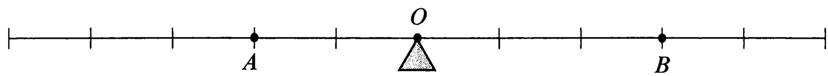
In fact, the whole path from Newton's Laws, which basically concern "point masses", to bodies whose shape and extent are significant, is often rather dubiously traversed, even though elementary physics courses blithely pose such problems of the most diverse sorts.

Our progress in explaining the lever can be used as a measure of how far we have managed to bridge this divide, an appealing test of whether our description of modern mechanics manages to account for what is generally considered one of the first mechanical principals ever discovered.

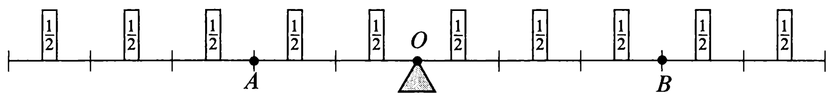
Archimedes, of course, didn't simply enunciate the law of the lever, but as a true mathematical theoretician, he devised a *proof*. The crux of Archimedes' proof can be illustrated by the particular case where we have two objects,



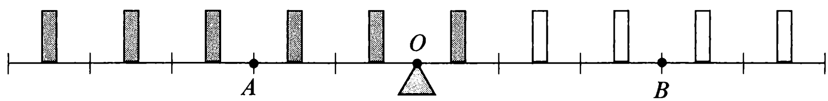
weighing 3 units and 2 units, respectively, positioned at points A and B whose distances from the fulcrum O are in the ratio 2 to 3, so that some unit distance fits twice into the segment AO and three times into the segment OB. To create a symmetrical situation, we begin by laying off 3 additional units, or the length of OB, to the left of A, and 2 additional units, or the length of AO, to the right of B, for a total of 10 units.



Then we take a weight of $\frac{1}{2}$ and place 10 copies at the centers of each of our units. This arrangement is clearly balanced, because it is completely symmet-



rical to the left and right of O. On the other hand, the 6 left-most weights (shaded below) have a total weight of 6, and their center of gravity is exactly



at A, so they have the same effect as our original weight of 3 units at A. Similarly, the other 4 weights have a total weight of 2, and their center of gravity is exactly at B, so they have the same effect as our original weight of 2 units at B. And thus it follows that the original weights at A and B will balance!

Archimedes doesn't bother to illustrate the reasoning with a particular case, but simply provides the argument for the general case of commensurate distances and weights, which doesn't require any more generality, but does make the argument quite a bit harder to follow.

Moreover, Archimedes then uses the "method of exhaustion" (the Greek geometric way of using the density of the rationals) to extend the result to the case of incommensurable distances. This might seem like over-kill for a proposition about physical weights and distances, which can only be measured to within a certain accuracy, but then Archimedes was really a mathematician!

Archimedes' proof has been criticized, and emendments to it have been proffered, for centuries, culminating in a detailed examination in Mach [1], a book that has been enormously influential not only among historians and philosophers of science, but even among scientists themselves.

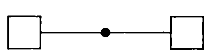
Mach's critique of Archimedes' analysis of the lever dismisses the whole enterprise as a symptom of the "Grecian mania for demonstration". He presents two axioms of Archimedes, which may be translated as follows:

1) Equal weights suspended at equal distances from a fulcrum are in equilibrium.
2) Equal weights suspended at unequal distances cannot be in equilibrium. The lever will be inclined towards the weight at the greater distance.

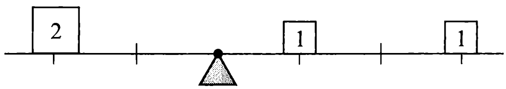
Mach points out that at best these axiom say that the effect of a weight W at distance L from the fulcrum has a dependence of the form $W \cdot f(L)$ [for a monotonic f, though Mach doesn't mention that]. How could the particular form $W \cdot L$ possibly be inferred from this? Instead, the replacement of one weight by several smaller weights implicitly assumes that the relationship is of this form.

Despite the apparently irrefutable logic of this assault, it must be said that Archimedes' argument certainly does seem awfully clever, and at first blush it even seems awfully convincing! So we should mention that Mach conveniently pretends that Archimedes proceeds from only two axioms, whereas Archimedes actually states eight, some of which specifically mention center of gravity, a concept which was apparently analyzed in an earlier lost work of Archimedes. In fact, Archimedes' proof of the law of the lever forms the first part of a work entitled *On the equilibrium of planes or centers of gravity of plane figures*, in which Archimedes finds the centers of gravity of triangles, rectangles, trapezoids, and finally, in a *bravura* calculation, the center of gravity of a section of a parabola.

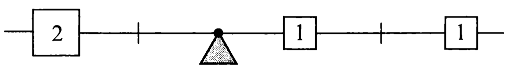
The details of Archimedes proof of the law of the lever may be found in Dugas [1; pp. 25–27]. It is clear from Archimedes' arguments that he regards the center of gravity of an extended body of weight W as a point where a "point mass" of weight W would have the same effect as the extended body. Moreover, the center of gravity of two equal bodies is supposed to be midway between them (presumably this is meant to apply only in the case where the two



bodies are connected, say by a rigid rod of negligible weight). So, for example, Archimedes would claim that the following weights are balanced because the two bodies of weight 1 act the same as a single body of weight 2 situated midway between them. One might object that in this case the two smaller weights



aren't connected by a rigid rod, but of course we are assuming that the weights don't slide along the lever, so that it might be better to think of the weights as attached to the lever. Notice that Archimedes' evaluation of the situation



actually amounts to some sort of assertion about rigid bodies—something about how the rod combines the effect of the two end weights.

I wouldn't want to attempt to defend Archimedes' proof too earnestly,[1] since our aim, after all, is to show in the next few chapters how the law of the lever arises as part of a systematic investigation of rigid bodies. But Archimedes' analysis serves as a good starting point from which we can jump ahead a couple of millennia, to Newton.

[1] For an extended discussion of Archimedes' proof along similar lines, together with further references, see Dijksterhuis [1]. Also see van der Waerden [1] for a critique of Mach's critique.

CHAPTER 1

NEWTONIAN MECHANICS

The terms classical mechanics and Newtonian mechanics are used virtually synonymously, attesting to the fact that all of classical mechanics flows from Newton [2], *Philosophiæ Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy), or simply The Principia.

It should be said at the outset that if you are trying to learn mechanics, the Principia is *not* the place to start! It has all the inherent difficulty and obscurity of any classic, with numerous approaches to fundamental considerations that are nowadays treated in quite different ways. But there is still considerable interest in examining parts of the Principia—after all, "Newton's Laws", with which all discussions of classical mechanics begin, come from this source, and there are some pleasant mathematical surprises awaiting us as well.

All our quotations from the Principia are taken from

> Newton, *The Principia.* Translated by Bernard Cohen and Anne Whitman, University of California Press, 1999.

This recent translation eliminates most obsolete terms, carefully explains any remaining old-fashioned terminology, and corrects various errors in the older translations. Moreover, it begins with an extensive Guide to reading the Principia that is almost as long as the Principia itself.

We will essentially cover a nowhere dense subset of the Principia, and another recent book, on which I relied for many topics, will be extremely useful to any modern reader interested in exploring the enormous amount of material of diverse sorts that it contains:

> S. Chandrasekhar [2] *Newton's Principia for the Common Reader*, Clarendon Press, 1995.

This book provides detailed modern mathematical arguments (though not always ones that correspond directly to Newton's proofs) for a large portion of the work, together with many intriguing conjectures about the course of Newton's thought.

In the manner of a preëmptive strike, I would like to say that the difficulty of reading the Principia is equaled only by the danger of commenting upon it. The work has been mined so thoroughly by experts and scholars, and its elaborate historical roots have been examined so carefully, that any amateur venture is sure to produce some remark or interpretation that will be met with scorn by the cognoscenti.

Nevertheless, we will be happily unconcerned while extracting and interpreting parts of the Principia in terms of modern physics. Several historical remarks about the Principia may be found sprinkled throughout Part I.

In the manner of Euclid's *Elements*, which it emulates, the Principia begins with two preliminary sections, "Definitions" and "Axioms, or the Laws of Motion". And it would be fair to say that it shares the same virtues and defects as the Elements, the latter a work that exhibits admirable rigor only after a shaky beginning where we have to contend with definitions like a point as "that which has no part" and a line as "breadthless length", while basic principles such as the "side-angle-side theorem" are presented as theorems with indefensible proofs, rather than as additional postulates.

Reworking the Elements into a rigorous system hardly troubles modern mathematicians. We simply give a list of undefined terms like "point", "straight line", and the notion of a "point lying on a line", declare any basic results that we can't prove adequately as axioms, and carry on. Of course, we might, for example, instead prefer to declare the straight lines to be particular collections of points, so that the notion of lying on a line reduces to set membership; this just illustrates that there will always be several possible ways of dealing with any collection of "undefined terms" that are supposed to have connections between them.

For modern physicists, the role of "undefined" terms may be replaced by the idea of "operationally defined" terms: a concept is defined once we explain how to measure it. However, as we shall see, even this idea can present some difficulties and subtleties.

Basic Concepts

Mass and force. In mechanics, the first basic concept is that of the **mass** of an object. Newton initially calls this the "quantity of matter" in the Principia, which immediately commences with Definition 1:

> *Quantity of matter is a measure of matter that arises from its density and volume jointly.*

Great acumen is hardly needed to realize that this definition is hopelessly circular, since density is normally defined as the ratio of mass to volume, but Newton's unhelpful phrase does have some implicit implications. For example, we expect the mass of an object to remain unchanged if we change its shape, that the mass of a quantity of water remains the same after it has been frozen into a piece of ice, or even reconstituted as powdery snow, and the mass of a quantity of air or other gas remains the same if we confine it to a smaller region.

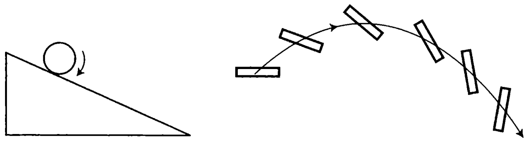
But our modern conception of mass is better reflected in what Newton soon afterwards refers to as the "inertia of the mass": a body's resistance to being moved if it is at rest—or of having its velocity changed if it is already moving with a uniform velocity.

For example, consider two solid balls, say about 3 feet in diameter, one made of cork, the other of iron, resting on a smooth floor. In order to get the cork sphere rolling at a speed of about 5 feet per second, we would have to push on it a bit, or, in everyday terms, exert a small force on it. But in order to get the iron sphere rolling at that speed we would have to push much harder, or, in everyday terms, exert a much large force on it. It is these different experiences that lead us to regard the iron ball as having much greater mass than the cork ball.

Thus, roughly speaking, the mass of an object is measured by its resistance to being moved by a force, and of course **force** is the other basic concept of mechanics, and we simply have to assume that we have some idea of what we mean when we say that a magnet exerts a force on a piece of iron, or that gravity exerts a force on any object, or that a force must be applied to the end of a spring to stretch it out, and the spring exerts a force pulling the stretched end back towards its original point. The force that a person must exert for some purpose, like moving a cork or iron ball, is the one that connects most directly with our everyday experiences, though it is naturally the one that we might be hardest pressed to quantify accurately.

Implicit in this discussion, by the way, is the important idea that forces have not only a magnitude, but also a *direction*; the force of gravity is exerted toward the center of the earth, and the force exerted by a spring is directed along the axis of the spring. Thus it is natural to represent forces mathematically by vectors, i.e., by elements of $\mathbb{R}^3$, although we often picture them as arrows.

Newton makes one other important definition, which seems much less problematical: "Quantity of motion", or what we now call the **momentum** of an object, is simply the product $m \cdot \mathbf{v}$ of its mass m and its velocity vector $\mathbf{v}$. But even this definition hides difficulties, because it assumes that the motion of an object can be described in terms of a curve $c\colon \mathbb{R} \to \mathbb{R}^3$, with velocity vector $\mathbf{v} = c'$, which is certainly not the case for a ball rolling down an inclined plane, or a rod revolving as it is thrown. In essence, we are constricting our initial considerations to "particles ", or "point masses", abstractions that don't



actually exist, but that represent reasonable first approximations in appropriate cases—for example, in discussing the motion of the planets around the sun. Mathematically, a particle is just a path $c\colon \mathbb{R} \to \mathbb{R}^3$, with derivative $c' = \mathbf{v}$, together with a number $m > 0 \in \mathbb{R}$.

Newton doesn't offer any operational definition of mass, or of force, leaving us with these rather vague intuitive concepts. He ends the "Definitions" section with a Scholium (commentary) rather longer than the definitions themselves, but it sheds no light on this matter, instead treating other topics that we will be in a better position to consider after we examine the material that appears in the next section of the Principia.

Newton's Three Laws

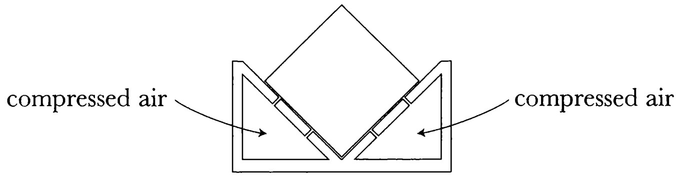
The first law. The second section of the Principia, "Axioms, or the Laws of Motion", begins immediately with the statement of Law 1:

> *Every body perseveres in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by forces impressed.*

Nowadays we might state this as follows:

> An object not acted upon by any force has a constant velocity $\mathbf{v}$, and, in particular, if it is initially at rest, then it remains at rest.

Newton, of course, did not call his law "Newton's First Law", and in the lengthy Scholium to the second section he describes it as one of the principles "accepted by mathematicians and confirmed by experiments of many kinds", explicitly mentioning Galileo, who was the first to enunciate it. Galileo's contribution was not so much an experimental "verification" of this law as an accumulation of experiments and reasonings to explain why everyday experiences seem to contradict it. Nowadays we can illustrate the law rather dramatically by sliding an object along a glass table with dry ice evaporating from it, forming a cushion of gas that practically eliminates friction. Or we can slide objects along an "air trough", a track with compressed air blown through holes in its sides, so that our object is sliding along a thin layer of air (see Neher and Leighton [1]).



side view of a block sliding on an air trough

In addition to the leap of imagination required by the first law, further clarification is required because the notion of position, and thus of velocity, depends on the coordinate system used by the observer. Newton's first law essentially distinguishes certain spatial coordinate systems, like those set up in a stationary room, from others, like the coordinate system used by some one in an accelerating train, or by an observer on a stone being swung in an arc by a sling shot. It basically states that there are certain coordinate systems in which $\mathbf{v}$ is constant for any body not acted upon by forces. Such coordinate systems are often called "inertial systems", because they exhibit a body's *inertia*, its tendency to stay at rest or in uniform motion unless acted upon by some force. Thus, the first law might be stated more completely as

> Newton's First Law: There is at least one coordinate system—an *inertial* system—in which any object not acted upon by any forces has constant velocity.

Often, the first law is simply referred to as the *law of inertia.*

Of course, all sorts of philosophical problems might arise if we enquired too closely into the distinction between definitions and observational facts that the first law entails—a danger with any axiomatically developed system—but two specific points should be made.

First, it is clear that any coordinate system moving with a uniform velocity with respect to an inertial system is itself an inertial system. On the other hand, determining even one inertial system might be a little more difficult than expected. Although a coordinate system set up in a stationary room acts as an inertial system for various earthbound experiments, it clearly isn't really an inertial system, since it is not only rotating in a 24 hour period around an axis, but also rotating in an annual period around the sun. A much better approximation to an inertial system is one based on the "fixed stars", although we now know that the stars aren't all that fixed either!

In Newton's time the "fixed stars" (i.e., the heavenly bodies that were not planets or comets) were indeed thought to be at rest with respect to each other, but in any case Newton would simply have regarded a reference frame based on the fixed stars as an excellent experimental approximation to the "absolute space" that he refers to in the Scholium at the end of the "Definitions" section, this "absolute space" presumably being the inertial system that is "really at rest", rather than one moving with some uniform velocity with respect to it.

Newton's Scholium (Newton [2; pp. 408–415]) includes a long discussion of the distinction between absolute and relative motion, ideas that have led, especially with the advent of relativity theory, to considerable fundamental physical/philosophical questions. Many elementary mechanics texts now include a discussion of these matters, motivated by the impulse to introduce ideas of modern physics as rapidly as possible. On the whole, we will allow them to fester until the proper time, in another volume, although Chapter 10 and some remarks in Chapter 7 touch on some of these topics.

The second law. Newton also credits Galileo with the next law, which may be formulated in modern terminology as

> Newton's Second Law: In an inertial system, the rate of change of momentum of a particle is directly proportional to the force $\mathbf{F}$ acting on it:
>
> $$\begin{aligned}\mathbf{F} &= (m\mathbf{v})' \\ &= m \cdot \mathbf{v}'.\end{aligned}$$

Here we have taken the constant of proportionality simply to be 1, since this just amounts to choosing a unit of force once units for mass and length and time have been determined.

Note, by the way, that our final version of the first law emphasizes that it cannot be regarded simply as the special case $\mathbf{F} = 0$ of the second law, since it is the first law that defines an inertial system.

In the Principia, where concepts of calculus are eschewed as much as possible at the beginning, Newton speaks merely of the *change* of momentum, rather than its derivative, so the second law seems to be stated in terms of "impulsive forces" that act "instantaneously"—a hockey stick hitting a puck might be a good approximation to this ideal—though Newton certainly uses the second law in the more general sense whenever he needs it:

> *A change in motion [momentum] is proportional to the motive force impressed and takes place along the straight line in which that force is impressed.*

At first sight, it's hard to imagine how the second law could play a central role, or indeed any role, in the foundations of mechanics. On the left side of the equation

$$\mathbf{F} = m \cdot \mathbf{v}'$$

we have the quantity **F** that we have discussed only in vague terms, without determining how to measure it, and on the right side of the equation we have the quantity m that has also been discussed only in vague terms, again without determining how to measure it. The equation $\mathbf{F} = m \cdot \mathbf{v}'$ might well be called an "axiom", in the purely mathematical sense of the word, but why would we think of it as a "law of motion", presumably something that we could check by experiment!?

Nevertheless, the situation is not quite so hopelessly muddled as it seems. In fact, in his Scholium to the second section, Newton specifically cites Galileo in regard to the second law because of Galileo's observations "that the descent of heavy bodies is in the squared ratio of the time". Newton proceeds to explain that this is a consequence of the constant acceleration of gravity, working from the impulsive case to the continuous case:

> When a body falls, uniform gravity, by acting equally in individual equal particles of time, impresses equal forces upon that body and generates equal velocities; and in the total time it impresses a total force and generates a total velocity proportional to the time. And the spaces described in proportional times are as the velocities and the times jointly, that is, in the squared ratio of the times.

This is the sort of explanation that makes you glad that you aren't trying to learn mathematics in the 17th century! Here, perhaps, is what Newton seems to be saying. Suppose we divide the time t of descent into small intervals of length $\Delta t = t/N$ for a large number N, and regard the motion with constant acceleration a as being uniform on each interval, with an "instantaneous" change of speed of $a \cdot \Delta t$ at the beginning of each interval. Our body, starting at rest, has velocity $a \cdot \Delta t$ during the first time interval of length Δt, falling through a distance of $a(\Delta t)^2$; it has the velocity $2a \cdot \Delta t$ during the next time interval of length t, falling through a distance of $2a(\Delta t)^2, \dots$. Thus, at the end of time t it has fallen a distance of $(1 + 2 + 3 + \cdots + N) \cdot a(\Delta t)^2 = \frac{1}{2}aN(N+1)(\Delta t)^2$. Since $N\Delta t = t$, this is close to $\frac{1}{2}at^2$.

Nowadays, of course, we just say

> If $s'' = a$ for a constant a, then $s' = at$, and thus $s = \frac{1}{2}at^2$.

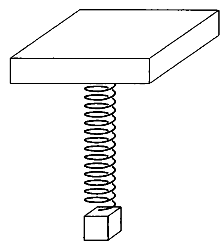
Even after all this explanation, we seem to be concerned with a purely mathematical result about second derivatives. So it's important to go back and look at one particular phrase in Newton's argument:

When a body falls, *uniform gravity*, by acting ...

How do we know that gravity *is* uniform?

Note that here we are not considering the question famously associated with Galileo concerning the effect of gravity on two different objects. We have only one object, and are asking how we know that the force of gravity on the object is constant throughout its descent, i.e., independent of its height. (Of course, that isn't actually true, since the force varies inversely as the square of its distance from the center of the earth, but the change is insignificant for any distances close to the radius of the earth.)

We can exhibit the effect of the force of gravity by a measurement with a primitive scale consisting of a spring attached to an upper support which is kept fixed, say by nailing it to a wall, with our object attached to the bottom, pulling the spring down. Experience tells us that if we nail the upper support higher up on the wall, the spring still stretches to the same length; if we try the experiment on the third floor of the laboratory, the length is still the same; if we carry out the experiment at the top of the leaning tower of Pisa, the length is still the same. Thus we see that the downward force of gravity on a particular object doesn't depend on the height of our object above the earth's surface—we can use our primitive scale as a way of ascertaining that two forces are *equal* without having to worry about the question of just how we should assign a magnitude to forces.

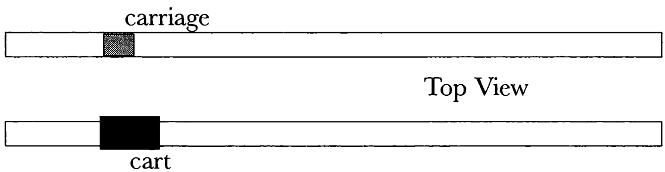


Note that this experiment doesn't involve the acceleration our object would have if we allowed it to fall, it simply ascertains that the *force* of gravity is the same throughout the downward path of the object when it does fall. Once we've determined that the force of gravity is constant, we note that the assumption $\mathbf{F} = mc''$ amounts to asserting that the downward *acceleration* is constant, and the latter assertion is independently verified by the observed result that the distance traveled by the body at time t is proportional to t^2.

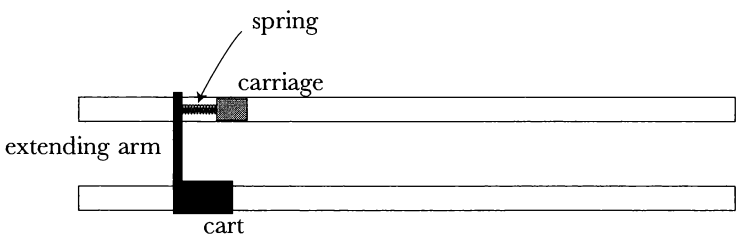
Of course, Newton and others didn't perform preliminary experiments of the exact sort we have outlined, but innumerable weighings had been recorded at different heights without any mysterious discrepancies; in more prosaic terms, the force needed to raise an object seems to be the same whether it is resting on the top floor of a tower or on the ground floor.

Now that we've seen that the second law does have some significance, we can try to concoct an operational definition of mass. We'll begin with a definition that is conceptually very straightforward, although it would certainly be rather awkward to use in practice.

First we want to have a very long air-trough, with a carriage, of negligible mass, in which we can place a body whose mass we want to measure. Parallel to this air-trough we have a track with a little cart that can be pulled along the track with any desired acceleration a; for simplicity, let's imagine that we merely have to turn a knob on an instrument panel to vary a, without worrying about the clever mechanism that would be required to produce this effect. Of course, our track will have to be very long if we expect to pull the cart with constant acceleration a for any reasonable amount of time.



We attach an extending arm to our cart that can be placed behind the carriage on the air-trough so that the carriage is moved in tandem with the cart. But instead of placing this arm directly behind the carriage, we will put a nice strong spring between them.



Now let's choose a particular body B_0 that we want to be our "unit mass", so that we will assign it mass $m = 1$. We place this body in the carriage and pull our cart with some convenient constant acceleration a_0. Initially, of course, the carriage will not move with the same acceleration, because the spring will compress somewhat, so that the carriage won't move exactly in tandem with the cart. But we very quickly reach the point where the spring is no longer compressing, or at any rate the length of the spring is constant within the limits of accuracy of our measurements. We carefully measure this final length, and call it L_0.

Note, by the way, that this whole set-up is dependent upon our original experimental observation that equal forces produce equal acceleration: the compression of the spring measures the force that is being applied to the carriage, and once the carriage is moving with a constant acceleration, the force applied to it must be constant.

Now let's take some other body B whose mass m we want to determine. We place B in the carriage instead of B_0 and once again pull our cart with constant acceleration a_0, and observe the final length of the compressed spring. It probably isn't L_0 any more, so we try adjusting the acceleration a in order to make it become L_0: if the spring was compressed more, to a length $< L_0$, we try an acceleration $< a_0$, if it was compressed to a length $> L_0$ we try an acceleration $> a_0$. After lots of trial and error, we finally find an acceleration a_1 which compresses the spring to exactly the length L_0. We now define

$$\text{mass } m \text{ of } B = a_0/a_1.$$

This definition makes the law $\mathbf{F} = m\mathbf{a}$ work for any particular fixed $\mathbf{F}$, and much experimentation would show that it works just as well for any other $\mathbf{F}$; in other words, if we repeated this whole process using a different spring, and thus a different L_0, we would still end up assigning the same masses to all bodies. Then, of course, we can use the equation in reverse, as a way of measuring force, by seeing what acceleration is produced on a body of some known mass m.

A few subtle points ought to be mentioned here. First of all, although our original little experiment with the spring is certainly consistent with $\mathbf{F} = m\mathbf{a}$, it would hardly seem to be very conclusive. After all, how do we know that the correct law isn't something like $\mathbf{F} = m\mathbf{a} + k\mathbf{a}'$ for some constant k, so that third derivatives, or even higher derivatives, are involved? I don't know of any experiments to directly test this, but there is an enormous body of experience that attests to it: the force of gravity isn't constant over large distances, so all the calculations that keep satellites in motion, guide space ships to the moon and land them, etc., present a great deal of evidence. Newton provided the most important evidence of all, by showing that this law, together with the inverse square law for the force of gravity accounts for the elliptical orbits of the planets. (You might think that the argument is somewhat circular, for Newton essentially postulated an inverse square law, rather than relying on numerous experiments to verify it. However, Newton did test his law in one very important case, the motion of the moon around the earth. As a matter of fact, Newton initially shelved his whole theory of gravity because this case didn't agree with measurements; later on, more accurate measurements of the radius of the earth then confirmed the idea—see Problem 2-2.)

A second point involves the "additivity" of mass: if an object is made up of two parts, we expect its original mass to be the sum of the masses of the two new pieces, an idea inherent in the very notion of mass as "quantity of matter".

But this hardly seems clear with our nice precise operational definition! If have two objects of masses m_1 and m_2, our operational definition means that a_1 is the acceleration that the first body must be subjected to in order to compress the spring to length L_0, while a_2 is the acceleration that the second body must be subjected to in order to compress the spring the same amount. If we join the two objects together by placing them together in the carriage on our air-trough, then the new object should have mass $m_1 + m_2$, which means, according to our definition, that to obtain the same compression for the two objects together, they must be subject to an acceleration a satisfying

$$\frac{1}{a} = \frac{1}{a_1} + \frac{1}{a_2} \qquad \text{or} \qquad a = \frac{a_1 a_2}{a_1 + a_2}.$$

At first glance, this might seem to be a strange fact that one might never have anticipated, one that could only be verified by a large number of experiments with varying masses, and perhaps never even divined from the experimental data obtained! Problem 25 gives a more promising approach.

Finally, there is one other point that needs to be emphasized. When we first investigated the "uniformity" of gravity (page 14), we said that "the force of gravity on the object is constant throughout its descent, i.e., independent of its height." But we should also have said that it is *independent of its velocity*; this was implicit in our mathematical description of the problem by the equation $s'' = a$.

It is worth pointing this out, not only because it was clearly implicit in Newton's statement of the second law, a subtlety obviously clear to Newton, but also because it is *false*, a subtlety that was *not* known to Newton, or to any one else until 1905, when Einstein discovered it by purely theoretical means.

If we were to measure the acceleration produced by a force of magnitude F on an object moving with varying speeds v, including speeds v close to the speed of light c, we would find that if $F = m_0 v'(0)$ when the object starts at rest, then when the speed is v we have

$$F = \frac{m_0}{(1 - v^2/c^2)^{3/2}} v'.$$

To be sure, no one ever states the result this way. Instead of saying that the same force produces a *smaller acceleration* on an object when it is moving faster, physicists always say that a moving object has a *larger mass*. You might think they would say that the mass of of a body moving with speed v is $(1 - v^2/c^2)^{-3/2}$ times its mass at rest, but they don't say that either (see Problem 8).

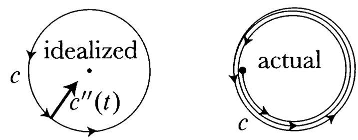
Leaving aside this relativistic subtlety, which won't pop up again anywhere in this volume, we simply want to mention a less direct, but much more convenient, operational definition of mass, based on the mathematician's and physicist's common view that a straight line is just a circle of infinite radius.

Instead of using an air-trough, we simply attach our body B to the end of a very stiff spring that is being rotated horizontally with some large constant "angular frequency" α, so that B moves along the circle

$$c(t) = R(\cos \alpha t, \sin \alpha t)$$

for some radius R. This radius R will be somewhat larger than the unstretched length of the spring, because B actually begins moving along a spiral, pulling the spring out, though its path soon becomes indistinguishable from a circle. For the acceleration we simply have

(A) $$c''(t) = -R\alpha^2(\cos \alpha t, \sin \alpha t),$$



so that the acceleration always points directly inward, and has magnitude $R\alpha^2$. This means that the force $\mathbf{F}$ that the spring exerts on B also always points directly inward and has constant magnitude.

Declaring B to be our unit of mass amounts to saying that

(a) $$|\mathbf{F}| = 1 \cdot R\alpha^2.$$

To determine the mass m of any other body, we attach it to the end of our spring and vary the angular frequency with which we rotate it until we arrive at an angular frequency β for which our body is moving along a circle

$$c(t) = R(\cos \beta t, \sin \beta t)$$

of the *same* radius R. Now we should have

(b) $$|\mathbf{F}| = m \cdot R\beta^2,$$

with the $|\mathbf{F}|$ in equations (a) and (b) having the *same* value, since in both cases the spring has been stretched by the same amount. In other words, we can determine m by

$$m = \alpha^2/\beta^2.$$

We've ignored the effect of gravity on these bodies, but that would become negligible in comparison to the force of our stiff spring when α is large, or we might imagine the measurements being made in outer space.

I'm sure that the basic mechanism for this definition could be greatly refined. Instead of a spring, one might whirl a tube filled with mercury, and measure the compression of this mercury column, etc. But I don't think any one has ever actually produced a mechanism of this sort. In fact, as far as I know, no one has ever measured the mass of *anything* accurately. This statement obviously requires a bit of explanation!

Mass and weight are different... The downward force of gravity that Newton refers to in reference to the second law and Galileo's observations of falling bodies is, of course, what we usually call the *weight* of the body.

Neophyte physicists are always warned not to confuse mass, a measure of an object's resistance to being moved, and weight, a measure of the gravitational force exerted on it by the earth. Nowadays we might emphasize that distinction by pointing out that the mass of a body on the moon would still be the same even though its "weight"—the gravitational force exerted on it by the moon—would be much less. Even more dramatically, we could consider a space ship cruising at a constant speed through space, far from any planets; in this space ship, a large iron ball and a large cork ball would both float effortlessly without any support from the floor(s), but *moving* the iron ball would be much harder than moving the cork ball! If the two balls were painted the same color, one can imagine all sorts of unpleasant practical jokes that might be perpetrated on a naive space cadet.

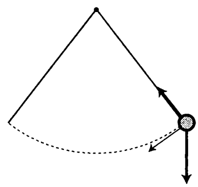
... **yet not so different.** On the other hand, even after comprehending the distinction, we might very well suspect that the *relative* masses of two bodies are the same as their relative weights.

On a crude scale, this idea certainly conforms to our experience. For example, knowing that the weight of our iron ball—the upward force that we would need to exert to keep it from falling—is much greater than the weight of our cork ball, we are hardly surprised that the force required to set the iron ball rolling at a certain speed is much greater than the force required for the cork ball, and we shouldn't be surprised that the same holds true in the cruising space ship.

In terms of the second law, however, we can make a much more specific correlation: Since the weight of an object, of mass m, say, is the force $\mathbf{F}$ of gravity on it, the law $\mathbf{F} = m \cdot \mathbf{v}'$ means that the ratio $\mathbf{F}/m$ of weight to mass is simply the acceleration that an object undergoes under free fall. Thus, proportionality of weight to mass is equivalent to the assertion that all bodies fall with the same acceleration, the famous fact usually attributed to Galileo.

Incidentally, even before Galileo's experiments, J.-B. Benedetti (1530-1590) had argued that different size bodies of the same material must fall with the same velocity, since a large body could simply be considered as two smaller bodies side by side (compare the arguments on page 27). But Benedetti still thought that denser bodies would fall more rapidly than less dense ones, and Galileo is usually credited with being the first, or at least one of the first, to assert the equal acceleration of bodies with differing compositions, like wood and iron—or, more to the taste of a modern physicist, like aluminum and gold, which have such different proportions of protons and neutrons.

Although Galileo's experiments—real or mythical—may have been the first investigation of this fact, it can be verified with much greater accuracy by resorting once again to the trick of replacing linear motion by circular motion, in this case by considering a pendulum. Problem 17 talks about the pendulum in



a little detail, but basically the pendulum bob moves in a circle about the pivot point because the downward gravitational force on the pendulum bob is partially offset by a force directed along the string that is just large enough to result in a motion tangent to the circle. Instead of trying to compare the downward path of two falling objects, we can instead use them as the bobs for two pendulums of the same length, and see if they move at the same rate—something that one can verify with great accuracy by letting the pendulums swing through many cycles. Addendum 1B discusses experiments of this sort in much greater detail.

Once we've found that weight is strictly proportional to mass, it becomes unnecessary to measure mass, i.e., to find the ratio of masses of different objects, because this ratio is just the same as the ratio of their weights. Even though weight may vary with location, the ratios of weights at an given location give the ratios of the corresponding masses.

Naturally, the proportionality of weight to mass only adds to the usual confusion between these two concepts. In fact, before Newton, the distinction had hardly ever been made, and the weight of an object was generally regarded as an invariable property of the object. It's not surprising that hardly any thought was given to the idea that the weight of an object might vary with its distance from the earth, since any such variation would have been extremely difficult to observe directly.

Indeed, it was indirect measurements that accidently provided the first evidence that the weight of a body changes at different distances from the earth. In 1672, the astronomer Jean Richer, making observations at Cayenne (latitude 5° north), found that his pendulum clock, which kept perfect time in Paris, was going, compared to the mean motion of the sun, more slowly by 2 minutes and 28 seconds per day, though he couldn't explain the difference; and similar observations were made later by others, including Halley. Newton interpreted these observations as an indication that the earth bulged near the equator, because of its rotation, so that locations near the equator were further from the center of the earth. In the third book of the Principia, Newton used this data to estimate that the equatorial radius of the earth was about 17 miles greater than its polar radius.[1] Modern estimates give about 21.4 km, or about 13 miles, a discrepancy mainly due to the higher density of the earth's core.

Separating the idea of mass and weight, deciding on mass and force as the basic concepts in terms of which others should be defined, and choosing the first two laws as the basis for deducing other results, was one of Newton's main achievements. The success of his choices is one of the reasons that we still speak of classical physics as Newtonian mechanics.

The third law. Strenuous exertions have allowed us to tease out a bit of meaning from the first two laws, both of which involve individual bodies, but say nothing about the interactions between different bodies. This information is given by Newton's third law, and since all of mechanics supposedly rests on Newton's three laws, this one must really be a doozy. In fact, it is usually stated as a memorable apothegm: "Every action has an equal and opposite reaction". In this form it is ideally suited to misappropriation by armchair philosophers, moral and political thinkers, and others of that ilk.

But Newton's statement was much more specific:

> Newton's Third Law:
> *To any action there is always an opposite and equal reaction; in other words, the actions of two bodies upon each other are always equal and always opposite in direction.*

Thus, if one object exerts a force $\mathbf{F}$ on a second object, then the second object exerts the force $-\mathbf{F}$ on the first. Numerous misuses and invalid analogues of the third law ignore this basic fact that the two actions in question are exerted on two different bodies.

A simple example of the third law is provided by the gravitational force $\mathbf{F}$ that the earth exerts on an object: according to the third law, the object itself exerts a force $-\mathbf{F}$ on the earth; of course, the resulting change in momentum of the massive earth is basically unnoticeable, because it involves such a small change in the earth's velocity.

Newton gives various everyday examples of the third law: "If anyone presses a stone with a finger, the finger is also pressed by the stone. If a horse draws a stone tied to a rope, the horse will (so to speak) also be drawn back equally toward the stone ... " Like all "simple" physics examples, these are actually

[1] The argument (using some material from Chapter 4), and some historical background, can be found in Chandrasekhar [2; pp. 381 ff.], or quickly set out at the beginning of the comprehensive book Chandrasekhar [1], examining the many later extentions of the method; a simple outline may also be found in Hand and Finch [1; pp. 339 ff.].

complicated phenomena that nowadays we might consider to be compounded from myriad instances of the third law applied to the atoms of which these everyday objects are composed. But for the study of mechanics, objects like billiard balls often serve as basic constituents.

As a particular instance of the third law, we consider the collision of two such objects: B_1, having mass m_1, and B_2, having mass m_2. During the collision, B_2 will be exerting a force $\mathbf{F}_{12}$ on B_1, while B_1 will be exerting a force $\mathbf{F}_{21}$ on B_2 (the first subscript indicates the body on which the force acts). We should really write $\mathbf{F}_{12}(t)$ and $\mathbf{F}_{21}(t)$ because these forces may vary with time; in fact, they presumably vary in an incredibly complicated way, depending on the particular way that the two bodies are compressed, spin, vibrate, undulate, bobble, etc. But we always have $\mathbf{F}_{12}(t) = -\mathbf{F}_{21}(t)$, so for all times t we have

$$\begin{aligned}(m_1\mathbf{v}_1 + m_2\mathbf{v}_2)'(t) &= (m_1\mathbf{v}_1)'(t) + (m_2\mathbf{v}_2)'(t)\\ &= \mathbf{F}_{12}(t) + \mathbf{F}_{21}(t)\\ &= 0.\end{aligned}$$

Thus, no matter how complicated the collision may be, the "total momentum" $m_1\mathbf{v}_1 + m_2\mathbf{v}_2$ is constant. Or, as we like to say, momentum is conserved.

Newton explicitly stated this "conservation of momentum" law as a Corollary of the laws of motion:

> *The quantity of motion, which is determined by adding the motions made in one direction and subtracting the motions made in the opposite direction, is not changed by the action of bodies on one another.*

In fact, it was the experimental verification of conservation of momentum that Newton cited as the evidence for the third law. In his Scholium he states that "Sir Christopher Wren, Dr. John Wallis, and Mr. Christiaan Huygens, easily the foremost geometers of the previous generation independently found the rules of [collisions of bodies], and communicated them to the Royal Society at nearly the same time," specifically mentioning that "Wren additionally proved the truth of these rules ... by means of an experiment with pendulums ... ." Pendulums, with their advantage of replacing linear motion with circular motion, were also the device of choice in the 18^{th} century for practically eliminating friction, with the added advantage that one could estimate the speed of the pendulum bob at the bottom of its path by noting how high the pendulum rose on the next swing. Newton himself conducted experiments with colliding pendulums carried out with special care to take into account the effects of air resistance, and gives a full account of them in the Principia (Newton [2; pp. 425–427].)

Thus, Newton took a basic experimental fact, the conservation of momentum in collisions, and recast it as a corollary of an apparently equivalent formulation, Newton's Third Law. Nowadays, physics texts may allude to the essential role played by the third law—"forces always appear in pairs"—but they give scant attention to the fact that Newton's decision to cast those experimental results in terms of the third law was an incredibly audacious generalization! Based on results involving the completely unknown repulsive forces between colliding bodies, Newton hypothesized a more general law concerning *all* forces.

And, despite the snide remark offered previously concerning non-scientists' misunderstanding of the third law, it may be said with equal justification that the significance of the third law is almost completely obscured in elementary physics classes, where it is usually regarded as no more than a tool for solving problems—"draw a force diagram, remember that if A exerts a force on B, then B exerts an equal and opposite force on A, blah, blah, blah, ... "

In this regard, it is instructive to consider a little experiment measuring the attraction of a magnet and a piece of iron. In the 18th century manner, we might employ the magnet and iron as the bobs of two pendulums, but let's take advantage of modern technology and simply place them on a surface with dry ice, or on an air-trough. We start by holding the magnet and iron fixed—so that the total momentum is 0—and then release them at the same moment. As soon as we do this, the iron starts moving towards the magnet (the magnet attracts the iron), and likewise the magnet starts moving in the opposite direction, toward the iron (the iron attracts the magnet). With modern technology it is easy to make accurate measurements of their positions over very small intervals of time, and thereby determine their velocities at these times, giving us an enormous amount of data to show quite convincingly that we do indeed have conservation of momentum.

Even without this modern technology, we can still make a convincing case, on the basis of one crucial observation: when the magnet and iron collide, they come to a *dead stop*; no matter the relative size of the magnet and iron, the total momentum at the moment of collision, and thereafter, is 0. So it must have been very close to 0 just before the collision. But we can obtain a whole range of velocities just before collision simply by varying the initial distance between the magnet and iron—or, equivalently, by varying the strength of the magnet. Thus the total momentum must always be zero!

Our experiment could naturally be carried out using two magnets, but we specifically chose a magnet and a piece of iron to emphasize the *non-intuitive* nature of the third law: after all, just because the magnet pulls the iron, why should the iron be pulling the magnet (let alone with exactly the reverse of the force with which the magnet is pulling the iron)? In everyday life, we seldom give much attention to this second force. If we take one of those cute little magnets that are used to keep pieces of paper on refrigerator doors and hold it close to the refrigerator, we usually don't notice the refrigerator moving toward the magnet! What we do notice is the magnet being pulled toward the refrigerator door; nevertheless, we tend to think of this as a property of magnets, not as a property of refrigerator doors, hardly ever thinking of the refrigerator door as exerting a mysterious force.

As it happens, our modern conceptions of magnetism explain rather nicely the mutual attractions of a magnet and a piece of iron: the individual atoms of the iron each act as magnets, except that they are oriented randomly, and the magnet causes them to align, so that the iron now acts as a magnet also. Ultimately, it's all a matter of iron atoms attracting *each other*.

Similarly, when an object with a static electric charge is used to give a similar charge to a second object, we would expect to find (repulsive) forces of equal magnitude, since ultimately it's all due to the mutual repulsive forces between electrons. Of course, we also find (attractive) forces of equal magnitude between objects with opposite charges, an example of that mysterious and pervasive duality one finds throughout nature.

This seems to leave the least esoteric example—the equal (repulsive) forces between colliding bodies, no matter how different their composition—as the most difficult to understand! Eventually, this must come down to the equality of certain forces between atomic particles. Indeed, the non-intuitive nature of the third law is highlighted when we consider the most important of these atomic particles, the protons and neutrons. We naturally expect that the third law should apply to the reactions of two protons and to the reactions of two neutrons, since these are both completely symmetric situations, involving identical particles. But that doesn't explain at all why the forces between a proton and a neutron should exhibit this same sort of symmetry. In other words, in order to really understood the third law, it must be necessary to understand all sorts of nuclear physics (thus I was quite delighted to learn that a proton is made up of two "up" quarks and one "down" quark, while a neutron is made up of two "down" quarks and one "up" quark, even though I haven't the slightest idea what quarks are).

Finally, let us return to our first example of the third law, the gravitational force that the earth exerts on an object. In this case the reciprocal force of the object on the earth is essentially negligible, and the same observation prevails when we are considering the earth, or some other planet, in relation to the sun: for all practical purposes we might just as well consider only the gravitational force that the sun produces on the planet, and ignore any force that the planet produces on the sun (Jupiter is something of an exception). The situation is quite different for the case of the earth and the moon, however, and this case was of enormous interest for Newton, who made it the subject of the Principia's final Book 3, The System of the World.

The force of gravity, unlike the repulsive forces between colliding bodies, is an attractive force, and it is also a much weaker force; the gravitational attraction between two laboratory-sized objects is so minute that it was impossible for Newton to provide direct experimental evidence for the third law in the case of gravitational forces (see also Chapter 7 in this regard). Instead, this evidence is ultimately provided by Newton's analysis of the earth-moon system, and the resulting tides, and the accuracy with which it agrees with observation. As for the question of *why* the third law should hold for gravitational forces, the ultimate answer to that question presumably lies in the theory of General Relativity.

Though it might seem that we have provided far too extensive a discussion of the third law, which merits no comparable attention in physics books, there is an equal and opposite reaction, that physics books provide far too little discussion. One might well wonder why critical readers readily accept so general a law buttressed by so little experimental evidence, as if it somehow expresses a morally compelling symmetry. Perhaps it's just because it's so easy to confuse laws of nature expressing the symmetry of space with other laws that merely seem to. This will be discussed in greater detail in relation to Lagrangian mechanics (at the end of Chapter 13), but the role of symmetry in mechanics bears examination even in our current setting.

The lures of symmetry. Our statement of the second law, as $\mathbf{F} = m\mathbf{v}'$, contains within it an assertion specifically enunciated in Newton's version: A change in momentum is proportional to the motive force impressed *and takes place along the straight line in which that force is impressed.* The second phrase would probably be taken for granted even if Newton had not mentioned it. After all, what other direction could it have, other than the reverse direction? If we assume that some specific direction must be determined, and also assume that the laws of physics must be invariant under rotations, it would appear that this is the only possibility.

Similarly, in the third law it is usually taken for granted that the equal and opposite forces of reaction between two point masses are directed along the line between them. This more specific statement, often referred to as the "strong form" of the third law, is needed for later theorems in mechanics found in Chapter 3, and plays a crucial role in the analysis of rigid bodies in Chapter 5. It should be noted, however, that Newton didn't assert this stronger version of the third law—note that it wasn't used in proving conservation of momentum for the collision of two bodies—and that Newton never took up the topic of rigid bodies in the Principia. Moreover, to be candid, it must be remarked that this conclusion isn't even true without quite a bit of amplification and modification, as briefly indicated in Addendum 5A.

This leaves the other part of the third law, that the forces of reaction between two point masses are always of the same magnitude, though in opposite directions. Like Euclid's notorious Fifth Postulate, this assertion has a long history of arguments designed to demonstrate it. In fact, a wonderfully ingenious demonstration was advanced by Huygens, in his arguments concerning the rules of collisions that were referred to in Newton's Scholium (page 22). Here is a slight modification and simplification of his arguments.

First consider two identical bodies, say two steel balls, moving toward each other with equal speeds, i.e., with velocities $\mathbf{v}$ and $-\mathbf{v}$. In this simple situation it is obviously reasonable to assume, on the basis of symmetry, that their rebound velocities will also be negatives of each other, $\mathbf{w}$ and $-\mathbf{w}$, so that conservation of momentum holds: it is 0 both before and after the collision.

Now let us imagine the same experiment as observed in a coordinate system that is moving with uniform velocity $\mathbf{u}$ with respect to us, like a boat moving with respect to the shore, to take Huygens' example. In this coordinate system, the objects are moving with the initial velocities

$$\mathbf{v}_1 = \mathbf{v} + \mathbf{u} \qquad \text{and} \qquad \mathbf{v}_2 = -\mathbf{v} + \mathbf{u},$$

while their rebound velocities are

$$\mathbf{w}_1 = \mathbf{w} + \mathbf{u} \qquad \text{and} \qquad \mathbf{w}_2 = -\mathbf{w} + \mathbf{u},$$

so $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{w}_1 + \mathbf{w}_2$ $(= 2\mathbf{u})$. Since we can obtain any pair $\mathbf{v}_1, \mathbf{v}_2$ by choosing the appropriate $\mathbf{u}$ and $\mathbf{v}$, we find that in the coordinate system of the boat, moving uniformly with respect to the shore, conservation of momentum holds for two identical bodies approaching each other with *arbitrary* velocities. Of course, we could just as well interchange the role of the boat and the observer on shore, to reach the same conclusion for our observer on shore.

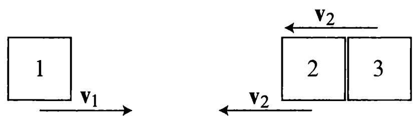
Rather than following the succeeding course of Huygens' arguments (see Problem 3-9 and remarks in Chapter 7), we will add some considerations from Volume 1 of

Feynman [1], *The Feynman Lectures on Physics*, Addison Wesley, 1963.

an essential book in the library of any one interested in physics.

Let's use steel cubes for convenience, and suppose that glue has been applied to opposing faces so that they will stick together when they meet. Symmetry dictates that when they approach each other with the same velocity and then stick together, they will end up at rest (compare page 23), so that conservation of momentum holds. Huygens' argument then implies that a collision with initial velocities $\mathbf{v}_1$ and $\mathbf{v}_2$ results in a "double cube" moving with velocity $\frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$.

We can apply these results to the case of one cube with velocity $\mathbf{v}_1$ colliding with two other cubes that have glue applied to opposing faces, but are moving



in tandem, separated by a tiny distance, with velocity $\mathbf{v}_2$. Immediately after cube 1 and cube 2 collide, they have velocities $\mathbf{w}_1$ and $\mathbf{w}_2$ satisfying

$$\text{(a)} \qquad \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{v}_1 + \mathbf{v}_2.$$

A moment later, cube 2 collides with and sticks to cube 3, after which the resulting double cube moves with velocity $\mathbf{w}_{12}$ satisfying

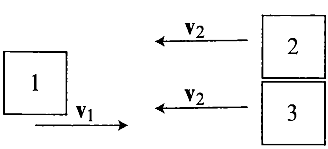
$$\text{(b)} \qquad 2\mathbf{w}_{12} = \mathbf{w}_2 + \mathbf{v}_2.$$

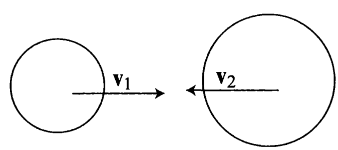
It follows that

$$\begin{aligned}
\text{initial total momentum} &= m\mathbf{v}_1 + 2m\mathbf{v}_2 \\
&= m(\mathbf{v}_1 + \mathbf{v}_2) + m\mathbf{v}_2 \\
&= m(\mathbf{w}_1 + \mathbf{w}_2) + m\mathbf{v}_2 && \text{by (a)} \\
&= m\mathbf{w}_1 + m(\mathbf{w}_2 + \mathbf{v}_2) \\
&= m\mathbf{w}_1 + 2m\mathbf{w}_{12} && \text{by (b)} \\
&= \text{final total momentum.}
\end{aligned}$$

Imagining the tiny distance decreased to 0, we conclude that conservation of momentum holds for a collision of a cube with a double cube, and we can easily generalize the argument for any multiple cube colliding with any other.

This clever argument might require supplementary considerations to deal with steel cubes stacked differently, let alone with objects of arbitrary shape,



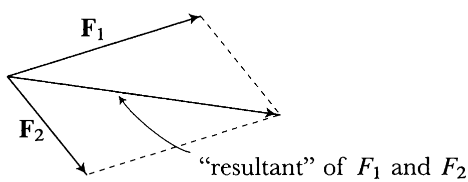


but the real problem is that it applies only to two objects made of the same "homogeneous" material. As soon as we consider objects made of different materials we are at an impasse. Given a steel cube and an aluminum cube of the same mass, approaching each other with velocities $\mathbf{v}$ and $-\mathbf{v}$, no symmetry argument allows us to conclude that the rebound velocities are also negatives of each other; having the same mass simply means that they are given the same acceleration by a given force, it says nothing about why they should react symmetrically in this situation. And that is the whole mystery of the third law.

As Feynman also points out, one cannot somehow get around this problem simply by *defining* two bodies to have the same mass if they rebound with equal speeds in such an experiment, because then there is no logical reason why "having the same mass" should be a transitive relation.

Composition of forces. Newton's statement of the conservation of momentum is actually presented as Corollary 3 of the three laws of motion. Before proving this Corollary, Newton has proceeded in the classical manner of Euclid, with a dubious proposition, providing, as Corollary 1 to his axioms, the rule for determining the effect of two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ acting on an object simultaneously.

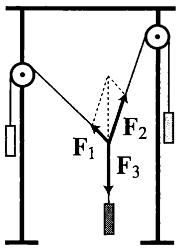
Nowadays the vector space structure of $\mathbb{R}^n$ is so ubiquitous and appears so natural that we might unhesitatingly aver that the effect must simply be the standard vector sum $\mathbf{F}_1 + \mathbf{F}_2$, constructed geometrically by the familiar parallelogram construction. But that can't simply follow from the fact that we've



decided to represent forces by vectors! In fact, the whole reason for introducing vector addition in the first place was because it represented the "addition" of forces.

Modern mathematicians have no trouble accepting the fact that Euclid's "proof" of the side-angle-side theorem is irreparably defective, and that the statement must be taken as a basic axiom. Even in physics, where the logical underpinnings of concepts is never so removed from its actual application in the real world, it should be clear that the parallelogram rule for combining forces can't possible be proved on the basis of the three basic laws, since none of them say a thing about two forces acting at once.[1] That hasn't prevented many eminent physicists and mathematicians from attempting to fashion a proof, but consideration of those attempts, requiring such involved and ultimately futile maneuvers, has been relegated to Chapter 7.

Physicists nowadays seem resigned to the stance of regarding the parallelogram law as just another law based on observation, and mechanisms like the one pictured below may be used to illustrate it in classroom settings. Strangely



$\mathbf{F}_3 = -(\mathbf{F}_1 + \mathbf{F}_2)$
at equilbrium

enough, however, if the parallelogram law really is an experimental fact, then one would expect physicists to be testing it with great precision, but no one every mentions such experiments!

I imagine physicists would say that the parallelogram law of forces is demonstrated by the consistency with which its many uses intermesh, but I suspect the real reason no one bothers to test this law is because everyone thinks that it really *has* to be true, as Chapter 7 so eloquently testifies.

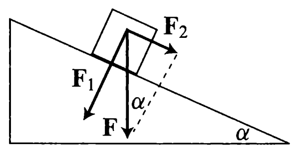
With a law for the composition of forces at hand, we can now consider the total momentum of a general system of particles $c_1, \dots, c_K$, with masses $m_1, \dots, m_K$. Assuming there are no "external" forces (like the force of gravity), but only the "internal" forces $\mathbf{F}_{ij}$ that the particle c_j exerts on c_i, with $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$, in accord with the third law, the total force on particle c_i is $\sum_j \mathbf{F}_{ij}$, and consequently

$$\sum_i m_i \mathbf{v}_i{}' = \sum_i \Big(\sum_j \mathbf{F}_{ij}\Big) = 0,$$

which means that $\sum_i m_i \mathbf{v}_i$, the total momentum of the system, must be a constant.

[1] From a strictly logical point of view, it is not even clear that two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ acting simultaneously should have the same effect as any other single force $\mathbf{F}$: while it's true that the combined forces must end up producing an acceleration of *some* sort on each object, that acceleration might not be proportional to the mass of the object, even though the accelerations produced by $\mathbf{F}_1$ and $\mathbf{F}_2$ individually are.

It should also be noted that the parallelogram law is not only used to determine the combined effect of two forces, but just as frequently it is used to *decompose* a given force into two hypothetical ones. For example, the usual elementary way to analyze the motion of a block sliding down a stationery inclined plane is to decompose the force $\mathbf{F}$ of gravity on the block into a force $\mathbf{F}_2$ parallel to the inclined plane, and a force $\mathbf{F}_1$ perpendicular to the inclined plane. Letting α be the angle from the horizontal to the incline of the plane, we see that the magnitude $|\mathbf{F}_2|$ of $\mathbf{F}_2$ is $\sin\alpha \cdot |\mathbf{F}|$.



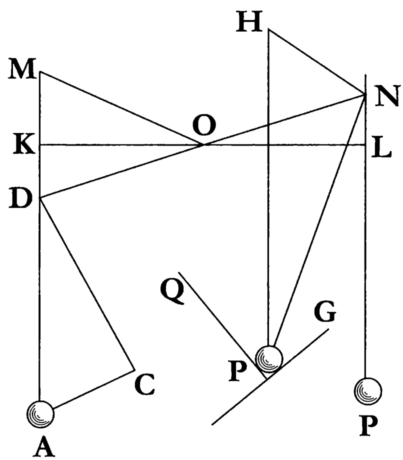
With this decomposition, the force of the block on the inclined plane is $\mathbf{F}_1$, so the inclined plane must exert a force of $-\mathbf{F}_1$ on the block, by the third law; the net effect is that the total force on the block is simply $\mathbf{F}_2$, from which we can determine the motion of the block explicitly—it undergoes uniform acceleration that is $\sin\alpha$ times its acceleration in free fall.

Of course, this all depends on the slight-of-hand introduction of a force canceling $\mathbf{F}_1$, with the implicit assumption that the force $\mathbf{F}_2$ has no effect on the incline plane. We will discuss the justification for such machinations in a little greater detail in Chapter 6, but for now we simply note that they are commonplace in elementary mechanics courses, and Newton had no aversion to them.

As a matter of fact, right after his proof of the parallelogram law of forces, which is Corollary 1 of the three laws, Newton states Corollary 2:

> *And hence the composition of a direct force* AD *out of any oblique forces* AB *and* BD *is evident, and conversely the resolution of any direct force* AD *into any oblique forces* AB *and* BD. *And this kind of composition and resolution is indeed abundantly confirmed from mechanics.*

To illustrate, Newton then launches a long investigation of a mechanical situation. A replica of the rather complex diagram that appears in the Principia is shown on the opposite page, but we won't reproduce his discussion. We merely mention that along the way he proves the law of the lever, and if you believe that his proof can be valid, then you can go read it yourself (Newton [2; pp. 418–420] as well as Chandrasekhar [2; pp. 24–25]).



Newton, in a couple of pages, has offered a Student Guide to solving mechanics problems: "the whole of mechanics—demonstrated in different ways by those who have written on the subject—depends on what has just now been said." But Newton doesn't intend to add much more, explaining near the end of the Scholium, "But my purpose here is not to write a treatise on mechanics."

Our purpose, on the other hand, *is* to write a treatise on mechanics, so we will soon be parting company with Newton. Before doing that, however, in the next chapter we will discuss Book 1, "The Motion of Bodies", which comes immediately after the "Definitions" and "Axioms, or the Laws of Motion", to see one of the reasons why Newton *did* write the Principia.

ADDENDUM 1A

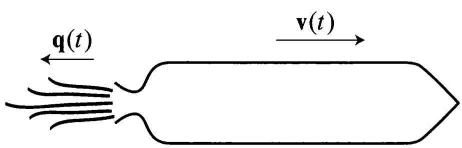
IT ISN'T ROCKET SCIENCE
(Why Easy Physics is So Hard: I)

School children are customarily told that rockets work because of the third law, instead of being provided with the more prosaic explanation that the burning fuel explodes both toward the front of the rocket and toward the rear, with the fuel that explodes toward the front pushing the rocket forward, while the fuel that explodes toward the rear simply escapes out the open end of the rocket.

As a matter of fact, the action of a rocket doesn't follow so clearly from the third law itself as from it's consequence, the conservation of momentum: since the burning fuel has momentum in one direction, the remaining fuel and rocket must have momentum in the other direction in order to conserve the total momentum—namely 0 if the rocket starts at rest. As this example illustrates, one of the main attractions of "conservation" laws is that they often allow us to consider quite complicated situations involving myriad particles at once.

Of course, a few idealizations are required here: rocket fuel is usually a liquid, but it is not unreasonable to regard it as a bunch of particles (which, at the atomic level, it really is); and the (empty) rocket itself is hardly a particle, and we might demand a more careful analysis of how the force exerted on one part gets transmitted to other parts, but we will defer that to a later chapter.

Now it would seem fairly straightforward to get an analytic expression for the motion of a rocket. We'll first consider a rocket in empty space, so that there is no external force acting on it. Let $\mathbf{v}$ be its velocity, and let $m(t)$ be the mass at time t, by which we mean the constant mass of the empty rocket plus the mass of the fuel still in the rocket at time t. Equivalently, $-m' = \mu$ is the rate at which the fuel is burned; this depends ultimately on the chemical characteristics of the fuel, the design of the combustion mechanism, etc., etc. We also need to consider the velocity $\mathbf{q}$ with which the fuel is ejected from the rocket; we'll use $\mathbf{q}$ for the velocity *with respect to the rocket*, so that $\mathbf{q} + \mathbf{v}$ is the velocity with respect to our inertial system. If we ignored $\mathbf{q}$, then we could just as well assume that the fuel was simply being dumped off the rocket ($\mathbf{q} = 0$), which wouldn't result in any motion at all!



The simplest analysis proceeds from the prosaic point of view suggested at the beginning. In a short time interval $[t, t + h]$, the amount of fuel ejected is

$m(t) - m(t+h)$, and therefore the momentum of expelled fuel will be close to $[m(t) - m(t+h)] \cdot \mathbf{q}(t)$. Thus, the momentum of the fuel in the other direction, pushing the rocket forward, will be the negative of this. So the force on the rocket must be the derivative, $m'(t) \cdot \mathbf{q}(t)$, and by the second law this means that

$$(*) \qquad m'(t)\mathbf{q}(t) = m(t)\mathbf{v}'(t).$$

This can also be written as

$$\mathbf{v}'(t) = \frac{m'(t)}{m(t)}\mathbf{q}(t) = \frac{d}{dt}\log(m(t))\mathbf{q}(t).$$

In our example, $\mathbf{v}$ and $\mathbf{q}$ always point along the same straight line, and if we let v and q denote their lengths—the speed of the rocket and the speed of fuel ejection—then, remembering that $\mathbf{q}$ and $\mathbf{v}$ point in opposite directions, we can simply write

$$\text{(R)} \qquad v'(t) = -q(t)\frac{m'(t)}{m(t)} = -q(t)\frac{d}{dt}\log(m(t)).$$

This argument might seem suspect, since we appear to be working in a coordinate system based on the rocket, which is not an inertial system, but that isn't really the case. Although we derived $(*)$ from the "rocket's point of view", at each particular time t_0 we were essentially working in the inertial system that is moving with the same velocity as the rocket at time t_0, and since the derived equation $(*)$ involves only the *change* $\mathbf{v}'(t_0)$ of velocity at time t_0, it holds just as well in the inertial system where we are making our measurements of position. Nevertheless, most physics books avoid tackling an explanation of this sort and instead present the following analysis.

Since the ejection velocity of the fuel with respect to our inertial system is $\mathbf{q}+\mathbf{v}$, in a small time interval $[t, t+h]$, the amount of fuel ejected, $m(t)-m(t+h)$, has a velocity close to $\mathbf{v}(t) + \mathbf{q}(t)$, so the total momentum of this expelled fuel is close to

$$[m(t) - m(t+h)] \cdot \big(\mathbf{v}(t) + \mathbf{q}(t)\big).$$

The derivative at time t of the momentum of the expelled fuel is

$$\lim_{h\to 0} \frac{m(t) - m(t+h)}{h} \cdot \big(\mathbf{v}(t) + \mathbf{q}(t)\big) = -m'(t) \cdot \big(\mathbf{v}(t) + \mathbf{q}(t)\big).$$

Setting this equal to the derivative of $-m(t)\mathbf{v}(t)$, we get

$$\frac{d}{dt}[m(t)\mathbf{v}(t)] = m'(t) \cdot \big(\mathbf{v}(t) + \mathbf{q}(t)\big),$$

which can also be written as our original equation

$$(*) \qquad m'(t)\mathbf{q}(t) = m(t)\mathbf{v}'(t).$$

The funny thing about this problem is that we tend to think of it as a "real-life" problem, involving a continuously changing fuel mass, and then find ourselves in the position of having to use laws that apply only to individual particles. But if we made our "real-life" problem really real, by considering the fuel as a collection of particles being ejected individually in tiny increments of time, then we would view our rocket as receiving tiny changes of velocity at these times, but moving with constant velocity in the intervals between. In other words, our rocket is an inertial system on these intervals, which makes the validity of the first argument much more transparent.

A completely independent source of confusion is offered by some less recent mechanics texts, which like to point out that in special relativity, which we will barely mention in this volume, the mass m of even an individual particle is not constant, but depends upon its velocity; it then turns out that the second law, which we have always stated as

$$\begin{aligned}\mathbf{F} &= (m\mathbf{v})'\\ &= m\mathbf{v}'\end{aligned}$$

has to be corrected to read

$$\mathbf{F} = (m\mathbf{v})' = m'\mathbf{v} + m\mathbf{v}'.$$

Of course, the mass of a particle doesn't vary with velocity in classical mechanics, but when we bring an external force $\mathbf{F}$ into the picture, we might wonder whether this more general equation is the right one to use for a rocket, with variable mass $m(t)$. If $\mathbf{F}$ is an external force on the rocket, e.g., gravity, should we use

$$(1) \qquad \mathbf{F}(t) = m(t)\mathbf{v}'(t)$$

or

$$(2) \qquad \mathbf{F}(t) = (m\mathbf{v})'(t)\,?$$

The short answer is: neither. When $\mathbf{F} = 0$, equation (1) contradicts $(*)$ unless $\mathbf{q} = 0$; this is hardly surprising, since $\mathbf{F}$ is simply the "external" force on the rocket, and we still have to account for the force exerted by the escaping fuel. And (2) likewise contradicts $(*)$ when $\mathbf{F} = 0$ except in the special case where $\mathbf{q} = -\mathbf{v}$.

In fact, using the same reasoning by which we established $(*)$ we can conclude more generally that when an external force $\mathbf{F}$ acts on the rocket we have

$$
\boxed{
\begin{array}{ll}
(**) \qquad & \mathbf{F}(t) = m(t)\mathbf{v}'(t) - m'(t)\mathbf{q}(t) \\
(**') \qquad & \mathbf{F}(t) = (m\mathbf{v})'(t) - m'(t)[\mathbf{v}(t) + \mathbf{q}(t)].
\end{array}
}
$$

A more complete answer is that the question is completely misleading. It purports to be studying an object, "the rocket", that has a variable mass. But the objects that we really have are the empty rocket, together with a myriad of particles of rocket fuel, some of which are moving along with the empty rocket, while others are moving in the opposite direction.

At any time t the force of the burning fuel acts on the empty rocket together with the part of the fuel still moving along with it. By additivity of mass, this composite object is given the same acceleration as a single object whose mass is what we have called $m(t)$, and which we misleadingly called the "mass of the rocket at time t".

In any case, the formula $(**)$ or $(**')$ states the proper result and this analysis applies just as well to any "variable mass" problem, where a body's mass is changing as particles are continually dispersed, or are added on, with velocity $\mathbf{q}$ (or velocity $\mathbf{v} + \mathbf{q}$ with respect to an inertial system).

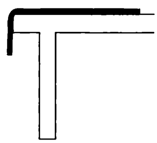
Unfortunately, some textbooks like to claim that the proper law is, in fact, (2), perhaps under the sway of the influential text *Mechanics* by Sommerfeld [2], where Newton's statement of the second law is specifically identified as (2), with the remark that "Newton's formulation ... prophetically turns out to be the correct one."

Naturally, considerable backtracking is required when (2) has been perpetrated as the proper relationship, as shown, for example, by the thorough, and thoroughly confusing, discussion in §I.4 of Sommerfeld's book.

To make matters worse, one class of variable mass problems $(**')$ reduces to (2)—namely, those in which $\mathbf{v} + \mathbf{q} = 0$. For example, consider a satellite moving in empty space uniformly filed with stationery interplanetary debris that sticks to the satellite when it hits it, thereby increasing the satellite's mass $m(t)$. In this case, (2) is applicable because it amounts to $(**')$ with $\mathbf{v} + \mathbf{q} = 0$, which merely says that the accumulated debris is initially at rest before it starts moving along with the satellite. (Since our analysis was made in terms of mass removed from a body, it might help to think of the reverse time picture, where bits of the satellite are being expelled with exactly the velocity that they need to become part of the stationery interplanetary debris.)

A similar situation, where again the proper equation $(**')$ reduces to (2), occurs for a raindrop falling through an atmosphere saturated with water vapor, and accumulating mass by condensation (Problem 12).

A typical contrasting case is illustrated by a fine chain with very small links lying on a table, with a small piece hanging over the edge, initially held at rest



and then released (Problem 13). Now the added particles, the small links, are already moving with the velocity of the hanging piece, so we have $\mathbf{q}(t) = 0$, and equation $(**)$ instead reduces to (1).

Further consideration of these examples will be found in the Problems, and in Addendum 3A of Chapter 3 (page 100).

ADDENDUM 1B
WEIGHT VERSUS MASS

Newton had good reason for wanting to check the proportionality of mass to weight with a high degree of accuracy, since it is crucial to his "universal law of gravity", that the force $\mathbf{F}$ between two bodies of masses m_1 and m_2, separated by a distance d, has magnitude

$$|\mathbf{F}| = G\frac{m_1 m_2}{d^2},$$

where G is some "universal constant".[1]

In fact, at the very beginning of the Principia, Newton states, right after the definition of mass, "It can always be known from a body's weight, for—by making very accurate experiments with pendulums—I have found it to be proportional to the weight, as will be shown below." The experiments in question are completely different from those mentioned at the bottom of page 22, and Newton's reference to these other experiments must have caused many a reader to scurry vainly through the succeeding pages of the Principia, because the presentation of these particular experiments actually occurs several hundred pages later in the Principia, in Proposition 6 of Book 3!

To discuss Newton's experiments, we will use the result of Problem 17(c), showing that the periods T_1 and T_2 of a pendulum bob undergoing different accelerations g_1 and g_2 stand in the relation.

$$\frac{g_2}{g_1} = \frac{{T_1}^2}{{T_2}^2}.$$

We are trying to compare the accelerations g_1 and g_2 on two objects, of masses m_1 and m_2. Denoting their *weights* by $W_1 = g_1 m_1$ and $W_2 = g_2 m_2$, the above equation can be written

$$(*) \qquad \frac{m_1}{m_2} = \frac{W_1}{W_2} \cdot \frac{{T_1}^2}{{T_2}^2},$$

and in particular, if $W_1 = W_2$ we have

$$(**) \qquad \frac{m_1}{m_2} = \frac{{T_1}^2}{{T_2}^2}.$$

---

[1] More precisely, the factor m_2 involves the proportionality of a body's mass m_2 and of the force exerted on it by another body, while the fact that the masses of both bodies enter symmetrically, as the double factor $m_1 m_2$, is basically equivalent to the third law for gravitational forces; as we shall see in the next chapter, the factor $1/d^2$ is consistent with the elliptical orbits of planets, and the fact that G is the same for all bodies is consistent with Kepler's third law.

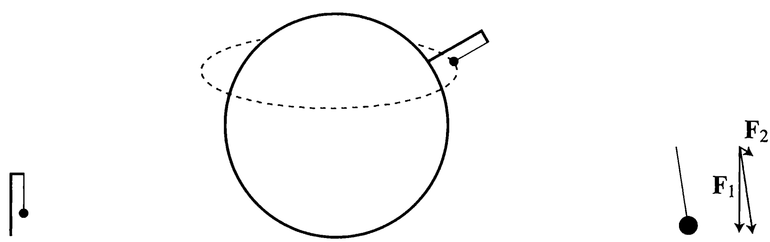
The result (∗) occurs as Proposition 24 of Book 2 of the Principia, Newton [2; pg. 700]; presumably Newton's proof —with nary an equation, or even an algebraic symbol, in sight—is equivalent to that in Problem 17.

Newton tested (∗∗) with equal weights of "gold, silver, lead, glass, sand, common salt, wood, water, and wheat". Each pair of materials to be tested was enclosed within one of two rounded, equal-sized wooden boxes. For the wood bob he simply filled the inside of the box with more wood, but for the gold bob he suspended the gold at the center of the box; he then hung each of the two boxes by eleven-foot cords, which "made pendulums exactly like each other with respect to their weight, shape, and air resistance."

He then placed them close to each other, and started them swinging from the same height, noting that "they kept swinging back and forth together with equal oscillations for a very long time. ... And it was so for the rest of the materials. In these experiments, in bodies of the same weight, a difference of matter that would be even less than a thousandth part of the whole could have been clearly noticed." Newton doesn't give any further details, but it should be noted that Problem 17 (a) shows that an inaccuracy in measuring the lengths of the pendulums would produce errors of the same magnitude as those we are trying to eliminate.

In 1832 Bessel made improved pendulum experiments that established proportionality within an error of 2 parts in 10^5 (cf. Problem 6-3). But the most accurate experiments depend on an ingenious idea first introduced by Eötvös in 1889, which enables us to consider only static measurements.

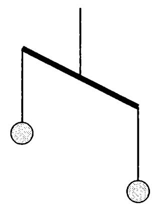
Although a suspended weight may seem to be hanging vertically in the laboratory (a), because of the earth's rotation it is really traveling along a circle (b),



(a) (b) (c)

and its natural tendency to travel in a straight line means that it acts as if there were a small "centrifugal force" pulling it away from the the axis of rotation. So the weight is really hanging at a slight angle (c), whose direction is compounded from the downward force $\mathbf{F}_1$ of gravity on the object, and the acceleration $\mathbf{F}_2$ from its circular motion (centrifugal force is discussed in Chapter 10, where detailed computations of the angle can be found, on page 380).

The angle of $\mathbf{F}_2$ is determined by the latitude of the laboratory, and its length depends—aside from the rate of revolution of the earth, and the latitude and radius of the earth—only on the *mass* of the object. On the other hand, the force $\mathbf{F}_1$ depends on the *weight* of the object. Consequently, if two objects have differing ratios of mass and weight, the angles at which they are hanging will be slightly different. So if they are suspended from opposite ends of a bar



attached to a sensitive torsion balance, they ought to twist the torsion fibre in one direction.[1]

Of course, this twist would be quite small, and various side effects might produce a twist that overwhelmed it. But Eötvös was able to look for the existence of a net twist by enclosing the whole apparatus within a chamber that could be rotated through 180°, which would reverse the twist due to differing ratios of mass and weight. As a result of his experiments, Eötvös concluded that weight and mass must be proportional within an error of 1 part in 10^7.

One crucial point about the Eötvös experiment is that the tested objects need not have exactly the same mass: so long as their masses are close, a difference in the ratios of force to mass would produce a net twist. Thus, as in Newton's pendulum experiment, we can show proportionality of weight to mass to a high degree of accuracy without having to measure mass itself accurately at all.

At the other extreme from the naive belief that heavier bodies fall faster than lighter ones, we have Einstein's extremely sophisticated view regarding the proportionality of mass and weight, that the identical acceleration experienced by all bodies must indicate that this acceleration is not really due to a force at all. This led to the general theory of relativity, which interprets free fall trajectories as natural geodesics in a curved space-time.

[1] The use of a torsion balance goes back to Henry Cavendish (1731–1810), who used it, in one of the first physics experiments performed to a high degree of accuracy, to measure the extremely small gravitational force between two lead spheres. He expressed the result in terms of the density of the earth, though his neighbors always described the building where this was done as the place where the world was weighed.

By the way, as far as I know, no one has ever conducted an experiment of this nature to test Newton's third law for gravitational forces.

This foundational role for the proportionality of mass and weight has inspired even more refined experiments of the Eötvös type, starting with the 1964 paper of Roll, Krotkov, and Dicke [1], which verified proportionality within 1 part in 10^{11} for gold and aluminum.

The book *Gravitation*, by Misner, Thorne, and Wheeler [1], gives a brief description of the experiment, and points out how much care is required to rule out extraneous influences. Many more details are provided by the paper itself, which undoubtedly can serve as an inspiration for any aspiring experimental physicist.

PROBLEMS

1. The purpose of this first problem is simply to remind us once again how slippery elementary physics problems can be.

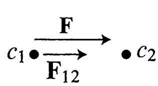
A man pushes against a cart with his hand, and the cart, initially at rest, starts to move. So his hand must be exerting a force on the cart. By the third law, the cart must be exerting a force on his hand in the opposite direction. How can that be, when his hand is obviously not accelerating in the opposite direction? (Compare the situation where the man is thrown against the cart with his hand outstretched.)

His hand has other forces on it, transmitted through his body by the force that the floor exerts on his feet. This force is due to the fact that his feet exert force on the floor (that's how he pushes), and the floor has friction (otherwise he would just slip backwards when he tried to push the cart). Altogether, we have an incredibly complicated system whose details we thankfully needn't explore right now.

2. (a) Consider two objects B_1 and B_2 of equal mass m. The object B_2 is stationary, while B_1 is moving towards B_2 with velocity $\mathbf{v}$. Suppose that after the collision, B_1 and B_2 have velocities $\mathbf{w}_1$ and $\mathbf{w}_2$ that are collinear with $\mathbf{v}$ and, moreover, point in the same direction as $\mathbf{v}$. Show that $|\mathbf{w}_1| \le |\mathbf{v}|/2$. (This is a very simple physics problem, not a mathematics problem!)
(b) Suppose that a system of particles $c_1, \dots, c_K$ with masses $m_1, \dots, m_K$ satisfies conservation of momentum, i.e., $\sum_i m_i \mathbf{v}_i{}' = 0$. Show that there are forces $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ such that $m_i \mathbf{v}_i{}' = \sum_j \mathbf{F}_{ij}$. (This is a simple mathematics problem, not a physics problem.)
(c) Can we always choose the $\mathbf{F}_{ij}$ so that they satisfy the "strong form" of the third law (page 25)?

3. Nowadays, rather than working from impulsive forces to continuous ones, we are more apt to reverse the process. Suppose that on the time interval $[0, h]$ we have a force acting along the x-axis with magnitude $f(t) \ge 0$ at time t, and $f(0) = f(h) = 0$. For a particle of mass m starting at rest, so that its distance $x(t)$ from 0 satisfies $0 = x(0) = x'(0)$, show that $x'(h) = F = \int_0^h f$. So we can basically think of an "impulsive" force at time 0 as one that suddenly changes our particle from one at rest to one with velocity F.

176 **4.** This problem has two parts. (See page viii for the role of the symbol.)
(a) Here is the precisely stated part. Let c_1 and c_2 be two particles, of masses m_1 and m_2, let $\mathbf{F}_{12} = -\mathbf{F}_{21}$ be the "internal" forces between them, and let $\mathbf{F}$ be



an external force on c_1 that always points in the direction from c_1 to c_2, so that under these combined forces the particles move along a straight line.

Now suppose that c_1 and c_2 *always remain the same distance apart.* Find $\mathbf{F}_{12}$ in terms of $\mathbf{F}$. (In particular, if $\mathbf{F}$ is a non-constant force, then $\mathbf{F}_{12}$ must also vary with time, even though the distance between c_1 and c_2 remains the same. Since we usually think of the internal forces as functions of the distance between the particles, rather weird internal forces would thus be necessary for truly rigid bodies, which are the subject matter of Chapter 5.)

(b) Here is the version of this problem as it appears in elementary physics books. Two blocks B_1 and B_2, of masses m_1 and m_2, are in contact on a frictionless

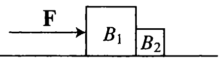
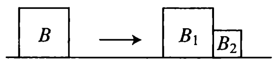


table. A horizontal force $\mathbf{F}$ is applied to B_1. Find the forces between B_1 and B_2 in terms of $\mathbf{F}$.

In this form, the problem appears to assume tacitly that the blocks remain in contact. Suppose the force is delivered as the result of another block B of mass M colliding with B_1. The fact that the system of all three blocks satisfies



conservation of momentum does not limit us very much: show that B_1 and B_2 can move in such a way that they remain in contact, and also in such a way that they do not. If we consider $\mathbf{F}$ to be an "instantaneous" force applied at some time t_0 can we find $\mathbf{F}_{12}(t_0)$?

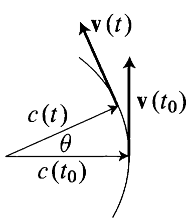
68 **5.** (a) For circular motion with constant angular frequency,

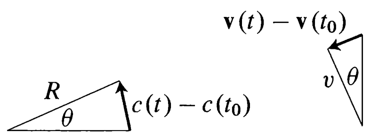
$$c(t) = R(\cos \alpha t, \sin \alpha t)$$

calculate that the magnitude of the acceleration can be written as v^2/R, where v is its constant speed.

(b) More generally [and more easily proved]: for any motion on a sphere of radius R, the inward component of the acceleration at time t has magnitude $v(t)^2/R$.

Note, by the way, that part (a) can easily be proved geometrically:





$$\frac{|\mathbf{v}(t) - \mathbf{v}(t_0)|}{v} = \frac{|c(t) - c(t_0)|}{R}$$

6. Suppose we have an accurately calibrated spring scale and we weigh an object in an elevator that has an acceleration a. What will the scale read? Distinguish between the cases of an upward acceleration and a downward acceleration.

7. Here is a problem that might appeal to bright physics students, especially those interested in experimental physics; I wouldn't even know where to begin.

(a) Dr. Fignewton publishes a paper in the *Journal of Irreproducible Results* in which he claims to have performed experiments suggesting that the second law should actually read

$$\mathbf{F} = m(\mathbf{v}' + \alpha\mathbf{v}'')$$

for a very small "universal constant" $\alpha \approx 0.000327\ldots$. Devise an experiment of sufficient accuracy to refute this.

It would appear that one must investigate motion under some non-uniform force $\mathbf{F}$, but it is important to make sure that the argument doesn't beg the question, by implicitly measuring $\mathbf{F}$ through an appeal to the second law.

(b) Devise an experiment to verify the parallelogram law of forces within 1 part in 10^{11} (all right, I'm actually willing to settle for 1 part in 10^5).

8. In the theory of special relativity, the mass m of an object moving with speed v is related to its "rest mass" m_0 by the equation

(a) $$m = m_0(1 - v^2/c^2)^{-1/2},$$

and the second law is modified to

(b) $$\mathbf{F} = (m\mathbf{v})'$$

(so that the argument on page 22 can still be used to show that momentum is conserved). Use (a) and (b) to derive the formula on page 17.

For the following problems, we adopt the usual notation that g denotes the acceleration produced by gravity on a body near the earth's surface. Equivalently, gm is the magnitude of the force on the body of mass m.

9. Assume that the earth is exactly spherical, and that a particle c, under the influence of the earth's gravity, circles the earth just above the surface of the earth—also ignore air resistance, the moon and the sun, and so forth. Show that the speed of the particle is $\sqrt{gR_e}$, where R_e is the radius of the earth.

10. (a) Consider a rocket in empty space, initially at rest, and assume that q in equation (R) on page 33 is constant. Let the mass of the fuel be $r \cdot m(0)$ $(0 \le r < 1)$. Show that the final velocity achieved by burning all the fuel is $q \log 1/(1-r)$. Notice that this is independent of the particular manner in which the fuel is burned (i.e., independent of m').
(b) Suppose our rocket is initially at rest on the earth, pointing upwards, and for simplicity suppose that the downward acceleration due to gravity is the constant g. Show that if the final velocity is achieved at time t_0, then this final velocity is $[q \log 1/(1-r)] - gt_0$ (so the faster we burn the fuel the better).

The next problem is a companion to this one.

11. (a) A group of people, each with the same mass, stand at one end of a cart that can roll without friction, and jump off with the same speed. Show that if everyone jumps off at once, then the resulting speed of the cart is greater than the final speed the cart acquires if they jump off one at a time. ("Jumping off" should be regarded as involving an instantaneous force).
(b) Analogously, consider the rocket of the previous problem, initially at rest in empty space, and suppose that all the fuel is ejected "instantaneously" with speed q. Show that the rocket achieves the final velocity $qr/(1-r)$, which is greater than the velocity $q \log 1/(1-r)$ found in part (a) of that problem. How can these results be reconciled?

12. Consider a falling raindrop that gains mass from the surrounding atmosphere saturated with water vapor, so that equation (2) on page 34 holds.

(a) If the rate at which it gains mass is proportional to its surface area, say $m' = \alpha \cdot (\text{surface area})$, then its radius $r(t)$ satisfies $r(t) = r_0 + \alpha t$, where r_0 is its radius at time 0, and its speed $v(t)$ satisfies

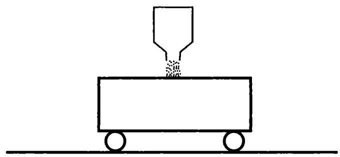
$$v(t) = \frac{g}{4}\left(t + \frac{r_0}{\alpha}\right) + \left(v_0 - \frac{g}{\alpha}\cdot\frac{r_0}{4}\right)\left(1 + \frac{\alpha t}{r_0}\right)^{-3},$$

where v_0 is its speed at time 0.
(b) Suppose instead that it gains mass at a rate proportional to the product of its surface area and speed v. Show that its acceleration asymptotically approaches a value.
(c) Suppose that it gains mass at a rate proportional to the product of its mass and its speed v. Show that the speed v asymptotically approaches a value.

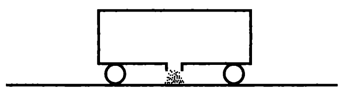
100 **13.** Let the chain on page 36 have length l, mass M, and uniform density $\mu = M/l$, and let l_0 be the amount initially hanging down. Find the length $x(t)$ hanging down after time t, and the time at which the whole chain has fallen off

the table. (Note that the whole chain is always being pulled, and recall that the equation $f'' = \alpha f$ has the solutions $f(t) = e^{\pm\sqrt{\alpha}\,t}$.) (Compare Problem 3-23.)

14. Physics textbooks like to consider a railroad cart rolling without friction, while its mass increases at a steady rate ρ because of sand from a hopper falling into it. This variable mass problem can be considered to have $\mathbf{v} + \mathbf{q} = 0$, since we really only want to consider the horizontal components of forces and velocities.



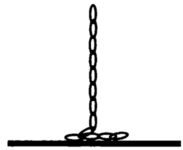
(a) Suppose that the empty cart, of mass M, is at rest at time 0, and a constant horizontal force of magnitude F is applied as the sand starts to fall. Find the speed $v(t)$ at time t, and the limiting velocity as $t \to \infty$.
(b) Now consider the opposite case where the cart has an opening in the bottom that allows sand to pour out, decreasing the mass at the steady rate ρ. At time 0



the cart is at rest, and the constant horizontal force of magnitude F is applied to it. If the sand has mass m at time 0, find the velocity of the cart when all the sand has poured out.
(c) Finally, suppose that mass is increasing at the constant rate ρ_i from sand pouring in the top, and decreasing at the constant rate ρ_o from sand pouring out the bottom. What is the proper equation of motion?

118 **15.** (a) Suppose that the railroad cart of the previous problem has been set moving with a constant speed v. Sand starts pouring into it in some arbitrary way, with $m(t)$ being the total amount poured in by time t. What force $F(t)$ is needed to insure that the cart will continue to have constant speed v? (Compare Problem 3-22.)

16. A fine chain of length l and mass M, with uniform density $\mu = M/l$, is held at one end, with the free end just touching the floor, and then released at time $t = 0$. The problem is to compute the force that the chain exerts on the



floor, as a function of t, or equivalently, as a function of the length $x = \frac{1}{2}gt^2$ of chain that has fallen.

(a) A straightforward answer can be obtained in the same way as our rocket equation was derived. The fallen part itself contributes a force equal to its weight $\frac{1}{2}\mu g^2t^2$, and we need to determine the force contributed by the falling chain, the end of which has speed gt. The amount of momentum contributed by the falling chain in a small time interval h is close to $(\mu gt \cdot h) \cdot gt$. Conclude that the force of the chain on the floor at time t is

$$\frac{3}{2}\mu g^2t^2 = 3\mu gx.$$

When the chain has just finished falling, this has the value $3gM = 3W$, where W is the weight of the chain, although one would assume that it should simply be the weight W from that time on; we'll worry about this discrepancy in part (c).

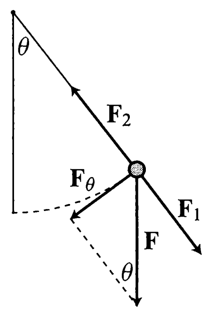
(b) A nifty way to solve the problem directly from our general equation $(**)$ on page 35 is to consider the force $\bar{\mathbf{F}}$ that the floor exerts back on the part of the chain lying on the ground, which we will consider as an object of variable mass $m(t) = \frac{1}{2}\mu gt^2$, which happens to have velocity 0 when acted upon by gravity and $\bar{\mathbf{F}}$. We have $m'(t) = \mu gt$ and gt is also the magnitude of $\mathbf{q}(t)$ in this situation. Use equation $(**)$ to show that $\bar{\mathbf{F}}(t)$ has magnitude $3\mu gx$.

(c) This problem is usually stated in terms of the reading of a scale onto which a chain is falling (a), assuming, though not explicitly (unless the textbook writers have a conscience), that the mechanism of the scale is so tight that the pan

(a) (b)

stays at nearly the same height throughout. Of course, the pan actually dips a bit, and after the chain has finished falling the pan will suddenly pop back up, with the scale reading of $3W$ quickly correcting itself to W. Even when the chain is dropped directly on the floor, the floor will be pushed down a tiny bit, and eventually pop back up a bit at the end of the fall. To get a more realistic problem, we need to consider a less realistic picture (b) where the weighing mechanism is explicitly based on the compression of a spring. Using *Hooke's Law*, which says that when one end of the spring is compressed by a distance s, the spring exerts a force in the other direction of magnitude $k \cdot s$ for some constant k, find an equation for s in terms of t and this "spring constant" k. By the way, with most scales, one can observe a scale reading greater than W, but it's practically impossible to observe a reading anywhere close to $3W$, since the chain falls so rapidly and the scale usually responds so slowly.

37 **17.** The usual analysis of a pendulum involves decomposing the gravitational force $\mathbf{F}$ of magnitude gm on the bob into a force $\mathbf{F}_1$ in the direction of the



pendulum string and another force $\mathbf{F}_\theta$ tangent to the path of the bob. The string is also exerting a force, $\mathbf{F}_2$, on the bob, which is assumed to point along the direction of the string. We must have $\mathbf{F}_2 = -\mathbf{F}_1$, since we assume that the bob stays at a constant distance from the pivot point, keeping the string taut but not stretching it out. (This reasoning is all examined more carefully in Chapter 6.) Thus, the net force on the bob is $\mathbf{F} + (-\mathbf{F}_1) = \mathbf{F}_\theta$, and consequently the acceleration of the bob, tangent to the circular path, has magnitude

$$a_\theta = g \sin\theta. \tag{1}$$

If we consider θ as a function of time, and let l be the length of the string (the radius of the circle on which the pendulum bob moves), then equation (1) yields

$$\theta'' + \frac{g}{l} \sin\theta = 0. \tag{P}$$

Although this simple little equation can't be solved explicitly, some important exact information can be derived using only trivial mathematical manipulations:

(a) For convenience, choose the origin O to be the point from which the pendulum hangs. For any $\alpha > 0$ consider the path

$$\gamma(t) = \alpha \cdot c\big(t/\sqrt{\alpha}\big),$$

which follows a circle with radius α times the radius of the path c, but with the time reparameterized by the factor $1/\sqrt{\alpha}$. Then the angle $\vartheta(t)$ for γ satisfies

$$\vartheta(t) = \theta\big(t/\sqrt{\alpha}\big).$$

Conclude that

$$\vartheta'' + \frac{g}{\alpha l}\sin\vartheta = 0,$$

so that γ gives the path of pendulum bob with length α times that of the original.

(b) The period of a pendulum, for a given fixed initial angle θ_0, is the time required for the pendulum to return to this position, after first swinging through to the angle $-\theta_0$. Conclude that the period of the pendulum described by γ for the angle θ_0 is $\sqrt{\alpha}$ times the period of the pendulum described by c for the same angle θ_0. Briefly, the period of a pendulum is proportional to the square root of its length.

(c) Instead of varying the length of the pendulum, let us instead assume that we have changed the acceleration g to $g \cdot \alpha$. (One way of doing this would be move the pendulum to the moon, for example, though in Addendum 1B we instead consider the possibility that g depends on the material comprising the pendulum bob.) In this case, show that the motion of the pendulum is now described by the curve

$$\gamma(t) = c\big(t/\sqrt{\alpha}\big),$$

and conclude that the period of a pendulum is inversely proportional to the square root of the acceleration g.

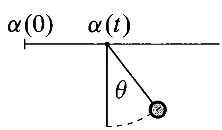
(d) The speed of the pendulum bob at time t is $v = l|\theta'(t)|$. Use the pendulum equation to show that if $h(t)$ is the height of the pendulum bob at time t, then

$$g \cdot h(t) + v^2(t) = \text{constant}$$

(compare page 93).

302

18. Suppose that the suspension point of a pendulum is moving horizontally,



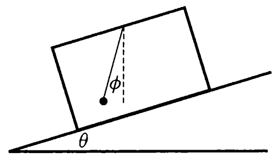
with its distance from the position at time 0 being $\alpha(t)$. Show that we now have

$$\theta'' + \frac{g}{l}\sin\theta = -\frac{\alpha''}{l}\cos\theta.$$

In particular, for an oscillating motion $\alpha(t) = a\cos kt$, we have

$$\theta'' + \frac{g}{l}\sin\theta = \frac{a}{l}k^2\cos kt.$$

19. A pendulum bob hangs without swinging in a car accelerating with constant acceleration a up a hill at angle θ.



(a) The total force on the bob, not counting the force in the direction of the string, is

$$G = \sqrt{g^2 + a^2 + 2ag\cos\theta},$$

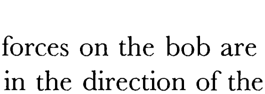
and the angle ϕ that the string makes with the vertical satisfies

$$\cos\phi = \frac{g + a\cos\theta}{G},$$

with $\phi < \theta$.

(b) Find the equation for the motion of the pendulum bob if it is set swinging.

291 **20.** A pendulum bob hangs at a constant angle $\theta > 0$ from the vertical, and moves along a horizontal circle with constant angular frequency α.



(a) Assuming that the only forces on the bob are the constant force of gravity downwards, and a force $\mathbf{F}_2$ in the direction of the string, of length l, conclude that

$$\cos\theta = \frac{g}{l\alpha^2}.$$

(b) This solution makes no sense for $\alpha < \sqrt{g/l}$. Note that in all cases, one possible solution is $\theta = 0$, i.e., the bob hangs straight down. For $\alpha > \sqrt{g/l}$, what happens when a bob moving in this way is perturbed slightly?

287 **21.** When our pendulum swings through a very small angle, so that θ in equation (P) is always very small, we use the fact that $\lim_{\theta \to 0} (\sin\theta)/\theta = 1$ to simplify our equation to

$$\text{(P}'\text{)} \qquad \theta'' + \frac{g}{l}\theta = 0,$$

obtaining an equation whose period we can easily determine [the accuracy with which we approximate the period is examined in Problem 23]. Setting $\omega = \sqrt{g/l}$, so that our equation becomes

$$\theta'' + \omega^2\theta = 0,$$

the solution is of the form $\theta(t) = \alpha \sin \omega t + \beta \cos \omega t$ for constants α and β. It is convenient to assume that the pendulum hangs straight down ($\theta = 0$) at time $t = 0$, so that we have

$$\theta(t) = \alpha \sin \omega t;$$

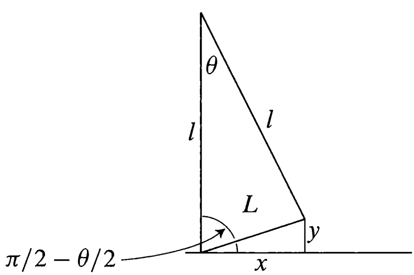
here α is clearly the maximum angle from the vertical that the pendulum reaches.

(a) Conclude that the period T of the pendulum is given by

$$T = 2\pi\sqrt{l/g}.$$

Since we can measure T very accurately by counting the number of periods through which the pendulum swings in a long time interval, we ought to be able to use this equation to find g accurately. In practice, however, it is impossible to determine the length of the string holding the pendulum with sufficient accuracy, since the string stretches a bit when the pendulum is set in motion, so we need something like a pendulum consisting of a solid rod, which we discuss in Chapter 6, with particular attention in Problem 6-3.

291 **22.** For later purposes it will be useful to consider what happens when we parameterize the position of the pendulum bob by x or y instead of θ.



(a) We have $y = L\sin\theta/2$, and $L = 2l^2 - 2l\cos\theta \approx 2l^2 - 2l$ for small θ. Conclude that for small oscillations, y satisfies the same equation as θ.
(b) We also have $x = l - \sqrt{l^2 - y^2}$, and it follows that x also satisfies the same equation.
Another approach is indicated in Problem 3-16.

23. Returning to the exact equation

$$\theta'' + \frac{g}{l}\sin\theta = 0,$$

multiply by θ' to conclude that

$$\theta'^2 = \frac{g}{l}\cos\theta + C$$

for a constant C, and then that

$$\theta'^2 = 2\omega^2(\cos\theta - \cos\alpha),$$

so that the period T is given by

$$T = \frac{2}{\sqrt{2}\,\omega}\int_{-\alpha}^{\alpha}\frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}}.$$

With a little work, this can be expressed in terms of standard elliptic integrals:

(a) Use the identity

$$\cos\theta - \cos\alpha = 2\left(\sin^2\frac{\alpha}{2} - \sin^2\frac{\theta}{2}\right)$$

to get

$$T = \frac{1}{\omega}\int_{-\alpha}^{\alpha}\frac{d\theta}{\sqrt{\sin^2\frac{1}{2}\alpha - \sin^2\frac{1}{2}\theta}}.$$

Now we are going to use the substitution

$$\sin\tfrac{1}{2}\theta = \sin\tfrac{1}{2}\alpha\cdot\sin x.$$

Noting that

$$\left(\sin^2\tfrac{1}{2}\alpha - \sin^2\tfrac{1}{2}\theta\right)^{1/2} = \sin\tfrac{1}{2}\alpha\cos x,$$

conclude that

$$T = \frac{4}{\omega}\int_0^{\pi/2}\frac{dx}{\sqrt{1 - \sin^2\frac{1}{2}\alpha\sin^2 x}}.$$

Setting

$$k = \sin \tfrac{1}{2}\alpha,$$

we have finally

$$T = 4\sqrt{l/g} \int_0^{\pi/2} \frac{dx}{\sqrt{1 - k^2 \sin^2 x}}$$

Note that this approaches $2\pi\sqrt{l/g}$ as $k \to 0$, i.e., for small α; on the other hand, it approaches ∞ as $k \to 1$, i.e., for $\alpha = \pi$, where the pendulum is swinging almost all the way to an upright position above the pivot point.

Expanding the term $(1 - k^2 \sin^2 x)^{-1/2}$ by the binomial theorem, we have

$$T = 2\pi\sqrt{l/g}\left(1 + \tfrac{1}{4}\sin^2 \tfrac{1}{2}\alpha + \tfrac{9}{64}\sin^4 \tfrac{1}{2}\alpha + \cdots\right).$$

For $\alpha \le 1\frac{1}{2}^\circ$, the first correction term is approximately 1 part in 20,000.

199 **24.** Although we won't be considering electricity and magnetism in any detail in this volume, there's one simple application of interest in regard to the basic laws, and we will refer to it later in Addendum 5A.

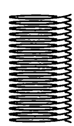
A magnet produces a "magnetic field" $\mathbf{B}$, which can be thought of as the force that the magnet produces on a small "test mass" of iron (like the iron filings used to show the "lines of force" around a magnet, which are the integral curves of the vector field $\mathbf{B}$).

The magnet also produces a force on a *moving* charged particle, of mass m and charge q, which depends upon the velocity $\mathbf{v}$ according to the so-called Lorentz Force Law, $\mathbf{F} = q(\mathbf{v} \times \mathbf{B})$ with proper choices of units for force and charge, so that the particle satisfies

$$m\mathbf{v}' = q(\mathbf{v} \times \mathbf{B}).$$

(If you sense something terribly wrong with this law, good for you! See Addendum 5A.)

Magnetic fields can also be created by currents, and we can get a constant magnetic field with a "solenoid", a long coil of wire with a current running



through it, which produces a magnetic field $\mathbf{B}$ that is nearly constant inside the coil, except near the ends. (Of course, the magnetic field of a magnet is also nearly constant at large distances from the magnet.)

(a) If $\mathbf{B}$ points along the third axis, with constant length B, then the components v_i of $\mathbf{v}$ satisfy

$$\left.\begin{aligned} v_1' &= \omega v_2 \\ v_2' &= -\omega v_1 \end{aligned}\right\} \quad \omega = qB/m \qquad \text{and} \quad v_3' = 0.$$

Using the obvious solution (compare the footnote on page 339)

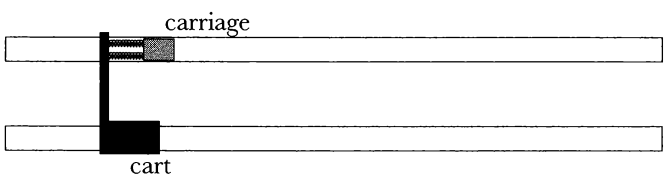
$$\begin{aligned} v_1 &= a\cos(\omega t) \\ v_2 &= a\sin(\omega t), \end{aligned}$$

show that the path of the particle is a helix with constant velocity. Notice that ω, and therefore the number of revolutions per second, is independent of both the initial velocity $|\mathbf{v}(0)|$ and the radius a of the orbit. It also depends on the ratio q/m of charge to mass, the basis for J. J. Thomson's experiments of 1897 to measure this ratio for the electron.

(b) The pitch h of the helix, the vertical distance gained through one revolution, is given by

$$h = \frac{2\pi m v_3(0)}{qB}.$$

25. In our discussion of Newton's proof of the parallelogram law for forces, we didn't mention that at the very least one has to assume that the law holds for *collinear* forces! In other words, we must appeal to experiment to conclude that "forces in the same direction are additive", remembering that our experimental cart is used to define the measurement of force, as on page 16: If our cart needs to be given the acceleration a_1 in order to compress the spring to length L_0 when some object is placed in the carriage of the air-trough, and we replace the single spring with two springs of the same construction, we find that the

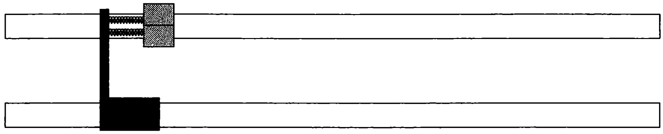


cart must be given the acceleration $\alpha_1 = 2a_1$ in order for the two spring to be compressed simultaneously to the length L_0. Equivalently, the mass m of our body can be determined by

$$m = \frac{a_0}{\alpha_1/2}$$

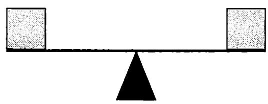
using two springs instead of one (a_0 being the convenient acceleration that we used on our "unit mass' to determine L_0 with just one spring).

By considering two copies of our object on the carriage, with a spring behind



each, conclude that the mass of this new object is $2m$ (the general rule for additivity of mass following in a fairly obvious way).

26. Accurate weighing basically requires a good "balance", which at heart is nothing more than a lever with equal arms; without knowing the law of the



lever, we can still say that two objects have equal weight if they balance. Just to be on the safe side, of course, we also switch their position to check that they still balance.

Once we've chosen some object for our "unit" weight, we can, for example, produce two new bodies that balance against each other, and that together

balance against the unit weight, each of which could then be called a half-unit weight. Similarly, we could create any number of $1/10$-unit weights, $1/100$-unit weights, etc., and then weigh any object to any desired degree of accuracy by balancing it against a suitable collection of the units.

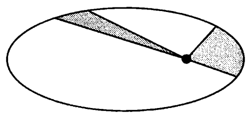
This procedure might make it appear that "additivity of weight", and thus "additivity of mass", has simply been declared by fiat. What implicit experimental results are involved?

CHAPTER 2

NEWTON'S ANALYSIS OF CENTRAL FORCES

Book 1 of the Principia is called The Motion of Bodies. The first section begins with geometric considerations that we would nowadays rephrase in terms of limits, but the second immediately begins with "Kepler's second law", though Newton does not mention Kepler's name in this regard.[1]

Kepler's second law, which was actually the first he discovered, says that the radius vector of a planet sweeps out equal areas in equal time, or equivalently, that the area swept out in time t is proportional to t.



Newton pointed out that this is a consequence of the fact that the gravitational force that the sun produces on the planet it always directed along the line from the planet to the sun, or equivalently, that the acceleration of the planet is always directed toward the sun—the specific magnitude of this force being irrelevant. Or as Newton expressed it

> Proposition 1. *The areas which bodies made to move in orbits describe by radii drawn to an unmoving center of forces lie in unmoving planes and are proportional to the times.*

This turns out to be extremely easy to prove analytically, especially if we use the cross-product of vectors. For simplicity assume that our force is always directed toward the origin O, and let c be any particle. We always have

$$\begin{aligned}(c \times \mathbf{v})' &= (c \times \mathbf{v}') + (c' \times \mathbf{v}) = (c \times \mathbf{v}') + (\mathbf{v} \times \mathbf{v}) \\ &= c \times \mathbf{v}',\end{aligned}$$

so if $\mathbf{v}'$ points along c, we just get $(c \times \mathbf{v})' = 0$, and consequently the relation

$$(*) \qquad c \times \mathbf{v} = \mathbf{w} \qquad \mathbf{w} \text{ a constant vector.}$$

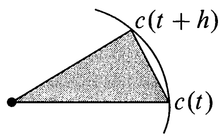
[1] Some discussion of this matter may be found in Cohen and Whitman [1; pg. 21], which throws light on numerous other questions of this sort.

If $\mathbf{w} = 0$, then $\mathbf{v}(t)$ always points along the line from O to $c(t)$, and our particle must simply be moving along a straight line towards O. If $\mathbf{w} \neq 0$, then, since the inner product satisfies

$$0 = \langle c(t) \times \mathbf{v}(t), c(t)\rangle = \langle \mathbf{w}, c(t)\rangle,$$

we see immediately that $c(t)$ always lies in one plane.

Moreover, $c(t) \times \mathbf{v}(t)$ has a natural interpretation in terms of the area swept out by the radius vector. For small h, this area, $S(t+h) - S(h)$, is approximately



the area of the shaded triangle, and thus approximately

$$\frac{1}{2}\big|c(t) \times [c(t+h) - c(t)]\big|.$$

Consequently, in the limit we have

$$\begin{aligned} S'(t) &= \frac{1}{2} \lim_{h \to 0} \left| c(t) \times \frac{c(t+h) - c(t)}{h} \right| \\ &= \frac{1}{2} |c(t) \times \mathbf{v}(t)|. \end{aligned}$$

Thus, $(*)$ implies that $S'(t)$ is constant, or that $S(t)$ is proportional to t.

This approximation argument isn't really necessary, for we we can simply write c as

$$c(t) = r(t) \cdot (\cos\theta(t), \sin\theta(t)),$$

compute $\mathbf{v} = c'$ and observe that

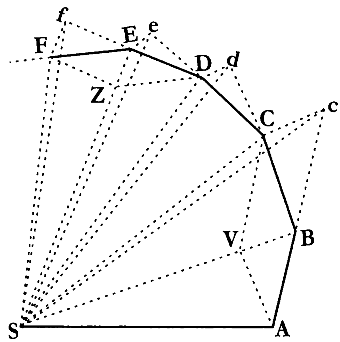
$$|c \times \mathbf{v}| = r^2\theta',$$

and $\frac{1}{2} \cdot r^2\theta'$ is just the integrand required to compute areas in polar coordinates. For future reference, we record the result

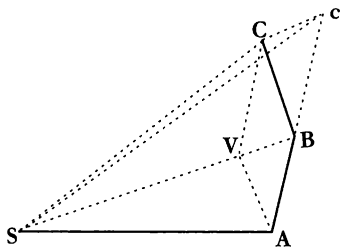
$$(*') \qquad r^2\theta' = |\mathbf{w}| = h, \text{ say.}$$

It might seem hard to improve upon this analytic proof, but Newton provides a geometric proof for Proposition 1, approximating the curve by a polygon, which is not only simple, but also seems to show just *why* the proposition is true.

As a sentimental gesture, we give a replica of the diagram that Newton uses in his proof



though we mainly concentrate on only a small portion of it:



Newton assumes that the particle follows the path $ABCD\ldots$, receiving "impulsive" forces at short equal intervals of time, and that these impulsive forces at B, C, $\ldots$ are always directed toward S, so that the path sweeps out the triangular areas $\triangle SAB$, $\triangle SBC$, $\ldots$. Newton merely has to point out that, if not for the impulsive force at B, the particle would move to c, with $Bc = AB$. In this case, it would sweep out the triangle $\triangle SBc$, which has the same area as $\triangle SAB$ (since they have equal bases, and the same height). The impulsive force applied at B will instead send the particle to C, which will be at the diagonal of the parallelogram formed by Bc and a line BV pointing along SB, since we are assuming that the force is directed toward S. This means that Cc is parallel to BV, and this in turn means that $\triangle SBc$ has the same area as $\triangle SBC$ (since these triangles have the common base SB and the same height above that base). In short, the area of $\triangle SAB$ is the same as the area of $\triangle SBC$, and so on, all along the path![1]

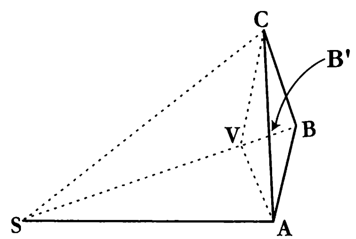
[1] Various details about the rigor of Newton's arguments may be found in Pourciau [4].

It is also noteworthy that Newton expressly states a converse of Proposition 1, the proof being pretty much the same, and one might not think it would be of much interest. In fact, Newton never mentions it again (even though, as we shall see later, it plays a crucial role):

> Proposition 2. *Every body that moves in some curved line described in a plane and, by a radius drawn to a point, ... describes areas around that point proportional to the times, is urged by a centripetal force tending toward that same point.*

Newton's proof of Kepler's second law is often presented in elementary physics books because of its dual virtues of simplicity and transparency. Unfortunately, this presentation usually marks the end of the exposition, with the lament that Newton's further investigations require many abstruse properties of conic sections which are unfamiliar to us nowadays. This turns out to be a double disappointment. While it is quite understandable that a geometric proof would use many geometric properties of conics, the real mystery is how an hypothesis about inverse square forces is going to be related to geometric properties of conic sections. Moreover, although we won't pursue Newton's argument in its entirety, it turns out that Newton's *strategy* for the proof is extremely clever.[1]

To see how Newton relates the forces to the geometry, we need only follow, with slight modifications, a few steps that he adds a bit later on. In the diagram for Proposition 1, breaking up the motion into small intervals δ of time, consider the segment BB' from B to the midpoint B' of the diagonal AC. This



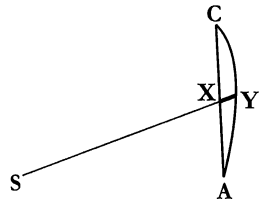
is half of BV, which represents the displacement due to the central force of magnitude F at B, and this distance is just $\frac{1}{2}F\delta^2$. So, in the limit,

$$\lim_{\delta\to 0} \frac{1}{\delta^2} BB' = \tfrac{1}{4}F.$$

Or as Newton phrases it, in inexact poetical-sounding terms, so much more beautiful than limits and epsilons and deltas,

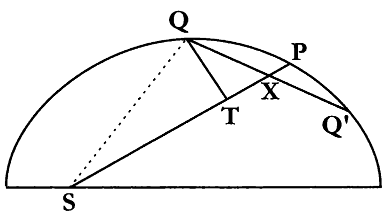
[1] So clever, in fact, that it has caused arguments up to recent times—see the remarks in the last section of Chapter 7.

> If ... a body revolves in any orbit about an immobile center and describes any just-nascent arc in a minimally small time, and if the sagitta of the arc is understood to be drawn so as to bisect the chord and, when produced, to pass through the center of forces, the centripetal force in the middle of the arc will be as the sagitta directly and as the time twice inversely.



It shouldn't be necessary to explicate the old-fashioned term sagitta [Latin for arrow], because in this instance Newton's statement explicitly indicates that he is referring to the segment XY of the line through S and the midpoint X of the chord AC. The fraction $\frac{1}{4}$ doesn't appear in Newton's statement because the result is phrased as a proportion: the ratio of the centripetal forces at points A and A' is the same as the ratio of the limits $\lim_{\delta\to 0} XY/\delta^2$ for arcs starting at A and A', respectively.

Newton next considers a point P at time t on a curved path around the center S, and two points Q, Q' on the path, at two nearby times $t-\delta$ and $t+\delta$.



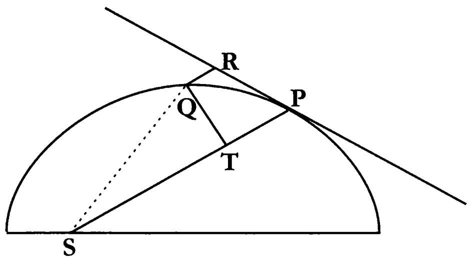
Then the force at P is proportional to

$$\lim_{\delta\to 0} \frac{PX}{\delta^2}.$$

But, by Proposition 1, δ is proportional to the area of the (curved) triangle SPQ, and thus, in the limit, to $QT \times SP$, where QT is perpendicular to SP. Thus, finally, the force at P is proportional to

$$\lim_{\delta\to 0} \frac{PX}{(SP)^2 \times (QT)^2}.$$

Newton, however, actually presents a figure that has a tangent line drawn at P, and the line QR drawn parallel to SP, with the assertion that the force



at P is proportional to

$$(*) \qquad \lim_{\delta \to 0} \frac{QR}{(SP)^2 \times (QT)^2}.$$

Of course, QR is not actually equal to PX, but it is apparently obvious to Newton that it is equal to second order so that the limit still holds.[1]

And now Newton is all prepared to show that the orbit of an object moving under an inverse square force is a conic section. Newton begins in a way that might seem strange to us, by proving a partial *converse* of this assertion:

> *Let a body revolve in an ellipse; it is required to find the law of the centripetal force tending toward a focus of the ellipse.*

In other words, given a path c lying along an ellipse, if c'' always points towards one focus of the ellipse, Newton is going to show that

$$|c''(t)| = \frac{k}{d(t)^2} \quad \text{for some constant } k,$$

where $d(t)$ is the distance from $c(t)$ to the focus.

The relation $(*)$ is the key to all this. In fact, in view of $(*)$, our assertion is equivalent to saying that for an ellipse we have

$$\lim_{\delta \to 0} \frac{QR}{(QT)^2} \quad \text{is a constant.}$$

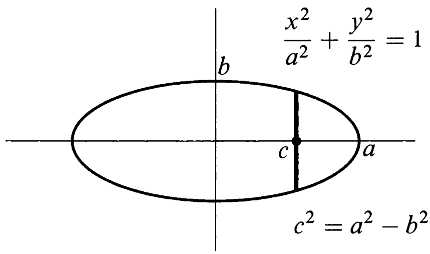
This limit has nothing to do with forces, and is completely determined by the shape of the ellipse. It could even be computed by a double application of

[1] For a more exact statement, and its somewhat intricate proof, see Pourciau [5].

L'Hôpital's Rule: If $F(\delta)$ denotes QR and $G(\delta)$ denotes $(QT)^2$, then we have $\lim_{\delta \to 0} F'(\delta) = \lim_{\delta \to 0} G'(\delta) = 0$, and

$$\lim_{\delta \to 0} \frac{QR}{(QT)^2} = \frac{F''(0)}{G''(0)};$$

when the unpleasant calculation (Problem 7) is carried through, it turns out that $F''(0)/G''(0)$ is independent of the point P, and in fact $= a/2b^2$ for the ellipse shown below; the reciprocal, $2b^2/a$, is the length of the classical *latus rectum* of the ellipse, the segment cut off by the ellipse on the vertical line through one of the foci.



Newton proves exactly this result geometrically, and the proof is indeed long, complicated, and depends on numerous results about the ellipse. For a complete exposition of this proof see Newton [pp. 325–330].

Newton then gives a similar proof for a body moving on a hyperbola, and finally a proof for a body moving on a parabola.

And immediately afterwards, the result we really wanted appears as a corollary:

> COROLLARY 1. From the last three propositions it follows that if any body P departs from the place P along any straight line PR with any velocity whatever and is at the same time acted upon by a centripetal force that is inversely proportional to the square of the distance of places from the center, this body will move in some one of the conics having a focus in the center of forces; and conversely.

In the first edition of the Principia this is *all* that appears—the result is claimed to be a corollary of its converse(s)—but in the second edition Newton added a few sentences to aid the reader (with a bit of rewording in the third edition):

COROLLARY 1. From the last three propositions it follows that if any body P departs from the place P along any straight line PR with any velocity whatever and is at the same time acted upon by a centripetal force that is inversely proportional to the square of the distance of places from the center, this body will move in some one of the conics having a focus in the center of forces; and conversely. For if the focus and the point of contact and the position of the tangent are given, a conic can be described that will have a given curvature at that point. But the curvature is given from the given centripetal force and velocity of the body; and two different orbits touching each other cannot be described with the same centripetal force and the same velocity.

Here, in slightly different terms, is a complete argument, where for simplicity we simply work in $\mathbb{R}^2$, and we choose the origin O as the point toward which the force is directed. Given a point P, and a tangent vector $\mathbf{v}$ at P, we want to find a curve $c = (c_1, c_2)$ with $c(0) = P$ and $c'(0) = \mathbf{v}$ satisfying

$$(*) \qquad c''(t) = \frac{k}{|c(t)|^2} \cdot \frac{-c(t)}{|c(t)|},$$

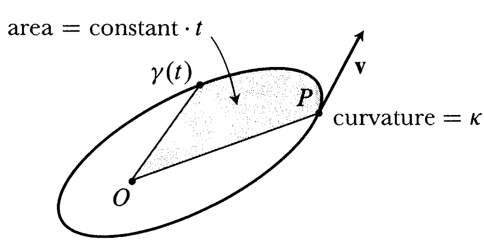
where k is a given constant (the factor $-c(t)/|c(t)|$ is just a unit vector pointing from $c(t)$ to the origin).

Since we know $c'(0)$, and $(*)$ gives us $c''(0)$, we know what the curvature κ of c at 0 should be, since this is given by

$$(**) \qquad \kappa = \frac{c_1{}'(0)c_2{}''(0) - c_2{}'(0)c_1{}''(0)}{(c_1{}'^2(0) + c_2{}'^2(0))^{3/2}},$$

a formula known, in essence, to Newton, though curvature was defined in terms of the osculating circle.

Now consider a conic section K having O as a focus, which passes through P, and is tangent to $\mathbf{v}$ at P, and whose curvature at P is this κ, assuming for the moment that such a conic section exists. Consider a curve γ with $\gamma(0) = P$, which traverses K in such a way that the areas cut out by radii from O is proportional to the time. Such curves are determined up to a multiplicative



change of parameter; by choosing the appropriate multiplicative constant, we can arrange for $\gamma'(0) = \mathbf{v}$. According to our converse Proposition 2, which Newton carefully provided, we have

$$\gamma''(t) = \frac{\bar{k}}{|\gamma(t)|^2} \cdot \frac{-\gamma(t)}{|\gamma(t)|}$$

for some $\bar{k}$. But we must have $k = \bar{k}$, since we chose γ so that its curvature at $\gamma(0) = P$ would be the κ given by $(**)$.

Thus, γ is a solution of our differential equation $(*)$, and by uniqueness (which of course Newton and all his contemporaries implicitly assumed) it is the only possible solution.

Newton's argument might strike us as a little weird, starting as it does with the converses of the result we want, but a *geometric* proof almost has to be of this nature: it's a lot easier to start with a geometric object—an ellipse, or hyperbola, or parabola—and deduce a formula for forces, then it would be to start with the formula for forces and somehow conjure up these geometric figures.

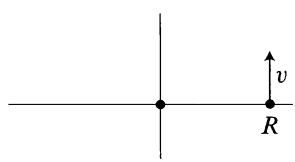
The only slight lacuna in the argument is the existence of a conic section with the required curvature, but Newton gives a detailed geometric solution to this problem a little later on in Proposition 17 of Book 1 (a somewhat unfortunate shuffling of the proper expository order). Analogously, there is the analytic problem of describing the solutions of $(*)$ in terms of the initial values. We will defer this problem, and other aspects of planetary motion, to Chapter 4, where we give a connected presentation of all the relevant material.

However, just to get some idea how the orbits are determined by the initial conditions, we will examine qualitatively the special case where our initial velocity $\mathbf{v}$ is perpendicular to the radius vector to the point. In other words, we consider the case where we start at a vertex of the conic section (more specifically, in the case of an ellipse, we are starting at the vertex at the end of the major axis).

We assume that our force is exerted toward the origin O, with magnitude

$$\frac{1}{|c(t)|^2} \qquad \text{(taking the constant } k \text{ as 1 for simplicity)}$$

and that we are looking for a solution c through a fixed point $c(0) = (R, 0)$, with an initial velocity $(0, v)$ of magnitude v. In this case, we easily compute



that the curvature κ of c at 0 given by (∗∗) on page 62 is

$$\kappa = \frac{1}{v^2 R^2},$$

so we have to find appropriate conics with varying κ, where small κ correspond to large v and *vice versa*.

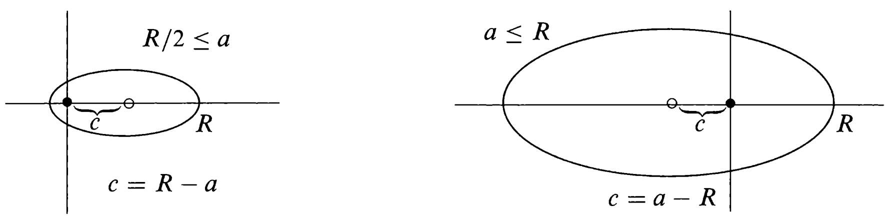
Using the parameterization

$$c(t) = (a\cos t, b\sin t)$$

of an ellipse with axes $a > b$, we easily compute that the curvature κ at the end of the major axis is

$$\kappa = \frac{a}{b^2}.$$

There are two classes of ellipses with a vertex at $(R, 0)$ and a focus at O, depending on whether $(R, 0)$ and O are on opposite sides of the center of the



ellipse, or are on the same side. In both cases we have

$$(a - R)^2 = c^2 = a^2 - b^2,$$

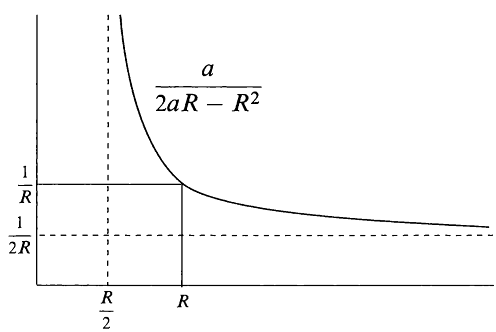
which gives

$$b^2 = 2aR - R^2,$$

so that the curvature at the end of the major axis is

$$\kappa = \frac{a}{2aR - R^2} \qquad a \ge R/2$$

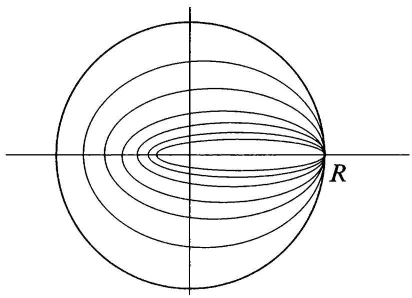
whose graph is shown below.



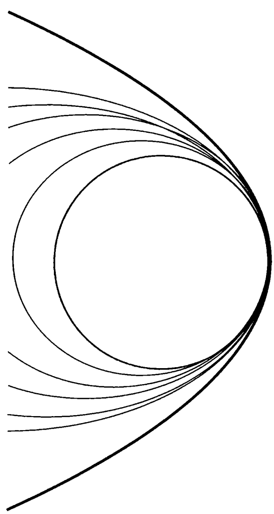
For $a = R$ we have $\kappa = 1/R$ which is a circle, corresponding to

$$\frac{1}{R} = \frac{1}{v^2 R^2}$$

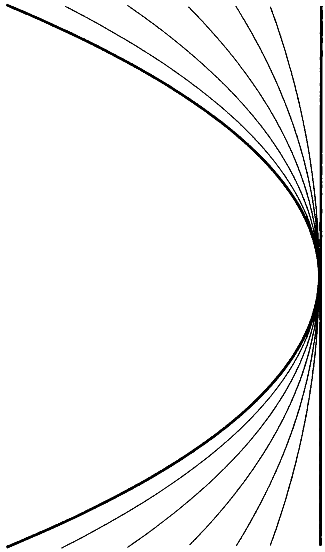
or $v = \sqrt{1/R}$. As we make v smaller we obtain a family of ellipses from the first class, which converge to a straight line from our point $(R, 0)$ to the origin.



As we make v larger, we obtain a family of ellipses from the second class, whose curvatures at the vertex approach $1/2R$, and these ellipses approach



a parabola with curvature $1/2R$ at its vertex. Still larger values of v give a family of hyperbolas approaching the vertical line through $(R, 0)$, which is the



orbit when v is "infinitely large", or equivalently, when the gravitational force is negligible.

PROBLEMS

117 1. (a) Right after the proof of Proposition 2, stated on page 58, Newton inserted a very simple result, concerning circular orbits (which the elliptical orbits of the planets very closely approach). The orbit can be described by

$$c(t) = r(\cos\alpha t, \sin\alpha t),$$

as on page 18, where we computed $c''(t)$. Correlate this with Newton's statement of the result (Newton [2; pg. 449]):

> *The centripetal forces of bodies that describe different circles with uniform motion tend toward the centers of those circles and are to one another as the squares of the arcs described in the same time divided by the radii of the circles.*

(b) Conclude that for two such orbits, the values of $|c''(t)|$ (and thus of the forces causing the motion) are inversely as the squares of the radii if and only if the periods are as the $3/2$ powers of the radii.

This is Kepler's third law, and Newton interjects a brief Scholium stating that this

> "holds for the heavenly bodies (as our compatriots Wren, Hooke, and Halley have also found out independently). Accordingly, I have decided that in what follows I shall deal more fully with questions relating to the centripetal forces that decrease as the squares of the distances from centers ... ".

Later on, after proving that inverse square forces give elliptical orbits for the planets, Newton proves the general case of Kepler's third law, that the periods are as the $3/2$ powers of the major axes. We will consider this result in Chapter 4.

2. It appears[1] that during the Plague Years of 1665–1666, Newton tested the inverse square law for the case of the moon revolving around the earth.

It was know, from triangulation, that the distance from the center of the earth to [the center of] the moon was close to 60 times the radius of the earth. For the length of a degree of latitude at the surface of the earth (i.e., $1/360$ of the earth's

[1] For details, suppositions, caveats, etc., see Chapter 4 of Herivel [1], as well as the discussion in Cohen and Whitman [1; pp. 67–70].

circumference) Newton took the common estimate at that time of 60 miles, the mile at that time apparently taken as 5,000 feet. Thus, the circumference of the earth was to be reckoned as $1{,}080 \times 10^5$ feet, the radius as $171{,}887 \times 10^2$ feet, and the distance from the center of the earth to the moon as $1{,}031{,}322 \times 10^3$ feet. The period of the moon is very close to 27 days and 8 hours, or 39,360 minutes.

(a) Assuming that the moon travels with uniform velocity in a circle around the earth, compute its velocity, and then use Problem 1-5 (a result known to Newton early on, presumably by something like the geometric proof in that problem) to compute that its acceleration towards the center of the earth is about 26.27 feet/min^2, or .0072972... feet/sec^2.
(b) Using the fact that the acceleration of bodies near the earth's surface is 32 feet/sec^2 (this was known from pendulum observations), compute that the acceleration of the moon towards the center of the earth should be about .00888... feet/sec^2.

In 1682, the French astronomer Picard announced a more accurate measurement of 69.1 miles for the length of a degree of latitude, which gave Newton a much better correlation, carried out in detail in Book 3 of the Principia, Newton [2; pp. 803–804]. Although the importance of this correlation was mentioned in Chapter 1, Newton actually tackled much more detailed problems in Book 3, a large part of which was devoted to showing how the intricate details of the moon's motion could be derived from the principles of gravitational physics. For an extremely instructive discussion of these matters, see Cohen and Whitman [1; pp. 246–264] (as well as other sections of Chapter 8 of that book's Guide).

After proving the relation ($*$) on page 60, Newton does not immediately tackle the problem of planetary orbits, but instead devotes the remainder of the section to other problems that indicate how ($*$) may be used. His first example is one that his methods handle quite nicely:

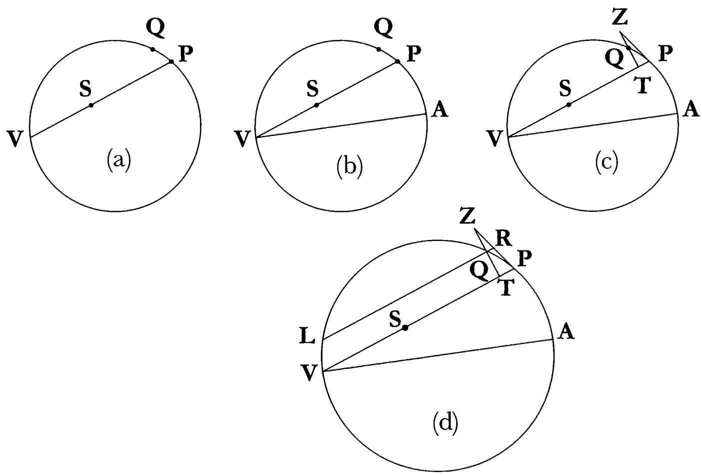
161 **3.** We want to find what central force at a point S will cause a body to move on the circumference of a circle when S is *not* the center of the circle. Newton presents a rather complicated figure that is most easily understood by considering it in steps, shown in the diagram on the next page:

(a) We consider a point P on the circle, and a nearby point Q, and draw the line from P to S, intersecting the circle at V.

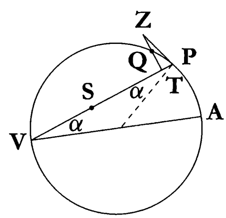
(b) We then draw the diameter VA.

(c) We next draw the line TQ perpendicular to SP, and extend it until meets the tangent line to the circle at P in the point Z.

(d) Finally, we draw the line LR parallel to VP intersecting the circle at L and the tangent line at R.

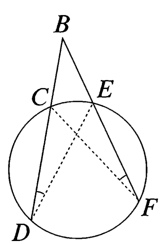


In (c) we also draw the radius to P, which is perpendicular to ZP and con-



clude that $\angle TZP = \alpha$, so that the right triangle ZTP is similar to the right triangle VPA

We also want to recall the elementary geometry theorem that for two secant lines through a point, we have the relation $BC \cdot BD = BE \cdot BF$, and the obvious



consequence when one of the secant lines is actually a tangent line ($E = F$).

(a) We have

$$\frac{RP}{QT} = \frac{ZP}{ZT} = \frac{AV}{PV},$$

so

$$\frac{QR \cdot RL}{(QT)^2} = \frac{(AV)^2}{(PV)^2},$$

and hence

$$\frac{(SP)^2 \cdot (QT)^2}{QR} = RL \cdot \frac{(SP)^2 \cdot (PV)^2}{(AV)^2}.$$

(b) Thus,

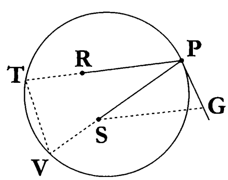
$$\lim_{Q \to P} \frac{(SP)^2 \cdot (QT)^2}{QR} = \frac{(SP)^2 \cdot (PV)^3}{(AV)^2},$$

and consequently by equation (∗) on page 60, the force is inversely proportional to $(SP)^2 \cdot (PV)^3$.

This result appears in the first edition of the Principia, but three remarkable corollaries first appear in the second edition:

(c) (Corollary 1) If a particle moves in a circle under a central force directed to a point V on the circle, then the force varies inversely as the fifth power of the distance. (And, since this is true for any point V on the circle, the exact same force law with two different centers can give the same orbits.)

(d) More generally, suppose that a particle moves on a circle under two different central forces, one with center at R and one with center at S. Let SG be drawn



parallel to RP, intersecting the tangent line to the circle at P in the point G. Then the ratio of the first force to the second is

$$\frac{(RP)^2 \cdot (PT)^3}{(SP)^2 \cdot (PV)^3} = \frac{SP \cdot (RP)^2}{\dfrac{(SP)^3 \cdot (PV)^3}{(PT)^3}}.$$

Conclude that the ratio of the first force to the second is

(Corollary 2) $$\frac{(RP)^2 \cdot SP}{(SG)^3}.$$

(e) By considering the osculating circle to the orbit at P, conclude that the same results hold for an *arbitrary* orbit: If a body moves on the same orbit under central forces directed toward R and S, then the ratios of the forces is again

(Corollary 3) $$\frac{(RP)^2 \cdot SP}{(SG)^3},$$

where SG is parallel to RP and meets the tangent line to the orbit at P in the point G.

4. After these corollaries, and a result about an inverse third power law that we will consider in Chapter 4, Newton considers elliptical orbits under a central force directed *toward the origin*, rather than toward a focus. His treatment of this problem is almost as complicated as his treatment of the inverse square law, but we easily compute that a particle moving under the orbit

$$c(t) = (a\cos\alpha t, b\sin\alpha t)$$

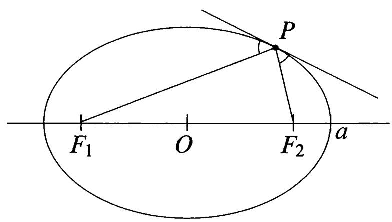
has $c''(t) = -\alpha^2 c(t)$, so that it is a possible motion under a central force varying directly as the distance to the origin, and obviously the same holds for any ellipse centered at the origin.

(a) Newton also states the converse as a corollary, again without any further indication of the proof. For an analytic proof, consider a particle $c(t) = (x(t), y(t))$ moving under a central force directed toward the origin, of magnitude $k^2 r$ at distance r. This gives the equations $x'' + k^2 x = 0$ and $y'' + k^2 y = 0$, so that x and y are each linear combinations of $\cos kt$ and $\sin kt$, say

$$(x, y) = (\cos kt, \sin kt)\begin{pmatrix} A & C \\ B & D \end{pmatrix}.$$

Solve for $\cos kt$ and $\sin kt$ in terms of x and y, and use $\cos^2 + \sin^2 = 1$, to obtain an equation for a conic centered at the origin, which is necessarily an ellipse, since x and y are bounded ("elliptic harmonic motion").

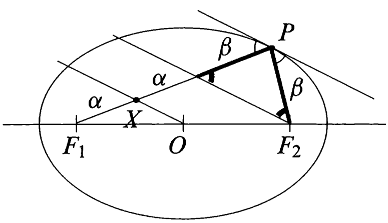
(b) Now comes the really interesting part, which also first appears in the second edition. We will need another property of the ellipse, but it is one that we can easily derive from familiar ones. Recall that the ellipse with major axis $2a$ and foci F_1 and F_2 is defined by the property that $F_1P + PF_2 = 2a$ for all points P.



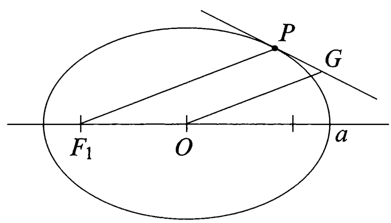
(a)

It also has the "focal point" property that a light ray starting from F_1 passes through F_2, i.e., that the two angles indicated in (a) are equal.

In diagram (b) below we have drawn lines parallel to the tangent line at P through O and F_2. The two angles indicated by thick arcs are equal, so the



(b)



(c)

two thick segments have the same length β. And the two segments with lengths indicated as α are equal because $F_1O = OF_2$. Thus

$$2\alpha + 2\beta = 2a.$$

Finally (c), moving XP over to OG we see that a line through the origin O parallel to the line F_1P always intersects the tangent line through P at a point G with $OG = a$.

Use this result together with part (a) and Corollary 3 in the previous problem to give an alternate proof that if a particle moves in an ellipse under a central force directed toward a focus, then the force must vary inversely as the square of the distance from the focus.

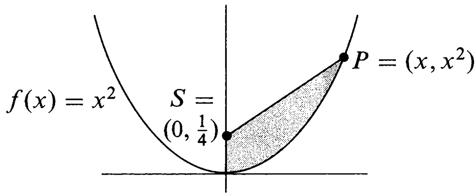
Addendum 4E considers in greater detail some of the questions raised by these observations of Newton.

5. Later on in the Principia Newton analyzes orbits much more thoroughly, essentially determining not only their shape, but also their parameterization, even allowing himself a formula or two now and then; as we will note later on,

the results correspond to standard modern formulas involving integrals. But first he disposes of the question of parabolic orbits:

> *If a body moves in a given parabolic trajectory, to find its position at an assigned time.*

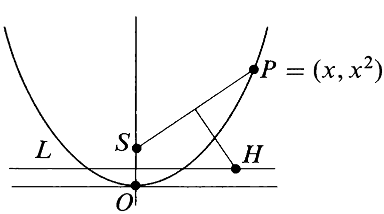
The graph of $f(x) = x^2$ is a parabola with focus at $(0, \frac{1}{4})$. Compute the shaded area in the figure on the next page, as a cubic expression in x, and



conclude that we can find the parameterization of a planet moving along this parabola, in terms of solutions of cubic equations.

The case of parabolic orbits is very special in this regard, because attempts to find the corresponding areas for elliptical orbits involve elliptic integrals, which cannot be expressed in elementary terms. Amazingly enough, Newton actually proves that we can't expect to find algebraic parameterizations in the case of an ellipse, and his argument constitutes, in the words of Arnold [3; pp. 83 ff.] "an astonishingly modern topological proof of a remarkable theorem on the transcendence of Abelian integrals". The remarks in Arnold are expanded somewhat in Chandrasekhar [2; pp. 133 ff.], scholastic grousing can be found in Cohen and Whitman [1; pp. 138–139], and a comprehensive account of the problem can be found in Pourciau [3].

By the way, Newton analyzed the motion along parabolic orbits purely geometrically (see Chandrasekhar [2; pp. 130–131]): If the line L is the perpendicular bisector of SO, and the perpendicular bisector of SP intersects L at H, then, as Chandrasekhar says, "With these constructions (which passes understanding) Newton proves" that the point H moves along L with a uniform velocity equal to $3/8$ of the velocity of the planet at O!



121 **6.** This problem gives a rather *ad hoc* solution of the equations of motion for a particle under a radially symmetric central force. A more systematic derivation is presented in Chapter 4, but this problem is important not only as preparation for that derivation, but as an introduction to other techniques.

Recall that for a particle c moving under a central force, the vector $c \times \mathbf{v}$ is a constant vector $\mathbf{w}$, so $|c \times \mathbf{v}|$ is the constant

$$(1) \qquad |c \times \mathbf{v}| = h, \qquad h = |\mathbf{w}|.$$

We write the equation of our particle as

$$c(t) = r(t) \cdot (\cos\theta(t), \sin\theta(t)),$$

so that we have

$$r^2 = \langle c, c \rangle,$$

and consequently

$$rr'' + r'^2 = \langle c, \mathbf{v}' \rangle + |\mathbf{v}|^2.$$

Since our force is radially symmetric, i.e., it depends only on the distance from O, we have

$$m\mathbf{v}' = -(f \circ r) \cdot \frac{c}{r}$$

for some function f (with the $-$ sign added so that a positive f corresponds to an attractive force). Then

$$\langle c, \mathbf{v}' \rangle = -\frac{r}{m} \cdot (f \circ r),$$

and consequently

$$(2) \qquad rr'' + r'^2 = -\frac{r}{m}(f \circ r) + |\mathbf{v}|^2.$$

(a) Show that for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ we have

$$\langle \mathbf{x}, \mathbf{y} \rangle^2 + |\mathbf{x} \times \mathbf{y}|^2 = |\mathbf{x}|^2 \cdot |\mathbf{y}|^2.$$

(You can either give a geometrically obvious explanation, or reduce to the case $\mathbf{x} = (1,0,0)$ and $\mathbf{y} = (y_1, y_2, 0)$ and calculate.)
(b) Apply this result to eliminate $\mathbf{v}$ from (2) and conclude that

$$r'' - \frac{h^2}{r^3} = -\frac{1}{m}(f \circ r),$$

where h is given by (1). Consequently,

$$\left(r'^2 + \frac{h^2}{r^2}\right)' = -\frac{2}{m}(f \circ r)r' = -\frac{2}{m}(F \circ r)',$$

where F is a primitive of f, i.e., $F' = f$. Thus, for an appropriate F we will have

$$(*) \qquad r'^2 = -\frac{2}{m}(F \circ r) - \frac{h^2}{r^2}.$$

(c) We thus have a differential equation for r. We also have the equation

$$(*') \qquad r^2\theta' = h$$

(page 56), so knowing r gives a differential equation for θ. However, we seldom expect to solve for r or θ directly. For example, in the case of an inverse square law, with an ellipse as solution, finding r or θ as a function of t would essentially involve finding the areas of sectors of an ellipse, which involves elliptic functions. Usually we concentrate on finding the *shape* of the solution, rather than its particular parameterization, by finding $r \circ \theta^{-1}$.

From the equation

$$(r \circ \theta^{-1})' = r' \circ \theta^{-1} \cdot \frac{1}{\theta' \circ \theta^{-1}},$$

which in the infinitely flexible Leibnizian notation can be written as

$$\frac{dr}{d\theta} = \frac{dr}{dt} \bigg/ \frac{d\theta}{dt},$$

equations $(*)$ and $(*')$ yield the differential equation

$$\frac{dr}{d\theta} = \frac{r^2}{h}\sqrt{-\frac{2}{m}(F \circ r) - \frac{h^2}{r^2}}.$$

We can also "separate variables" to obtain

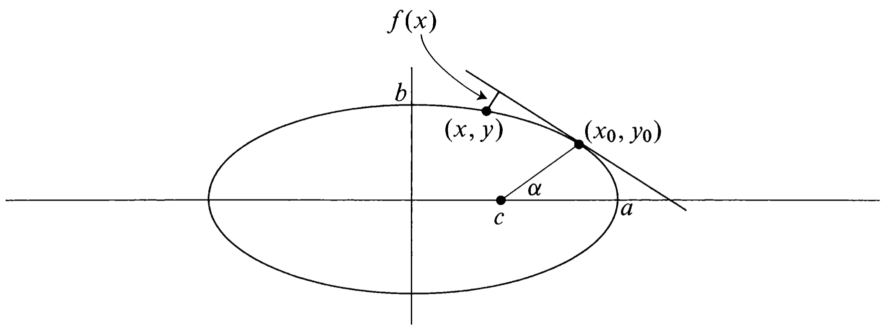
$$d\theta = \frac{h\,dr}{r^2\sqrt{-\dfrac{2}{m}(F \circ r) - \dfrac{h^2}{r^2}}},$$

so that

$$\theta(r) = \pm \int_{r_0}^{r} \frac{h\,dr}{r^2\sqrt{-\dfrac{2}{m}(F \circ r) - \dfrac{h^2}{r^2}}}.$$

(Only the r in the limit of integration on the right side is of significance—the other r's could just as well be replaced by any other "variable of integration", but notation of this sort, consistently used by physicists and others solving differential equations, is more convenient.) This determines $\theta \circ r^{-1}$ [the expression $\theta(r)$ means $\theta(r^{-1}(t))$], and thus $r \circ \theta^{-1}$. The extent to which specific formulas can be written down depends on our ability to determine the integral in elementary terms; we will explore this question further in Chapter 4.

7. Consider the ellipse $x^2/a^2 + y^2/b^2 = 1$, whose upper half is the graph of $x \mapsto (b/a)(a^2 - x^2)^{1/2}$. The focus is the point $(c, 0)$ with $c^2 = a^2 - b^2$. We



consider the tangent line L at a point (x_0, y_0) and the *perpendicular* distance $f(x)$ from L to the point $(x, y) = (x, (b/a)(a^2 - x^2)^{1/2})$ on the ellipse.

(a) Recall that the distance from the straight line with equation $y = Mx + B$ to the point (C, D) is

$$\frac{|CM - D + B|}{\sqrt{M^2 + 1}}.$$

So, if the slope of L is ν, then

$$\sqrt{\nu^2 + 1}\, f(x) = x\nu - y + \text{constant} = x\nu - \frac{b}{a}(a^2 - x^2)^{1/2} + \text{constant}.$$

Conclude that

$$\sqrt{\nu^2 + 1}\, f''(x_0) = ba(a^2 - x_0{}^2)^{-3/2},$$

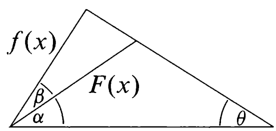
and then, using the value of ν, that

$$f''(x_0) = ba(a^2 - x_0{}^2)^{-3/2}\left[\frac{b^2}{a^2}\frac{x_0{}^2}{a^2 - x_0{}^2} + 1\right]^{-1/2}$$

$$= ba(a^2 - x_0{}^2)^{-3/2}\left[\frac{a^4 - c^2 x_0{}^2}{a^2(a^2 - x_0{}^2)}\right]^{-1/2}$$

$$= \frac{ba^2}{(a^2 - x_0{}^2)(a^4 - c^2 x_0{}^2)^{1/2}}.$$

(b) Let $F(x)$ be the length of the line from (x, y) to L that is *parallel* to the line from the focus $(c, 0)$ to (x_0, y_0), so that $F(x) = f(x)\sec\beta$.



Using

$$\tan\theta = \frac{b}{a}x_0(a^2 - x_0{}^2)^{-1/2},$$

we have

$$\tan(\alpha + \beta) = \frac{a}{bx_0}(a^2 - x_0{}^2)^{1/2},$$

while

$$\tan\alpha = \frac{y_0}{x_0 - c} = \frac{b}{a}\frac{(a^2 - x_0{}^2)^{1/2}}{x_0 - c}.$$

Conclude that

$$\tan\beta = \frac{\tan(\alpha+\beta) - \tan\alpha}{1 + \tan(\alpha+\beta)\tan\alpha}$$

$$= \frac{(a^2 - x_0{}^2)^{1/2}\left[\dfrac{a^2 x_0 - a^2 c - b^2 x_0}{abx_0(x_0 - c)}\right]}{\dfrac{a^2 - x_0 c}{x_0(x_0 - c)}}$$

$$= \frac{(a^2 - x_0{}^2)^{1/2}\cdot\dfrac{c(cx_0 - a^2)}{ab}}{a^2 - x_0 c}$$

$$= -\frac{(a^2 - x_0{}^2)^{1/2}c}{ab}.$$

Compute that

$$\sec\beta = \frac{(a^4 - c^2{x_0}^2)^{1/2}}{ab},$$

and then that

$$F''(x_0) = \frac{a}{a^2 - {x_0}^2}.$$

(c) Now let

$$\mu = \frac{y_0}{x_0 - c} = \frac{b}{a}\frac{(a^2 - {x_0}^2)^{1/2}}{x_0 - c}$$

be the slope of the line between the focus $(c, 0)$ and (x_0, y_0). For the distance $g(x)$ from (x, y) to this line we have

$$g'(x) = \left(\frac{b}{a}\frac{(a^2 - {x_0}^2)^{1/2}}{x_0 - c} + \frac{bx}{a}(a^2 - x^2)^{-1/2}\right) \cdot (\mu^2 + 1)^{-1/2}$$

and for $G(x) = g(x)^2$ we have

$$G''(x_0) = 2g'(x_0)^2 = \frac{2b^2}{a^2}\frac{(a^2 - cx_0)^2}{(x_0 - c)^2(a^2 - c_0{}^2)} \cdot \frac{1}{\mu^2 + 1}.$$

Conclude, finally, that

$$G''(x_0) = \frac{2b^2}{a^2 - {x_0}^2}.$$

CHAPTER 3

CONSERVATION LAWS

In this chapter we will look at the basic "conservation" laws of mechanics, one of which, conservation of momentum, has already been briefly mentioned in Chapter 1. The considerations of that chapter might lead us to consider a "system of particles" consisting of

(a) certain particles $c_1, \dots c_K : \mathbb{R} \to \mathbb{R}^3$,

(b) with positive masses $m_1, \dots, m_K \in \mathbb{R}$,

(c) functions $\mathbf{F}_i^e : \mathbb{R} \to \mathbb{R}^3$,

(d) functions $\mathbf{F}_{ij} = -\mathbf{F}_{ji} : \mathbb{R} \to \mathbb{R}^3$,

with the following basic property: If we set

$$\mathbf{F}_i = \mathbf{F}_i^e + \sum_j \mathbf{F}_{ij},$$

then

$$\mathbf{F}_i = m_i \cdot c_i{}''.$$

Here $\mathbf{F}_i^e(t)$ represents an "external force" on the particle c_i at time t, while the $\mathbf{F}_{ij}(t)$ represent "internal forces" between $c_i(t)$ and $c_j(t)$, satisfying Newton's third law, and consequently $\mathbf{F}_i(t)$ represents the total force on the particle c_i at time t. For forces satisfying the "strong version" of Newton's third law (page 25), condition (d) should also stipulate that $\mathbf{F}_{ij}(t)$ is a multiple of $c_i(t) - c_j(t)$.

Of course, in practice, the $\mathbf{F}_i^e$ and $\mathbf{F}_{ij}$ will often have simple expressions in terms of other functions. For example, the external force $\mathbf{F}_i^e(t)$ might be of the form $m_i \cdot \mathbf{f}(c_i(t))$ for a vector field $\mathbf{f}$ on $\mathbb{R}^3$, e.g., the gravitational attraction due to some external body, while the internal forces $\mathbf{F}_{ij}$ might be some function of m_i, m_j and the distance $|c_i - c_j|$.

Our more general definition allows all sorts of more complicated situations, for example one where the force $\mathbf{F}_i^e$ depends not only on t and $c_i(t)$ but on the whole collection of $\{c_j(t)\}$. A simple instance would be a system of many particles representing a space ship with rocket propulsion, where the direction of the force depends on the particular angle at which the space ship is rotated at any particular time. (Presumably, the space ship has more than one rocket, so that it can steer).

Conservation of momentum. Although conservation of momentum, as stated in Chapter 1, involved only internal forces, we can easily state a generalization allowing for external forces. We set $\mathbf{F} = \sum_i \mathbf{F}_i^e$, the total external force.

1. PROPOSITION (MOMENTUM LAW). The derivative of the total momentum is the total external force,

$$\mathbf{F} = \Big(\sum_i m_i \cdot \mathbf{v}_i\Big)'.$$

Here we have to regard the various $\mathbf{F}_i^e$ simply as elements of $\mathbb{R}^3$, rather than as tangent vectors at different points of $\mathbb{R}^3$, and similarly for the $\mathbf{v}_i$.

This formulation gains considerable significance when we introduce the concept of the **center of mass** of the system $\{c_i\}$, which represents the "average" position of the particles c_i weighted according to their masses:

$$C = \frac{\sum_i m_i \cdot c_i}{\sum_i m_i}.$$

More precisely, we should define the center of mass as the particle consisting of the path C with the mass $M = \sum_i m_i$.

If all $\mathbf{F}_i^e = 0$, so that $\sum_i m_i \cdot \mathbf{v}_i$ is constant, then $C'' = \frac{1}{M}\sum_i m_i \cdot c_i'' = \frac{1}{M}\sum_i m_i \cdot \mathbf{v}_i' = \frac{1}{M}\big(\sum_i m_i \cdot \mathbf{v}_i\big)'$, so that we also have $C'' = 0$. Thus, C' is constant; in other words, the center of mass moves with uniform velocity.

More generally, we have

2. PROPOSITION. If $\mathbf{F} = \sum_i \mathbf{F}_i^e$ is the total external force, then

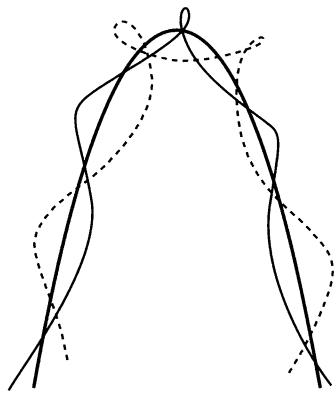
$$\mathbf{F} = M \cdot C'',$$

so that the center of mass particle simply moves as if it were acted upon by the total force $\mathbf{F}$.

PROOF. We have

$$\begin{aligned} M \cdot C'' &= \sum_i m_i \cdot c_i'' \\ &= \sum_i \mathbf{F}_i \\ &= \sum_i \mathbf{F}_i^e + \sum_i \sum_j \mathbf{F}_{ij} \\ &= \sum_i \mathbf{F}_i^e. \ ❖ \end{aligned}$$

Although the "particle" C might not be one of the particles in our system, this result is seldom regarded as particularly "theoretical"—instead it allows us to get a very simple picture of very complex phenomena. For example, in the figure on page 10, showing a rod executing a complicated revolving motion, the center of mass, which does happen to be a point on the rod in this case, simply moves in a parabola, just like a point mass. A striking illustration may be obtained with a time-exposure photograph taken when a baton is tossed in the air, with lights at the ends and the center of mass, giving a picture like this:

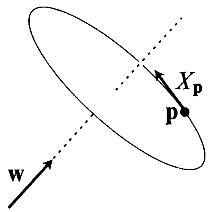


We usually think of a rod as a "rigid body", a concept whose analysis we have still shied away from. At first, that might seem to make the result even more impressive: in a real rod, with all sorts of complicated intermolecular forces, which make it approximately "rigid", but not truly so, it is still true that the center of mass moves according to a simple law. But that is a somewhat misleading way of construing the result, since rigidity of the rod is required in order to identify its center of mass with a particular point of the rod, on which we can attach one of the lights.

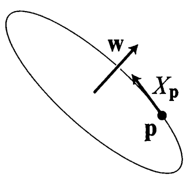
Center of mass is often called "center of gravity", a concept that goes back at least to Archimedes (cf. the Prologue). These concepts are not identical except in a uniform gravitational field—which applies, of course, to reasonable sized objects on the earth's surface—but the difference is often ignored.

Conservation of angular momentum. The use of the cross-product $\times$ at the beginning of the previous chapter could be regarded simply as a convenient abbreviation for manipulations with determinants. But there is a more important reason why this special product of $\mathbb{R}^3$ is significant.

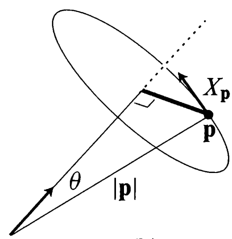
For any vector $\mathbf{w} \in \mathbb{R}^3$, consider the one-parameter family of maps $B(t)\colon \mathbb{R}^3 \to \mathbb{R}^3$, where $B(t)$ is a counter-clockwise rotation through an angle of $t|\mathbf{w}|$ radians around the axis through $\mathbf{w}$ [choosing an orientation $(\mathbf{v}_1, \mathbf{v}_2)$ of the plane perpendicular to $\mathbf{w}$ so that $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{w})$ is the usual orientation of $\mathbb{R}^3$]. Now consider the vector field generated by this one-parameter family. In other words, for each $\mathbf{p} \in \mathbb{R}^3$ consider the curve $B_{\mathbf{p}}(t) = B(t)(\mathbf{p})$, and then look at the tangent vector $X_{\mathbf{p}}$ of this curve at 0.



To compute $X_{\mathbf{p}}$ geometrically, we note that $X_{\mathbf{p}}$ is clearly perpendicular to both $\mathbf{p}$ and $\mathbf{w}$. Its length is also easy to determine. When $\mathbf{p}$ happens to lie in the plane perpendicular to $\mathbf{w}$, as in (a), the point $\mathbf{p}$ rotates in a circle of radius $|\mathbf{p}|$, and $X_{\mathbf{p}}$ has length $|\mathbf{p}| \cdot |\mathbf{w}|$. More generally (b), the point $\mathbf{p}$ rotates in a circle of radius $|\mathbf{p}| \cdot |\mathbf{w}| \cdot \sin\theta$, where θ is the angle between $\mathbf{w}$ and $\mathbf{p}$. Thus, $X_{\mathbf{p}}$ is just the geometrically defined cross-product $\mathbf{w} \times \mathbf{p}$.



(a)



(b)

For an analytic determination of $X_{\mathbf{p}} = B_{\mathbf{p}}'(0)$, we note that since the $B_{\mathbf{p}}(t)$ are all orthogonal, and $B_{\mathbf{p}}(0) = I$, the derivative $B_{\mathbf{p}}'(0)$ is skew-adjoint, with a skew-symmetric matrix M, which we will write in the form

$$M = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

Then the vector $X_{\mathbf{p}}$ is the 3-tuple whose transpose $X_{\mathbf{p}}{}^{\mathbf{t}}$ is given by

$$\begin{aligned} X_{\mathbf{p}}{}^{\mathbf{t}} &= M \cdot (p_1, p_2, p_3)^{\mathbf{t}} \\ &= \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \cdot (p_1, p_2, p_3)^{\mathbf{t}} \\ &= (-p_2\omega_3 + p_3\omega_2, -p_3\omega_1 + p_1\omega_3, -p_1\omega_2 + p_2\omega_1)^{\mathbf{t}}. \end{aligned}$$

Setting $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$, we then have $X_{\mathbf{p}} = \boldsymbol{\omega} \times \mathbf{p}$. Moreover, $\boldsymbol{\omega}$ is easy to identify, because

$$B_{\mathbf{w}}(t) = \mathbf{w} \text{ for all } t \implies 0 = X_{\mathbf{w}} = \boldsymbol{\omega} \times \mathbf{w},$$

so $\boldsymbol{\omega}$ must be a multiple of $\mathbf{w}$, and it easy to check, by considering some specially chosen vector, that in fact $\boldsymbol{\omega} = \mathbf{w}$. Thus, one might say that the cross-product $\times$ is special to $\mathbb{R}^3$ because $n = 3$ is the only dimension where $\mathrm{O}(n)$ has dimension n. More to the point, we have

PROPOSITION. The vector fields in $\mathbb{R}^3$ generated by rotations about an axis are of the form $\mathbf{p} \mapsto \boldsymbol{\omega} \times \mathbf{p}$ for $\boldsymbol{\omega} \in \mathbb{R}^3$.

For a particle c with velocity vector $\mathbf{v}$ we can consider the function $c \times \mathbf{v}$ from $\mathbb{R}$ to $\mathbb{R}^3$, which is called the **angular velocity** of the particle; if $c(t) = (x(t), y(t), z(t))$ for functions x, y, and z, then the angular velocity of c is

$$(\mathrm{A}_c) \qquad (yz' - y'z,\ x'z - xz',\ xy' - x'y).$$

For a particle whose mass is m, the cross-product $\mathbf{L} = c \times m\mathbf{v}$ is called its **angular momentum**. The angular velocity and momentum just defined are "with respect to the origin 0": for any other point P, the angular momentum with respect to P is the cross-product

$$\mathbf{L}_P = (c - P) \times m\mathbf{v}.$$

For a system of particles $(c_1, \dots, c_K)$ we define the angular momentum $\mathbf{L}$ of the system, with respect to 0, as

$$\mathbf{L} = \sum_{i=1}^{K} c_i \times m_i \mathbf{v}_i;$$

here it is naturally necessary to consider all $c_i \times m_i \mathbf{v}_i$ as vectors at a single point, rather than as tangent vectors at different points. Note that the equation $\mathbf{L}' = \sum_{i=1}^{K} (c_i \times m_i c_i{}')'$ reduces to

$$(\mathrm{L}') \qquad \mathbf{L}' = \sum_{i=1}^{K} c_i \times m_i c_i{}''.$$

More generally, we define the angular momentum $\mathbf{L}_P$ with respect to P as $\mathbf{L}_P = \sum_{i=1}^{K}(c_i - P) \times m_i\mathbf{v}_i$. In particular, suppose we take P to be the center of mass C of the system (this means that we may be considering the angular momentum with respect to different points at different times). Letting $M = \sum_i m_i$, the "mass" of the particle C, we then have

$$\begin{aligned}\sum_i m_i c_i \times \mathbf{v}_i &= \sum_i m_i(c_i - C) \times \mathbf{v}_i \ + \ \sum_i m_i C \times \mathbf{v}_i \\ &= \mathbf{L}_C + \left[C \times \left(\textstyle\sum_i m_i\mathbf{v}_i\right)\right] \\ &= \mathbf{L}_C + [C \times MC'],\end{aligned}$$

so that we can write

$$\mathbf{L} = \mathbf{L}_C + (C \times MC').$$

The vector $\mathbf{L}_C$, the angular momentum with respect to the center of mass, is also called the "rotational angular momentum", so our equation says that the total angular momentum $\mathbf{L}$ is the sum of the rotational angular momentum $\mathbf{L}_C$ and the angular momentum of the center of mass with respect to 0.

If instead of a momentum vector we consider an arbitrary force $\mathbf{F}$ at a point c, the cross-product

$$\boldsymbol{\tau} = c \times \mathbf{F}$$

is called the **torque** of the force with respect to 0, while $\boldsymbol{\tau}_P = (c - P) \times \mathbf{F}$ is the torque with respect to P. (Although I have used the physicists' $\mathbf{L}$ for angular momentum, I couldn't bring myself to use the standard $\mathbf{N}$ for torque.)

Similarly, we define the torque of a system of forces on a system of particles; here it is again necessary to consider the individual torques as being vectors at one point, even though we naturally think of the forces as being applied at different points.

3. PROPOSITION (ANGULAR MOMENTUM LAW). If our system satisfies the strong form of the third law, then the total torque is the derivative of the total angular momentum,

$$\boldsymbol{\tau} = \mathbf{L}'.$$

PROOF. We have

$$\begin{aligned}\mathbf{L}' &= \sum_i c_i \times m_i c_i'' \qquad \text{by equation } (\mathrm{L}') \\ &= \sum_i c_i \times \mathbf{F}_i \\ &= \sum_i c_i \times \mathbf{F}_i^e + \sum_i \sum_j c_i \times \mathbf{F}_{ij}\end{aligned}$$

$$= \boldsymbol{\tau} + \sum_i \sum_j c_i \times \mathbf{F}_{ij}.$$

The strong form of the third law allows us to write

$$\mathbf{F}_{ij} = \lambda_{ij}(c_i - c_j),$$

with $\lambda_{ij} = \lambda_{ji}$, so we have

$$\sum_i \sum_j c_i \times \mathbf{F}_{ij} = \sum_i \sum_j \lambda_{ij}[c_i \times c_i \quad - \quad c_i \times c_j],$$

which vanishes, since $c_i \times c_i = 0$, while $c_i \times c_j = -c_j \times c_i$ and $\lambda_{ij} = \lambda_{ji}$. ❖

Easy manipulations give us the more general

4. COROLLARY. For any point P,

$$\boldsymbol{\tau}_P = \mathbf{L}_P{}'$$

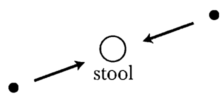
In particular, of course, if the torque is 0, then angular momentum is conserved. This certainly happens in the special case of a single particle moving under a *central force*, where the external force $\mathbf{F}$ is a multiple of c, so that $\boldsymbol{\tau} = c \times \mathbf{F} = 0$. This was noted by Newton as the first Corollary of his Proposition 1 (page 55):

> COROLLARY 1. In nonresisting spaces, the velocity of a body attracted to an immobile center is inversely as the perpendicular dropped from that center to the straight line which is tangent to the orbit.

As we saw in Chapter 2, the particle actually stays in a plane, and if we have $c(t) = (x(t), y(t), 0)$, say, then conservation of angular momentum just says, by equation (A_c), that $xy' - yx'$ is constant. Even the somewhat more general rule that angular momentum is conserved in the absence of external forces was not stated until quite some time afterwards, and this law was known for a long time simply as "the law of areas", or *Flächensatz* in German.

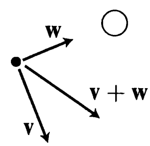
The evocative term "torque" (from the Latin *torquere*, to twist) was not introduced until the 19$^{\text{th}}$ century. Before that, the cross-product $c \times \mathbf{F}$ was called the **moment** of the force $\mathbf{F}$ at the point c, with respect to 0. Here "moment" is being used in the sense of "importance" or "significance" (e.g., a matter of great moment), this significance having been noted long before in terms of the law of the lever. Correspondingly, angular momentum was known as the "moment of momentum", a term which has not yet been totally expunged.

A standard elementary illustration of the law of conservation of angular momentum is provided by a person seated on a rotating stool with arms extended out holding weights, and then increasing the speed of the spin, often quite dra-



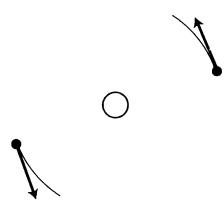
matically, simply by pulling the weights inward. Similarly, ice-skaters speed up their turns by pulling their arms in; divers, starting their dive with a small angular momentum, do rapid somersaults by pulling their arms and knees in; and gymnasts do all sorts of tricks.

By the way, without appealing to conservation of angular momentum we can explain the speed-up as a simple consequence of the parallelogram rule for forces, or even for velocities: the sum of the velocity $\mathbf{v}$ that the weight already has and the velocity $\mathbf{w}$ that it acquires as a result of the inward pull is the



diagonal of the rectangle spanned by these two, and consequently has a greater length.

In these examples, we merely altered the given non-zero angular momentum, but something interesting occurs even when we start with angular momentum 0. Moving the weights along a circle in one direction contributes a certain amount of angular momentum to the system of weights-plus-person, which must be



countered by an opposite amount of angular momentum in the system, so the seated person must rotate in the opposite direction. At the end of the motion, when the weights are no longer being rotated, the person will have stopped rotating, but will be facing in a different direction; cats use this mechanism to land on their paws even when dropped from an upside-down position.

In this respect, rotation is quite different from linear motion. A system cannot change its position using only internal forces, and no external forces. On a perfectly frictionless ice surface you can change the direction in which you are facing, but you can't move the position of your center of mass. (Of course, you can forcefully exhale, providing yourself with rocket propulsion, making use of the fact that the air inside your lungs is a part of your system that you aren't attached to—or you could just throw your coat away.)

The momentum law and the angular momentum law are the first two great conservation laws of mechanics, and they apply to all mechanical systems, although their application may also require further understanding about rigid bodies and other matters. We should also mention that they are vector equations, so, for example, if the x-component of the total external force $\mathbf{F}$ is 0, then the x-component of the total momentum is constant; or to put it more generally, if the total external force $\mathbf{F}$ is 0 in one particular direction, then the total momentum in that direction is also constant. This probably doesn't seem particularly useful, but the analogue for angular momentum definitely can be (cf. Problem 5): if the the torque is some direction is 0, then the component of the angular momentum in that direction is constant.

The third conservation law is quite different: it is both much more special and much more general.

Conservation of energy: kinetic and potential energy. As Galileo had noted, for a body falling under the acceleration of gravity, the distance s that it travels after being released from rest satisfies $s = at^2$ for some a, so that $v = s' = 2at$ satisfies $v^2 = \text{constant} \cdot s$, expressing v in terms of the distance traveled, rather than in terms of the time traveled.

More generally, suppose a body falls due to a force that depends only on its height x from the earth's surface,

$$mx''(t) = -f(x(t)),$$

for some function f; as in Problem 2-6, we add the $-$ sign so that a positive f corresponds to an attractive force *towards* the earth's surface, the direction in which x *decreases*. Although we may not be able to solve for x explicitly, we can still get out information about $v = x'$. We use the obvious trick of multiplying both sides of the above equation by $x'(t)$, so that the right side becomes a derivative,

$$mx'(t)x''(t) = -f(x(t)) \cdot x'(t) = -(F \circ x)'(t) \qquad \text{for } F' = f,$$

and then observe that the left side is also a derivative, so that we get

$$\left(\tfrac{1}{2}mx'^2\right)' = -(F \circ x)'.$$

The quantity $T = \frac{1}{2}mv^2$ is called the **kinetic energy** of the body, so if we let $x_i = x(t_i)$ for $i = 0, 1$ and $v_i = v(t_i) = x'(t_i)$, we have

$$(*) \qquad T(t_1) - T(t_0) = \tfrac{1}{2}m{v_1}^2 - \tfrac{1}{2}m{v_0}^2 = -F(x_1) + F(x_0),$$

so that the difference $T(t_1) - T(t_0)$ of the kinetic energy at two times depends only on the heights at the two times.

As an aside, we point out that an alternative approach is suggested if we know that we are looking for an expression for v in terms of x. As on page 75, we use Leibnizian notation to transform $mv'(t) = -f(x(t))$ into a formula for dv/dx:

$$\begin{aligned}
\frac{dv}{dx} &= \frac{dv}{dt}\Big/\frac{dx}{dt} = \frac{-f(x)}{mv} \\
-f(x) &= mv\frac{dv}{dx} = \frac{1}{2}m\frac{dv^2}{dx} \\
\tfrac{1}{2}mv^2 &= \int f(x)\,dx.
\end{aligned}$$

We've been discussing a one-dimensional situation, or equivalently one in which our force always points in one direction, but the same conclusion holds for the more general case of a radially symmetric central force. We introduce polar coordinates (r, θ) for the plane in which the motion takes place, and let $\mathbf{r}$ be the unit vector field pointing toward the origin, while $\boldsymbol{\theta}$ is the perpendicular unit vector field. At any point x our radially symmetric central force has the value $-f(|x|)\mathbf{r}$ for some function f. If for convenience we introduce the usual "abuse of notation" of allowing r to stand for $r \circ c$ and θ to stand for $\theta \circ c$, then

$$\begin{aligned}
\tfrac{1}{2}m(v^2)' &= \tfrac{1}{2}m \cdot \langle \mathbf{v}, \mathbf{v}\rangle' = m\langle \mathbf{v}, \mathbf{v}'\rangle = \langle \mathbf{v}, m\mathbf{v}'\rangle \\
&= \langle \mathbf{v}, -(f \circ r)\mathbf{r}\rangle \\
&= \langle r'\mathbf{r} + \theta'\boldsymbol{\theta}, -(f \circ r)\mathbf{r}\rangle \\
&= -(f \circ r)r' = -(F \circ r)' \qquad \text{for } F' = f.
\end{aligned}$$

We thus have

$$(**) \qquad T(t_1) - T(t_0) = -F(r(t_1)) + F(r(t_0)),$$

so that the difference in kinetic energy at two times depends only on the distances from the origin at the two times.

Nowadays this result is usually stated rather differently. If we let $V\colon \mathbb{R}^3 \to \mathbb{R}$ be the function

$$V = F \circ r,$$

then $(**)$ becomes

$$\begin{aligned} T(t_1) - T(t_0) &= -F(r(c(t_1))) + F(r(c(t_0))) \\ &= -V(c(t_1)) + V(c(t_0)), \end{aligned}$$

and if we choose a fixed t_0 we find that

$$(***) \qquad T(t) + V(c(t)) \text{ is constant.}$$

The quantity $V(p)$ is called the **potential energy** of the particle at $p \in \mathbb{R}^3$, so equation $(***)$, called **conservation of energy**, states that for radially symmetric central forces the sum of the kinetic energy and the potential energy of a particle is constant throughout its path. The important point here is that $V(c(t))$ depends only on the position $c(t)$ of the particle, not on the path c itself.

Obviously, the function V is only determined up to a constant. For elementary problems involving free falling bodies near the surface of the earth, it is customary to consider V to be 0 on the earth's surface, so that its value when the body is released from some height is positive. As the body falls, its potential energy decreases as its kinetic energy increases. This accords with the usual interpretation of V as the kinetic energy that the body "potentially" has, i.e., the kinetic energy that it can acquire by being released, and allowed to fall to earth.

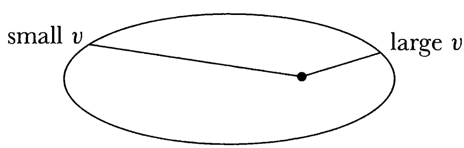
In the case of an inverse square force, a body falling radially toward the center, with distance $r(t)$ from the center given by

$$r''(t) = -K/r(t)^2,$$

has

$$v(t) = r'(t) = K/r \implies V(p) = -\frac{K}{r(p)} + \text{constant}.$$

It is convenient to take the constant to be 0, so that V is 0 at ∞. For a planet moving in an ellipse, V is larger (though negative) at points further from the sun, so the kinetic energy is smaller there (as implied by Kepler's second law).



More generally, we define a force $\mathbf{F} = (F_1, F_2, F_3)$ to be **conservative**, with **potential energy** function V, if the conservation of energy equation

$$\text{(C)} \qquad \tfrac{1}{2}m\langle \mathbf{v}(t), \mathbf{v}(t)\rangle + V(c(t)) = \text{constant}$$

holds for all particles $c(t)$ moving under the force $\mathbf{F}$. For the standard coordinate system (x^1, x^2, x^3) on $\mathbb{R}^3$, differentiating gives

$$\begin{aligned} 0 &= \langle \mathbf{v}(t), m\mathbf{v}'(t)\rangle + \frac{d}{dt}V(c(t)) \\ &= \langle \mathbf{v}(t), \mathbf{F}(c(t))\rangle + \sum_{i=1}^{3} \frac{\partial V}{\partial x^i}(c(t)) \cdot {c_i}'(t). \end{aligned}$$

Choosing a path c with $c(0) = p$, and evaluating at $t = 0$, we obtain

$$0 = \langle \mathbf{v}(0), \mathbf{F}(p)\rangle + \sum_{i=1}^{3} \frac{\partial V}{\partial x^i}(p) \cdot {c_i}'(0).$$

Since, under the standard identification of tangent vectors of $\mathbb{R}^3$ with $\mathbb{R}^3$ itself, we also have

$$\mathbf{v}(0) = ({c_1}'(0), {c_2}'(0), {c_3}'(0)),$$

we see, by choosing c with only one ${c_i}'(0) \neq 0$, that we must have

$$\text{(C}'\text{)} \qquad \mathbf{F} = (F_1, F_2, F_3) = -\left(\frac{\partial V}{\partial x^1}, \frac{\partial V}{\partial x^2}, \frac{\partial V}{\partial x^3}\right),$$

which physicists usually write as

$$\mathbf{F} = -\operatorname{grad} V.$$

Equivalently,

$$\left\langle \mathbf{F}, \frac{\partial}{\partial x^i}\right\rangle = F_i = -\frac{\partial}{\partial x^i}(V),$$

and thus, more generally, for any tangent vector $\mathbf{v}$ we have

$$\langle \mathbf{F}, \mathbf{v}\rangle = -\mathbf{v}(V),$$

for the usual operation of a tangent vector $\mathbf{v}$ on a function.

Conversely, suppose that **F** satisfies (C′). For any field $\mathbf{F} = (F_1, F_2, F_3)$ we have

$$\begin{aligned} T(t_1) - T(t_0) &= \tfrac{1}{2}m{v_1}^2 - \tfrac{1}{2}m{v_0}^2 \\ &= \int_{t_0}^{t_1} \langle \mathbf{v}(t), \mathbf{F}(c(t))\rangle \, dt \\ &= \int_{t_0}^{t_1} \sum_{i=1}^{3} c_i{}'(t) \cdot F_i(c(t)), \end{aligned}$$

and if we introduce the 1-form

$$\omega = F_1\, dx^1 + F_2\, dx^2 + F_3\, dx^3$$

and let γ be the curve $\gamma = c|[t_0, t_1]$, this can be written as

$$(*) \qquad T(t_1) - T(t_0) = \int_\gamma \omega.$$

Now if **F** satisfies (C′), then $\omega = -dV$, so we have

$$\begin{aligned} T(t_1) - T(t_0) &= \int_\gamma -dV \\ &= -V(c(t_1)) + V(c(t_0)), \end{aligned}$$

which implies that **F** is conservative, with potential function V.

The quantity

$$\int_{t_0}^{t_1} \langle \mathbf{v}(t), \mathbf{F}(c(t))\rangle \, dt = \int_\gamma \omega$$

is called the **work** done by the force **F** on the particle c as it moves along the path γ. We have just seen that for conservative forces this depends only on the end-points of the path. As a simple example, consider a closed elliptical path, on a time interval $[0, T_0]$, of a particle moving under an inverse square force **F**. The total work done by **F** along this path must be 0, since that is the total work done on the interval $[0, 0]$, which has the same end-points.

In general, of course, there usually won't be more than one trajectory between two points. The more interesting situation—and the one that connects with our every-day notion of work— arises when we consider the work done by a force as we move it along some other path, i.e., the work that has to be done *against* the force field **F** in order to move a particle from one point to another. For example, raising a particle of mass m from height h_0 to height h_1 near the earth's surface requires a total work of $mg(h_1 - h_0)$, no matter what path we take, provided that periods during which the particle is moving downward instead of upwards are regarded as contributing negative work.

This property of conservative forces is usually regarded as the more important one in physics books, so they define a force $\mathbf{F}$ to be conservative if $\int_\gamma \omega$ depends only on the end-points of γ. This immediately implies our previous definition, since we can define

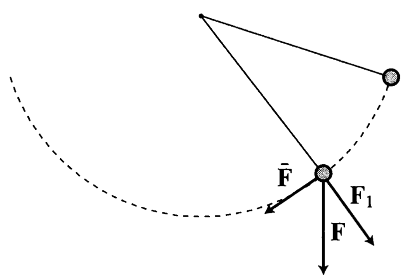
$$V(p) = \int_\gamma \omega$$

where γ is any path from a fixed point p_0 to p, and equation $(*)$ on page 91 immediately leads to the conservation of energy equation[1]

$$\text{(C)} \qquad \tfrac{1}{2}\langle \mathbf{v}(t), \mathbf{v}(t)\rangle + V(c(t)) = \text{constant}.$$

Looking at the calculation for equation $(**)$ on page 88 we see that it still holds if we replace $\mathbf{F}$ by $\mathbf{F} + \mathbf{F}_1$ where $\mathbf{F}_1$ is always perpendicular to $\mathbf{v}$; as physicists would express it, the work done by the extra force $\mathbf{F}_1$ is 0, by the very hypothesis that it is always perpendicular to $\mathbf{v}$.

In particular, instead of a body moving under the gravitational force of the earth, consider one suspended by a thread, so that we have a pendulum bob, which we have analyzed (somewhat informally) in Problem 1-17. The total force

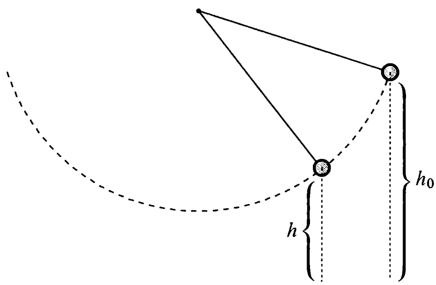


on the pendulum bob is then $\bar{\mathbf{F}} = \mathbf{F} - \mathbf{F}_1$, where $\mathbf{F}$ is the conservative force from the gravitation field of the earth, while $\mathbf{F}_1$ is the force that was introduced in our analysis, pointing along the thread. Since $\mathbf{F}_1$ points along the thread, $\langle \mathbf{v}, \mathbf{F}_1\rangle = 0$, so we still have the conservation of energy (C).

---

[1] For some strange reason, physics books (and even mathematics books) usually eschew this simple direct argument, instead noting that the dependence of $\int_\gamma \omega$ only on the end-points of γ implies that $\int_\gamma \omega = 0$ for all closed γ, so that for all 2-chains σ we have, by Stokes' theorem, $0 = \int_{\partial\sigma} \omega = \int_\sigma d\omega$, and thus we must have $d\omega = 0$. This implies (with proper conditions on the region where $\mathbf{F}$ is defined) that $\omega = -dV$ for some V.

This has the interesting consequence that although we cannot explicitly solve the pendulum equation derived on page 47, we can still say what the speed v



of the bob is at any height h, because we have

$$mgh + \tfrac{1}{2}mv^2 = \text{constant},$$

so we just have to know the height h_0 at which we released the bob, with $v = 0$ (compare Problem 1-17). The pendulum can be regarded as a mechanism that is continually interchanging potential energy and kinetic energy. At the top of the swing the kinetic energy is 0, while at the bottom of the swing, the difference in potential energy has been converted to kinetic energy, just sufficient to raise it up to the same height at which it started.

Similarly, on page 30 we mentioned the usual elementary analysis of a block sliding down an inclined plane, where we assume that the block is acted upon by the force of gravity $\mathbf{F}$ and another force $-\mathbf{F}_1$ perpendicular to the inclined plane. Thus this argument works in that case also, and the kinetic energy $\frac{1}{2}mv^2$ at the bottom must again be gh. So the speed of the block when it reaches the bottom must be the same as if it fell straight down, which agrees with our calculations, since the block's acceleration along the inclined plane is only $\sin\alpha$ of its falling acceleration, but it has $1/\sin\alpha$ as far to go. In the same way, instead of a pendulum bob hanging from a string, we could just as well allow the bob to slide along a plane with a circular profile, or indeed any profile, if we could really provide a frictionless surface.

Aside from its obvious physical interest, conservation of energy is important mathematically as a "first integral" of the laws of motion, i.e., an equation involving only first derivatives, rather than the second derivatives that appear in Newton's laws—all harking back to our original trick on page 87. In the next chapter, we will derive the result of Problem 2-6 in a more systematic way, starting from conservation of energy, with the sign of the total energy E of an orbit turning out to have a simple geometric significance.

Conservation of energy in collisions. While the role of kinetic energy with respect to conservative forces seems fairly straightforward, there was initially considerable confusion about kinetic energy because of the completely different role that it plays in that simplest, yet most essential, physical phenomenon, the collision of two bodies.

Consider two particles, c_1 and c_2, with masses m_1 and m_2, moving along a straight line with velocities v_1 and v_2; as usual, since our motion is confined to a straight line, we can represent the velocities simply by numbers. It seems natural to ask the question: if they collide, what are their new velocities w_1 and w_2 after the collision?

Conservation of momentum gives us only one equation,

$$(1) \qquad m_1w_1 + m_2w_2 = m_1v_1 + m_2v_2,$$

for the two unknowns w_1 and w_2, so it obviously can't determine an answer to the question, even under special circumstances, like the case where $m_1 = m_2$ and $v_2 = 0$, so that we have a moving object colliding with a stationary one of the same mass. One possible solution would be $w_1 = 0$ and $w_2 = v_1$, so that the first body stops and imparts all its motion to the second (something close to this happens when two steel balls collide). On the other hand, the second body might be "soft", like clay, so that it yields on impact, losing its shape and adhering to the first body, with the two then moving together as one (alternatively, we might consider bodies that will stick after contact because of glue, as on page 26, or perhaps carts with couplings that cause them to move as one after an impact), and in this case the final velocity of the two bodies will simply be $v_1/2$, just another possible solution of the infinitely many.

Elementary physics textbooks need to provide problems that have answers, of course, so, in the manner of a host nonchalantly introducing a celebrity at a party, they will often unobtrusively insert a new definition: a collision is called "completely elastic", if we also have conservation of kinetic energy,

$$(2) \qquad \begin{cases} m_1{w_1}^2 + m_2{w_2}^2 = m_1{v_1}^2 + m_2{v_2}^2 \\ \text{or} \\ m_1({w_1}^2 - {v_1}^2) = m_2({v_2}^2 - {w_2}^2) \end{cases}$$

Consorting with this new definition we have a contrasting one: a collision is "completely inelastic" if $w_1 = w_2$ (the two bodies stick together).

Once we've made a definition, it's possible to pose all sorts of simple problems about collisions that are "completely elastic", whatever that might mean. In general, writing (1) in the from

$$m_1(v_1 - w_1) = m_2(w_2 - v_2),$$

and dividing into [the second form of] (2), which is permissible so long as we don't have $w_1 = v_1$ (and $w_2 = v_2$), we get

$$(3) \qquad w_1 - w_2 = -(v_1 - v_2),$$

and solving (1) and (3) for the unknowns w_1 and w_2 gives

$$(*) \qquad \begin{aligned} w_1 &= \frac{m_1 - m_2}{m_1 + m_2} v_1 + \frac{2m_2}{m_1 + m_2} v_2 \\ w_2 &= \frac{2m_1}{m_1 + m_2} v_1 - \frac{m_1 - m_2}{m_1 + m_2} v_2; \end{aligned}$$

the other solution, $w_1 = v_1$ and $w_2 = v_2$, is discarded on physical grounds, since it represents the particles moving through each other.

This rather unenlightening formula will appear much more natural when we express it in terms of "center of mass coordinates" (Problem 10). In any event, we obtain the usual obvious cases: if $m_1 = m_2$ and $v_1 = -v_2$, so that we have two particles of equal mass approaching each other with opposite velocities, we get $w_1 = v_2$ and $w_2 = v_1$, so the two particles rebound with the same speeds at which they collided; if $m_1 = m_2$ and $v_2 = 0$, then $w_1 = 0$ and $w_2 = v_1$, so the first particle comes to a stop, while the second proceeds with the velocity of the first.

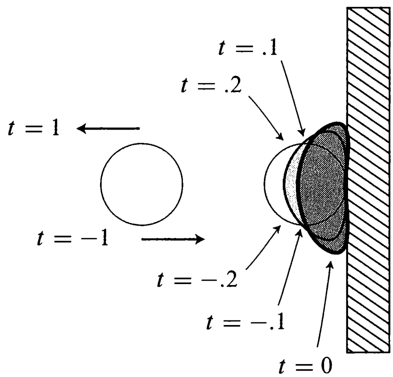
For general collisions, a *coefficient of restitution* e is sometimes defined by

$$(w_1 - w_2) = -e(v_1 - v_2) \qquad \text{or} \qquad e = \frac{w_2 - w_1}{v_1 - v_2},$$

which experimentally seems to be (somewhat) independent of the initial velocities v_1 and v_2. This is usually applied only when the two objects are moving towards each other and rebound in opposite directions after the collision, so that if we have, for example, $v_1 > 0$ and $v_2 < 0$, then also $w_1 < 0$ and $w_2 > 0$, which means that $e \geq 0$.

But having a definition of the coefficient of restitution hardly tell us anything; it simply give us a way of specifying how far experimental results differ from the theoretical ones that we obtain from our *ad hoc* definition of completely elastic collisions. We would like to understand why the modern definition of a completely elastic collision amounts to an idealization of the concept that lurks in the back of our minds when we think of an "elastic" body as one that pops back into shape after being squashed in a collision.

To simplify things, let's start by considering collisions of one body with a wall, whose mass may be regarded as so large that we don't have to worry about its motion. First we take some nice modeling clay, form it into a ball, and hurl it at the wall, where it sticks in some deformed shape. This is clearly an example of a "completely inelastic" collision. Then we throw a rubber ball at the wall. The rubber ball is also squashed when it hits the wall, but, unlike the clay, the



compressed rubber ball restores itself to its old shape, and bounces back in the reverse direction. Of course, it never bounces back with quite the same speed, but the term "completely elastic" was meant to describe an idealization of this situation, where the ball ends up pushing itself back with the same amount of force that caused the compression in the first place, so that it bounces back with the same speed. In this case, of course, we do have conservation of kinetic energy.

The general case of a "completely elastic" collision of two bodies, with velocities v_1 and v_2 along a straight line can be treated in a similar way. In this case, both bodies are deformed, and this deformation will continue until the two bodies have the same velocity u, which, by conservation of momentum, must be

$$(1) \qquad u = \frac{m_1 v_2 + m_2 v_2}{m_1 + m_2}.$$

During the compression, the first body's velocity will change from v_1 to u, so the compression will involve decreasing the velocity by the amount $u - v_1$. Consequently, when it decompresses, its velocity is then *increased* from u by the amount $u - v_1$, and of course the same reasoning applies to the second body. So the final velocities w_1 and w_2 are given by

$$(2) \qquad \begin{aligned} w_1 &= u + (u - v_1) = 2u - v_1 \\ w_2 &= 2u - v_2. \end{aligned}$$

Using the value of u from (1), we easily find that $m_1 w_1{}^2 + m_2 w_2{}^2 = m_1 v_1{}^2 + m_2 v_2{}^2$; in fact, substituting (1) into (2) gives exactly the equations (*) on page 95 that we obtained by assuming conservation of kinetic energy.

This entire discussion has been limited to "head-on" collisions, but Problems 12 and 13 have some information about the more general case.

Conservation of energy in general. Although "collisions" between atomic particles may be completely elastic, this is virtually never the case for everyday collisions between objects, where we can only hope to come fairly close to complete elasticity with objects like steel or ivory balls, and this is but one example where conservation of kinetic energy fails in general. A completely different example is illustrated by a rocket. Suppose that it is initially at rest, so that the initial kinetic energy is 0 (we assume that the rocket is in space away from any gravitational fields, so that there is no external force on the rocket, and thus no potential energy to consider). Once the rocket has expelled some fuel, so that it and the fuel are both moving, in opposite directions, the kinetic energy clearly isn't 0, since the kinetic energies are non-negative numbers, and therefore can't cancel out like momentum.

A similar phenomenon occurs when one stationary person shoots another with a rifle—the resulting motion of the bullet upsets conservation of kinetic energy (as well as the person being shot at). Since momentum is always conserved, the shooter experiences the recoil of the rifle, which is the negative of the momentum that the bullet obtains, and which will be transferred to the target, hopefully outfitted with a "bullet-proof" vest—the violent effects obtained without such protection are only indirectly a measure of the bullet's momentum, depending more on the fact that it is delivered to such a small area.

Of course, nowadays we would say that the loss of kinetic energy involved in collisions is due to its dissipation as heat, that the increase in kinetic energy of the rocket is due to the conversion of chemical energy in the fuel, and that the increase in kinetic energy of the bullet is similarly due to the conversion of chemical energy in the gunpowder. In all cases, the total energy—when we add up the heat energy and the chemical energy, and all the other types of energy which go into modern physics—is supposed to remain constant.

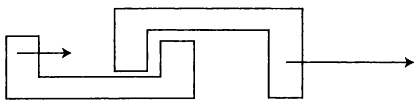
To quote from Feynman [1], the law of *conservation of energy*

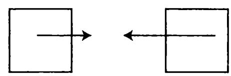
> ... states that there is a certain quantity, which we call energy, that does not change in the manifold changes which nature undergoes. ... It is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same.

Feynman goes on to discuss this by means of an analogy which is both very instructive and very entertaining, but much too long to quote here, so you should go read, or re-read, it yourself. In fact, chapters 4 through 13 of Feynman [1] may be regarded as a continuing exposition of the role that the concept of energy plays in physics.

In summary, as far as mechanics is concerned, conservation of energy—kinetic plus potential—is an important principal for conservative forces, which are generally the ones we wish to consider. On the other hand, for more complicated phenomena, which involve other forms of "energy", there are no such conservation laws; or, to put it another way, the conservation of energy involves factors which are basically outside the purview of mechanics itself.

In this regard, we may consider once again "completely inelastic" collisions. In addition to the case of a ball of clay hurled at a wall, or even at a more mobile object like a steel cube, we can also consider a completely inelastic collision between two rigid bodies that stick together because of couplings, or perhaps





glue on opposing surfaces. It is easy (Problem 15) to compute the energy lost in the collision; the result was first obtained by General Lazarus Carnot, the father of Sadi Carnot of thermodynamics fame, and has been dubbed the "Carnot energy loss" in Sommerfeld [2], mentioned on page 35.

This energy loss presumably shows up in shock waves coursing through the (not really rigid) bodies, dissipating as heat and sound waves, and possibly in some sort of chemical reaction involving the glue. In a way, this is the exact opposite of the rocket, where a large kinetic energy evolves from none at all—in that case, totally because of a chemical reaction.

A Carnot energy loss that we might judge to be fairly large produces only a small change in temperature, which might be difficult to observe experimentally. For a simple calculation, consider two iron cubes each with a mass of 1 kilogram, which smash together after moving toward each other, each with speeds of 1 meter/second. The energy loss would then be

$$1\,\mathrm{Kg}\,\frac{\mathrm{m}^2}{\mathrm{s}^2} = 1\mathrm{J},$$

by definition of the Joule. The specific heat of iron is

$$.45\ \frac{\mathrm{J}}{\mathrm{g}\,{}^\circ\mathrm{C}},$$

where °C denotes degrees centigrade, so 1 Joule raises the temperature of 1 gram of iron by .45°C, and our two iron cubes, of total mass 2 kilograms, would have their temperature raised by .45/2000 degrees centigrade. If instead of a speed of 1 m/s, which is only 3.6 km/hr, we chose speeds 30 times as fast, namely 108 km/hr, or roughly 67 miles/hr, then the temperature increase would be 900 times as great, or roughly .2 degrees centigrade; of course, this result holds just as well for two iron cubes of arbitrary masses that end up moving as one.

In mechanics problems we naturally do not expect to determine exactly how such energy losses occur in order to understand the underlying mechanical principles. But this circumstance provides a convenient rug under which all sorts of mysterious energy losses can be swept; a classic example is discussed in the following Addendum.

ADDENDUM 3A

WHIPS AND CHAINS
(Why Easy Physics is So Hard: II)

The progenitor of all those horrible "variable mass" problems introduced in Addendum 1A, which have been used to torment generations of physics students ever since, was a paper by a mathematician, Cayley [1], that begins: "There are a class of dynamical problems which, so far as I am aware, have not been considered in a general manner. The problems referred to ... are those in which the system is continually taking into connexion with itself particles of infinitesimal mass ... For instance, a problem of the sort arises when a portion of a heavy chain hangs over the edge of a table, the remainder of the chain being coiled or heaped up close to the edge of the table ... " (presumably an idealized case of a fine chain with very small links, as in Problem 1-13; notice that in Problem 1-13 the entire chain is always being pulled, so Cayley's problem is quite different).

We can apply the equations $(**)$ and $(**')$ on page 35, where our variable mass is the part of the chain hanging over the table, with the additional links added as the chain falls; since these additional links are initially at rest, the velocity at which they are added, relative to the falling chain, is $-\mathbf{v}$, so this is again a case where $\mathbf{v}+\mathbf{q}=0$, and our equation is simply

$$\mathbf{F}(t)=(m\mathbf{v})'(t).$$

Taking the uniform density of the chain to be 1 for convenience, if $x(t)$ is the length of chain hanging over the table at time t, then $m(t)=x(t)$, while $\mathbf{F}(t)$ has magnitude $gx(t)$. Thus our equation becomes simply

$$gx=(xx')'.$$

If we set $y=xx'$, so that

$$gx=y'=\frac{dy}{dt},$$

and then use the good old-fashioned trick of writing

$$\frac{dy}{dt}=\frac{dy}{dx}\cdot\frac{dx}{dt},$$

we obtain

$$gx=\frac{dy}{dx}\cdot\frac{dx}{dt}=\frac{dy}{dx}\cdot x',$$

so that

$$gx^2 = \frac{dy}{dx} \cdot xx' = \frac{dy}{dx} \cdot y,$$

or

$$gx^2\,dx = y\,dy.$$

Thus we have

$$g\frac{x^3}{3} + A = \frac{y^2}{2},$$

and if we assume the initial condition $x'(0) = 0$, we get $A = -ga^3/3$, where a is the initial length hanging over the table. This leads to

$$(x')^2 = \frac{2g}{3}\left(x - \frac{a^3}{x^2}\right). \tag{1}$$

Cayley considered only the special case $a = 0$ (which means that the whole chain is initially on the table, so that it shouldn't fall at all, but presumably he assumed that the result would be a good approximation to the case where a is small). Then (1) becomes $x' = \sqrt{2g/3}\,\sqrt{x}$, and integrating $x'/\sqrt{x} = \sqrt{2g/3}$ yields

$$2\sqrt{x(t)} = \sqrt{2g/3} \cdot t, \quad \text{or} \quad x(t) = (1/6)gt^2.$$

Consequently, the chain is always falling with $1/3$ the acceleration of a freely falling chain; of course, that applies only until the chain has cleared the table, after which its acceleration must simply be g.

At first sight this seems rather weird and unlikely, and we might suspect that it is an artifact of the strange choice $a = 0$. Cayley may have taken this case because he didn't want to deal with the general equation (1), which would require an elliptic integral. But, without solving explicitly, we can still compute from (1) that

$$2x'x'' = \frac{2g}{3}\left(x' + \frac{2a^3x'}{x^3}\right),$$

so that

$$x'' = \frac{g}{3}\left(1 + \frac{2a^3}{x^3}\right).$$

At the beginning ($x = a$) this has the value g, and it then decreases, approaching $g/3$ for long chains, with the same sort of discontinuity as before. (The equation for x'' follows from the previous equation when $x' \neq 0$, (i.e., $x \neq a$), and then for $x = a$ by continuity, although technically we must appeal to an elementary calculus theorem.[1])

[1] If f is continuous at a and $\lim_{x\to a} f'(x)$ exists, then $f'(a)$ exists and $= \lim_{x\to a} f'(x)$. See, e.g., Spivak [1], Theorem 11-7.

To make sense of this perplexing answer, which of course is only approximate for the case of an actual chain, it helps to note that each time another link is added to the falling chain, that link is suddenly yanked from velocity 0 to the velocity of the falling chain, and the resultant increase in momentum must be balanced by a decrease in momentum of the falling chain. As the falling chain gets longer and more massive, one might expect the effect to be less noticeable, but the longer falling chain also has a much greater velocity, so the momentum added to the next link also increases greatly.

This problem has appeared in many standard mechanics books—usually with Cayley's solution, though Sommerfeld [2] hints at the more general solution—not so much for its own sake, but in order to examine the question of conservation of energy.

When a piece of chain of length x is hanging over the table, the potential energy has decreased by $\int_a^x gu\,du = \frac{1}{2}g(x^2 - a^2)$, while the kinetic energy has increased by $\frac{1}{2}x(x')^2$, so that the total change of energy is

$$\Delta E = \tfrac{1}{2}x(x')^2 - \tfrac{1}{2}g(x^2 - a^2).$$

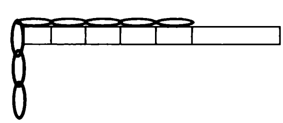
By (1) this is

$$-\tfrac{1}{6}gx^2 + \tfrac{1}{2}ga^2 - \tfrac{1}{3}g\frac{a^3}{x}$$

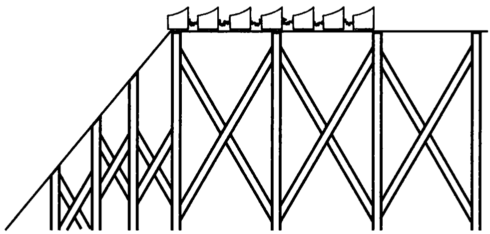
so for all $x \geq a$ it is negative, and increases rapidly as x gets large. It is hardly surprising that conservation of energy does not hold for this solution, since, as Problem 23 shows, it *does* hold for the solution to Problem 1-13.

This loss of energy is explained in terms of the Carnot energy losses each time the dangling chain pulls another link off the table, the point being that this is a completely inelastic collision, since the resulting velocities of the two bodies are the same; the energy loss presumably ends up heating the chain.

It actually turns out to be rather difficult to conduct experiments to check our answer, because one can't get a real chain "coiled or heaped up close to the edge of the table" in such a way that each link is right at the table edge, ready to be added to the falling part; in practice, there is an unpredictable jumble as individual links are released. One way to simulate the conditions of the problem might

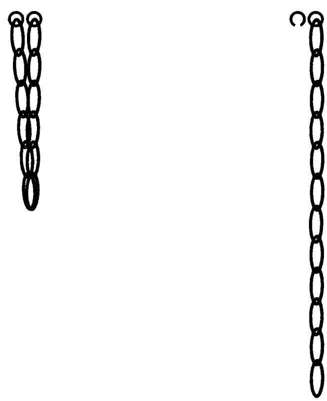


be to lay the additional links along a table made of slats, removing each leftmost slat once the link lying on it has been pulled off. An inexact, but instructive approximation could be provided by a carnival ride, the "whip", where individual cars are lined up close to the edge of a slide, but attached with bunched-up 50′ chains. After the first car is allowed to start sliding, the experience of its



riders is quite different from that of riders in the last car!)

There is another classical problem for which experiments can more easily be carried out. Consider a folded chain, initially hanging by supports at both



ends, and then released at one end. After the chain has finished falling, its total energy (all in the form of potential energy) is much smaller.

The classical description of this situation is that each of the links on the free end of the chain just falls with acceleration g until being jerked to a stop, with the loss of energy being accounted for by the corresponding Carnot energy losses. But a simple experiment shows that the acceleration must be considerably greater than g. It involves only a moderately heavy chain, a single link from such a chain, and two pieces of window glass. The specific chain used in one experiment was 5 feet long ($\approx$ 152 cm), with 50 links, and weighed about 14.3 oz ($\approx$ 405 gm). The thickness of the glass was 3/16 inches ($\approx$ .5 cm).

Opposite ends of one piece of glass were placed on rests of the same height (two copies of a book), and the single link was repeatedly dropped onto it from a height greater than 5 feet, with no apparent ill effect (sometimes the link was initially held horizontally, sometimes vertically). The piece of glass was replaced

with the second, fresh, piece, and the 5 foot long chain was secured so that it hung with only the last link touching the glass. The free end was then raised to the same height as the secured end (a short distance away from it horizontally, so that the chain wouldn't become entangled in itself as it fell) and released. The result was a dramatic shattering of the glass plate.

We can analyze the fall of the chain in this problem in the same way as the Cayley problem, taking as our body with variable mass the falling part of the chain, which is "losing" links to the fixed part. Since these links become stationary as they join the fixed part of the chain, we again have $\mathbf{q} = -\mathbf{v}$, so we still have the case where $\mathbf{v} + \mathbf{q} = 0$, leading to the same equation

$$\mathbf{F}(t) = (m\mathbf{v})'(t).$$

As before, we assume that the chain has uniform density 1. It will also be convenient to assume that the fully extended chain is hanging so that it just touches the ground, and then let x be the height of the free end of the chain, initially having the value x_0 (thus, $x_0 = L$ for a folded chain of length L with both ends initially at the same height). The falling part of the chain has length $x/2$, so our equation becomes

$$g\frac{x}{2} = -\frac{1}{2}(xx')',$$

or simply $gx = -(xx')'$. Setting $y = xx'$ we now have

$$\begin{aligned} gx &= -\frac{dy}{dt} \\ &= -\frac{dy}{dx} \cdot x', \end{aligned}$$

so that

$$gx^2 = -\frac{dy}{dx} \cdot xx' = -\frac{dy}{dx} \cdot y,$$

and hence

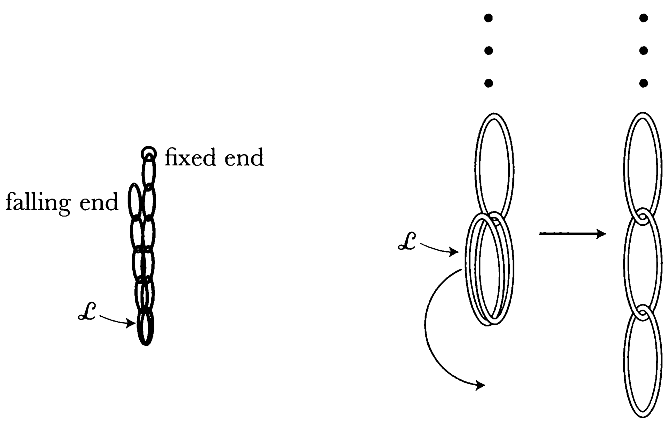
$$g\frac{x^3}{3} + A = -\frac{y^2}{2}.$$

At $t = 0$ we have $x = x_0$ and $x' = 0$, so we get $A = -g{x_0}^3/3$, leading to

$$\text{(A)} \qquad \begin{cases} (x')^2 = \dfrac{2}{3}g\left(\dfrac{{x_0}^3}{x^2} - x\right) \\[2ex] x'' = -\dfrac{g}{3}\left(\dfrac{2{x_0}^3}{x^3} + 1\right), \end{cases}$$

the second equation following by differentiation of the first, as before. Thus the downward acceleration starts at g and then increases, so the released end of the chain falls faster than a freely falling chain.

This increase in acceleration can be explained by considering a link $\mathcal{L}$ of the chain that has just reached the bottom, as in (a). This link has acquired a large velocity, but is now going to be stopped dead in its descent by the part of the



(a) (b) (c)

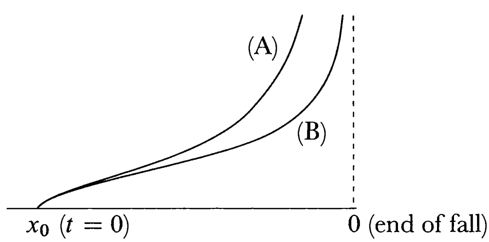
chain on the fixed end, and all that momentum will be used to yank the link around by 180°, as shown magnified in (b) and (c) of the figure. This yanking is going to pull the falling part of the chain even faster. This is basically just the opposite of what happens for the Cayley problem, where the falling chain yanks the next link off the table, resulting in the falling chain having its acceleration reduced; in the current situation the link that was falling, but now becomes part of the fixed chain, yanks the falling chain, resulting in the falling chain having its acceleration increased.

The fact that the falling chain has acceleration greater than g seems to have been first observed by Calkin and March [1]. To explain the results of their experiments (rather more sophisticated than the shattered glass experiment), they completely jettisoned the question of Carnot energy losses, and simply assumed that conservation of energy holds, obtaining the equations (Problem 24)

$$\text{(B)} \qquad \begin{cases} (x')^2 = g\left(\dfrac{{x_0}^2}{x} - x\right) \\[2ex] x'' = -\dfrac{g}{2}\left(\dfrac{{x_0}^2}{x^2} + 1\right), \end{cases}$$

which seem to agree quite well with their experimental results. (In the case of this solution, the increase in velocity is easily explained by the fact that the same amount of energy has to be concentrated in shorter and shorter pieces of chain, so that the velocities must increase.)

In the figure below, comparing the downward speeds for equations (A) and (B), the direction of the x-axis is reversed, so that $x = x_0$, at time $t = 0$, appears on the left, while $x = 0$, at the end of the fall, appears on the right. I do not



know how well the Calkin-March data would match up with equations (A) [for an actual chain, of only finitely many links, either set of equations becomes less reliable near the end of the fall, which is where the solutions diverge the most], or how to choose between them, or whether the solution for a real chain is some sort of compromise between the two. Or, for that matter, how one should treat the same problem when the chain is replaced by a rope.

Note that, with either solution, at the end of the fall the speed and acceleration actually become infinite, or at any rate very large for an actual chain of only finitely many links, where the friction between links also takes its toll. This possibly counter-intuitive behavior is also demonstrated by the crack of a whip;

here the force applied to the whip takes the place of gravity, and the crack of the whip is a shock wave caused by the very large velocity with which the end of the whip is traveling.

A more recent paper examining these questions, Wong and Yasui [1], approves of the Calkin-March solution, dismisses Cayley's solution of his problem, and by extension the one given here, in favor of a conservation of energy solution (Problem 25), and goes on to discuss the folded chain problem in great detail. This paper contains an extensive bibliography of previous solutions to both the folded chain problem and the Cayley problem, which may be very instructive to peruse. But it appears to me that all the conclusions of the paper itself are wrong.

Undoubtedly others will find that all the conclusions in this Addendum are wrong.

I must admit to being totally confused. I thought mechanics was a cookbook subject where one uses a few basic principles to translate physics into mathematics, and then revs up the calculus machine and grinds out the answer. I guess your book is intended to cure those of us who have this misapprehension.

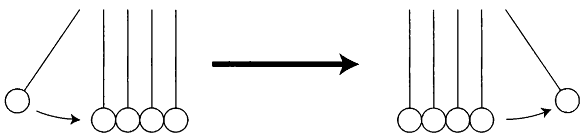
—An eminent mathematician

ADDENDUM 3B

FOLLOW THE BOUNCING BALL
(Why Easy Physics is So Hard: III)

I was fortunate enough to be a graduate student at Princeton when the mathematics department was still ensconced in the original Fine Hall, whose comforting common room, with its antique appointments and fireplace bearing Einstein's famous words "Raffiniert ist der Herr Gott, aber boshaft ist Er nicht" (God is subtle, but not malicious), was often equally a meeting place for physicists from the adjoining building.

At one point this led to a lively discussion of a demonstration often included in mechanics courses, involving a series of balls suspended from cords so that they just touch. If one ball is raised on the left, and allowed to strike the remaining

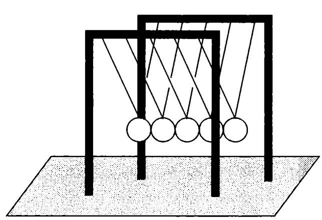


balls, the right-most ball is observed to swing out to the same height, while if two balls a raised together on the left, the right-most two balls swing out together

on the right. And if two balls are raised on the left while one is raised on the right, the final result seems to "interchange" the balls, with one ball flying off

on the left and two balls flying off together on the right.

The apparatus for such striking and amusing experiments has apparently been dubbed "Newton's cradle", a miniature version of which is often sold as a

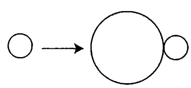


device to relieve an executive's boredom, as well as an "educational toy", since it supposedly illustrates the laws of conservation of momentum and energy.

Of course, the results are merely *consistent* with the laws of conservation of momentum and energy. Those two laws alone cannot possibly *predict* the results once there are more than two balls, since they give only two equations for three or more unknowns. There are, of course, a few other mathematical conditions, expressing the fact that one ball can't pass through another, but not enough to make the solution unique. In fact, to get another solution one need only

consider, for example, the case where one ball collides with two balls that simply behave as a single ball of twice the mass.

Similarly one can ask what would happen when a small ball strikes a large ball



that touches another small ball on the other side. There were those who claimed that the results would be exactly the same as if the large ball were replaced by a series of small balls: the incoming small ball would come to a stop, while the other small ball would fly off with the same velocity. It was pointed out that this would be rather strange, since one wouldn't expect that throwing a ball at the earth would cause it to stop dead while a ball at the antipodes would suddenly shoot up in the opposite direction! But that objection was rather disparaged, supposedly explained away by the fact that the earth and the balls were far from completely elastic.

At this point, it was helpfully pointed out that the physics department kept a handy room full of experimental equipment, including ivory balls suspended from strings. We enthusiastically decided to do an experiment, much to the disdain of several physicists, who considered this a complete waste of time, as they had already told us what the result would be. To our delight, we even found that the equipment included one large ivory ball. Although one passing physics professor warned us that we needed to know the speed of sound in ivory, we stubbornly proceeded merely to do the experiment and observe the results. Of course, as expected, the first ball bounces back quite a bit, rather than coming to a stop, allowing us to return to the common room triumphantly echoing Galileo's rebellious words "But it does move!"

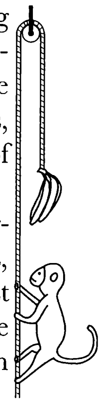
Ever since that time, I have regarded explanations given in mechanics texts with a healthy skepticism—which has helped shape much of the material in Chapters 5 and 6.

Mathematicians like to use the "epsilon-method" for solving this problem, considering what happens when the balls are separated by a small distance ε, and letting $\varepsilon \to 0$, but there's no physical justification for assuming that this will give the right result for balls starting in contact. In fact, since material spheres aren't perfectly spherical, and positions aren't precise, it's not even clear what "in contact" really means—in practice, the balls will actually be pressing against each other.

Some models of Newton's cradle do leave a slight gap between the balls, presumably because this gives better results. In fact, careful experiments when the balls are in contact, Chapman [1], showed discrepancies from the theoretical predictions larger than can be attributed to the slight deviation from complete elasticity, and led to several other investigations, e.g., Herrmann and Schmälzle [1], Herrmann and Seitz [1], and Auerbach [1].

This is the sort of problem that cannot be solved by a simple application of Newton's laws, treating each ball as a particle; we would have to consider all the individual molecules of the balls and their interacting forces, which would be a problem of impossible complexity, even if we actually knew these forces. Instead, we need a simplified model that gives good agreement with experiment. Not surprisingly, therefore, the question has received considerable attention from practitioners of "applied mechanics" or "mechanical engineering", who really need to solve such problems. See, for example, Ceanga and Hurmuzlu [1].

PROBLEMS



1. As an undergraduate math major, the hopelessness of understanding physics was borne down upon me every time I heard about some first-year physics problem. One such typical problem, usually presented on the second or third day of class, concerns a monkey, let's call him Tantalus, who is climbing a rope passing over a pulley with a bunch of bananas of exactly his weight attached to the other end.

Problems of this sort always drove me bananas because I never understood how one was supposed, on the basis of Newton's laws for particles, to divine the fact that the end of the rope attached to the bananas must be exerting exactly the same upward force that the other end of the rope exerts upwards on Tantalus (problems of this sort are considered briefly in Chapter 6).

But what really made a monkey out of me was the next step, concluding that Tantalus and the bananas rise at exactly the same rate. My mistake was that I kept trying to think about the mechanism of Tantalus climbing—just what happens when he lifts one arm to reach higher on the rope and then tugs on the rope (does it matter if he still holds on to the rope with the other hand)? Sometimes, thinking about physics problems just seems to make them harder.

Actually, the standard answer, that Tantalus and the bananas rise at exactly the same rate, is false—in fact, it is *meaningless*. What is the correct answer?

The *center of mass* of Tantalus rises at the same rate as (the center of mass) of the bananas. Of course, the center of mass of the bananas doesn't change (unless Tantalus gets lucky and a banana falls off), but Tantalus' center of mass continually changes as he climbs.

2. Let $A_1, \ldots, A_K$ be K collections of particles, with C_i the center of mass of A_i. Show that the center of mass C of the collection $A = A_1 \cup \cdots \cup A_K$ is the same as the center of mass of the collection $\{C_1, \ldots, C_K\}$.

3. (a) If two particles c_1 and c_2 satisfy the second law, $\mathbf{F} = m_i \mathbf{v}_i{}'$, as well as the Momentum Law $\mathbf{F} = \sum_i m_i \mathbf{v}_i{}'$, then the internal forces between them satisfy the third law.
(b) If they also satisfy the Angular Momentum Law $\boldsymbol{\tau} = \mathbf{L}'$, then the internal forces satisfy the strong form of the third law.
(c) Can similar conclusions be drawn for a system of more than two particles?

4. (a) For a collection of particles c_i the total angular momentum $\mathbf{L}_P$ is independent of P if and only if $\sum_i m_i \mathbf{v}_i = 0$.
(b) For a collection of forces $\mathbf{F}_i$ at c_i, the total torsion $\boldsymbol{\tau}_P$ is independent of P if and only if $\sum_i \mathbf{F}_i = 0$.

290

5. A "spherical pendulum" is just a pendulum that is not necessarily swinging in a plane, which can easily happen if the pendulum bob is given a push in some direction as it is released. The forces on the pendulum bob are the force of gravity downwards and the force exerted by the string, as on page 47.

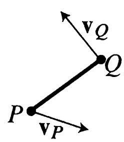
(a) The vertical component of $\boldsymbol{\tau}$ is always 0.
(b) If the path of the pendulum bob is $c(t) = (x(t), y(t), z(t))$, then $x'y - y'x$ is constant.
(c) If the pendulum is ever perpendicular, passing over $(0,0)$, then it is actually swinging in some vertical plane.

6. (a) Consider a system of particles $c_1, \dots, c_K$ with total mass M and center of mass C. If $\gamma_i = c_i - C$, then

$$T = \tfrac{1}{2}M|C'|^2 + \tfrac{1}{2}\sum_i m_i|\gamma_i{}'|^2.$$

(b) Consider inertial systems that differ only by the choice of origin. Show that the one having origin C is the one with the smallest kinetic energy.

7. Let $\mathbf{v}_P$ and $\mathbf{v}_Q$ be the velocities at some time of the endpoints P and Q of



a uniform rigid rod of mass m. Show that the kinetic energy of the rod at this time is

$$T = \tfrac{1}{6}\big[|\mathbf{v}_P|^2 + |\mathbf{v}_Q|^2 + \langle \mathbf{v}_P, \mathbf{v}_Q\rangle\big].$$

8. Assuming that the collision of two objects doesn't increase the total kinetic energy, show that the coefficient of restitution e satisfies $0 \le e \le 1$.

9. In Chapter 1, we mentioned Huygens' ingenious argument (page 26), based on the idea of examining a collision in two different coordinate systems, moving with uniform velocity with respect to each other. Huygens actually extended his argument in a strange and complicated way (see Dugas [1; pp. 177-180] and Mach [1; pg. 403ff.]) that essentially assumed conservation of kinetic energy in collisions.

Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be the initial velocities of two bodies, of masses m_1 and m_2, and $\mathbf{w}_1$ and $\mathbf{w}_2$ their velocities after a collision. Assuming conservation of kinetic energy we then have

$$m_1\langle \mathbf{v}_1, \mathbf{v}_1\rangle + m_2\langle \mathbf{v}_2, \mathbf{v}_2\rangle = m_1\langle \mathbf{w}_1, \mathbf{w}_1\rangle + m_2\langle \mathbf{w}_2, \mathbf{w}_2\rangle. \tag{1}$$

Obtain the corresponding equation assuming that conservation of energy also holds in an intertial system moving with velocity $\mathbf{u}$ with respect to the original one, and conclude that

$$m_1\langle \mathbf{v}_1, \mathbf{u}\rangle + m_2\langle \mathbf{v}_2, \mathbf{u}\rangle = m_1\langle \mathbf{w}_1, \mathbf{u}\rangle + m_2\langle \mathbf{w}_2, \mathbf{u}\rangle.$$

If this is true for all $\mathbf{v}$, then we have

$$m_1\mathbf{v}_1 + m_2\mathbf{v}_2 = m_1\mathbf{w}_1 + m_2\mathbf{w}_2;$$

so conservation of kinetic energy in collisions implies conservation of momentum (provided we assume that the conservation of kinetic energy holds all inertial systems).

10. (a) For two particles colliding along a straight line, if the total momentum of our two particle system is zero, we have

$$m_1v_1 + m_2v_2 = 0 \implies v_2 = -(m_1/m_2)v_1,$$

and thus also

$$m_1w_1 + m_2w_2 = 0 \implies w_2 = -(m_1/m_2)w_1.$$

Show that conservation of kinetic energy leads immediately to the solutions $w_1 = v_1, w_2 = v_2$ (ignored for physical reasons) and $w_1 = -v_1, w_2 = -v_2$, so that after the collision the velocities are simply reversed.
(b) For a system with no external forces, the center of mass has constant velocity, so we can choose an inertial system, the center of mass coordinates, in which the center of mass is the origin. Show that the total momentum in this inertial system is zero. Letting v_i^* and w_i^* denote the initial and final velocities in the center of mass coordinates, compute v_i^* and w_i^* in terms of the v_i and w_i and deduce (∗) on page 95 from the interchange of speeds in the center of mass coordinates.

11. For a particle of mass m_1 and velocity v_1, in a totally elastic head-on collision with a stationary particle of m_2, use (∗) to show that

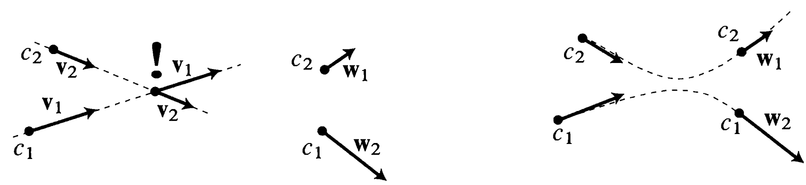
$$\frac{w_2}{v_1} = \frac{2m_1}{m_1 + m_2}.$$

Conclude that if m_1 and v_1 are unknown, but m_2 and w_2 are known for two different stationary particles of different masses, then m_1 and v_1 can be determined. In 1932 Chadwick used this method for certain unknown uncharged particles created in a nuclear reaction colliding with various nuclei of known masses, to determine that the mass of these particles (now known as neutrons) was practically equal to that of the proton.

259

12. Suppose that c_1 and c_2 are bodies of the same mass, with c_2 initially at rest, so that they have initial velocities $\mathbf{v}_1$ and $\mathbf{v}_2 = 0$. Show that after a perfectly elastic collision their velocities $\mathbf{w}_1$ and $\mathbf{w}_2$ are perpendicular.

13. Now consider the general situation where two particles c_1 and c_2, moving with velocities $\mathbf{v}_1$ and $\mathbf{v}_2$ lying in a plane, collide and end up with velocities $\mathbf{w}_1$ and $\mathbf{w}_2$. As illustrated in part (b) of the figure, this is often applied to situations



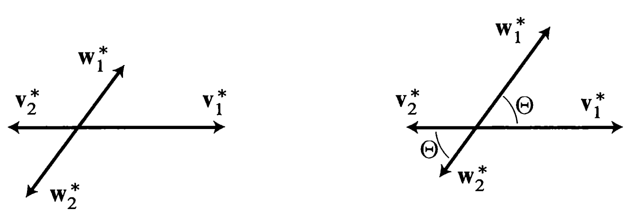
(a) (b)

where two particles don't actually collide, but are deflected from a straight path for other reasons. For example, two positively charged particles initially far from each other might follow paths like this, a situation we will encounter in Addendum 4C.

(a) The velocity of the center of mass C in our original coordinate system is $\mathbf{v}_C = (m_1\mathbf{v}_1 + m_2\mathbf{v}_2)/(m_1 + m_2)$, and the velocities

$$\mathbf{v}_1^* = \mathbf{v}_1 - \mathbf{v}_C$$
$$\mathbf{v}_2^* = \mathbf{v}_2 - \mathbf{v}_C$$

are negative multiples of each other, and similarly for $\mathbf{w}_1^*$ and $\mathbf{w}_2^*$.
(b) If the collision is perfectly elastic we have $v_1^* = w_1^*$, and $v_2^* = w_2^*$. Thus the speeds are the same before and after the collision, and the velocity vectors simply rotate by an angle Θ, known as the scattering angle.



(a) general collision (b) perfectly elastic collision

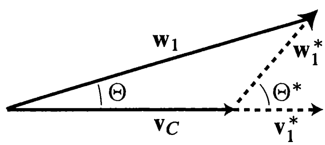
14. Many experiments involve a moving particle colliding with one that is initially at rest in the laboratory. Addendum 4C describes one such experiment, where the mass of the moving particle is very small compared to that of the stationary particle, so that the center of mass coordinates are practically the same as the "laboratory coordinates". In general, however, the observed "laboratory scattering angle" differs significantly from the scattering angle in the center of mass coordinates.

(a) If the second particle is stationary, $\mathbf{v}_2 = 0$, then the velocity $\mathbf{v}_C$ of the center of mass is parallel to $\mathbf{v}_1$ with magnitude $v_C = m_1 v_1/(m_1 + m_2)$, and

$$\left.\begin{aligned} \mathbf{v}_1^* &= \frac{m_2}{m_1 + m_2}\mathbf{v}_1 \\ \mathbf{v}_2^* &= \frac{-m_1}{m_1 + m_2}\mathbf{v}_1 \end{aligned}\right. \qquad \Longrightarrow \qquad v_C/v_1^* = m_1/m_2.$$

(b) If Θ^* is the scattering angle in the center of mass system, and Θ the scattering angle in the laboratory system, then we have

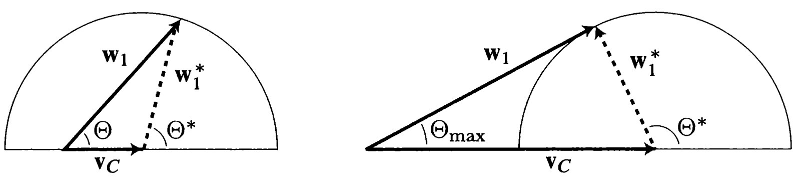
$$\tan\Theta = \frac{w_1^* \sin\Theta^*}{v_C + w_1^* \cos\Theta^*}.$$



If, moreover, the collision is completely elastic, so that $w_1^* = v_1^*$, then

$$\tan\Theta = \frac{\sin\Theta}{m_1/m_2 + \cos\Theta^*}.$$

(c) Show that if $m_1 < m_2$, then this expression has a maximum at $\cos\Theta^* = -m_2/m_1$, and conclude that Θ has a maximum possible value Θ_{max} with $\sin\Theta_{\text{max}} = m_2/m_1$. This is often derived geometrically from the figure below, where (a) illustrates the case that $v_C/v_1^* = m_1/m_2 < 1$, while in (b) we



(a) (b)

have $v_C/v_1^* = m_1/m_2 > 1$, and the minimum value of Θ occurs for $\mathbf{w}_1$ perpendicular to $\mathbf{w}_1^*$.

If $m_1 \gg m_2$, then Θ_{max} is close to 0; in fact, we then have $\Theta_{\text{max}} \approx m_2/m_1$. All of which is a fancy way of saying that when a body of large mass hits a body of small mass, the body of large mass is hardly deflected.

15. (a) Using the formulas for $\mathbf{v}_i^*$ in Problem 14, show that the momenta of the particles in the center of mass coordinates can be written as

$$\begin{aligned} m_1\mathbf{v}_1 &= \mu\mathbf{v} \\ m_2\mathbf{v}_2 &= -\mu\mathbf{v}, \end{aligned}$$

where $\mathbf{v} = (\mathbf{v}_1 - \mathbf{v}_2)$ and μ is the *reduced mass*,

$$\mu = \frac{m_1 m_2}{m_1 + m_2},$$

a quantity that frequently arises in two-particle problems (see page 136).
(b) For a completely inelastic collision between two bodies of masses m_1 and m_2 approaching each other along a line, with velocities v_1 and v_2, the resultant common velocity is given by equation (1) on page 96. Compute that the loss of kinetic energy is

$$\tfrac{1}{2}\mu v^2, \qquad v = v_1 - v_2,$$

(the kinetic energy of a body of mass μ moving with the relative velocity v of the two particles).

291 **16.** For Problem 1-22, use conservation of energy

$$\tfrac{1}{2}m(x'^2 + y'^2) + mgy = E$$

and $y = l - \sqrt{l^2 - x^2} \approx x^2/2l$ to deduce directly that for small oscillations we have $x'' + (g/l)x = 0$.

17. (a) The sum of conservative forces is conservative. (Trivial, but worth noting, in order to get simple examples of non-radially symmetric conservative forces!)
(b) In particular, find a non-radially symmetric conservative force with a singularity at the origin, and nowhere else.

18. Consider a force that is central, always pointing towards the origin, but not radially symmetric. Show that it cannot be conservative.

Hint: The work done moving along any sphere around the origin will be 0.

19. Consider a function $V\colon \mathbb{R}^3 \to \mathbb{R}$ that is homogeneous of order k, meaning that

$$V(ax) = a^k V(x) \qquad x \in \mathbb{R}^3,\ a \in \mathbb{R};$$

note that each $\partial V/\partial x^i$ is homogeneous of order $k-1$. Let c be a particle moving under the force with potential function V, so that

$$c_i{}'(t) = -\frac{\partial V}{\partial x^i}(c(t)),$$

and consider the new path

$$\bar{c}_i(t) = \alpha \cdot c(\beta t)$$

[we change the time by a factor of β and then the position by a factor of α].

(a) Show that

$$\bar{c}_i{}'(t) = -\alpha\beta^2 \frac{\partial V}{\partial x^i}(c(\beta t))$$

$$= -\frac{\alpha\beta^2}{\alpha^{k-1}} \frac{\partial V}{\partial x^i}(\bar{c}(t)).$$

Consequently, $\bar{c}$ satisfies the same equation as c if

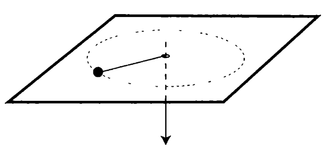
$$\beta = \alpha^{\frac{k}{2}-1}.$$

(b) In the case of a uniform gravitational field (like that near the earth's surface) we have $k = 1$. So for any path c, the path $\bar{c}(t) = \alpha \cdot (t/\sqrt{a})$ is also a solution. This basically reproves the result of Problem 1-17 (a).
(c) Generalize part (b) of that problem similarly.
(d) Also use this general result to reprove Problem 2-1 (a).

20. Find the "escape velocity" of a rocket on earth—the smallest initial velocity it must have so that it never falls back to earth (ignoring air resistance, the gravitational force of the sun, etc.), in terms of g and the earth's radius R_e. Notice that this does not depend on the rocket being fired directly upwards. (But rocket launches nearer the equator are better because the centrifugal force due to the earth's rotation is greater there, so that the rocket actually starts with a greater velocity.)

In contrast to Problem 1, the next problem shows that although it sometimes pays not to think too hard about a physics problem, at other times the difficulty is thinking about what the problem is coyly trying to say.

21. (a) A small object travels in a circle on a frictionless table, held in place by a string that passes through a hole in the table. The string is slowly pulled



through the hole so that the radius of the circle changes from r_0 to r_1. Show that the work done pulling the string equals the increase in kinetic energy of the object.

"Slowly" means: pretend that the object moves in a series of circles instead of a spiral (and that the force exerted upon it is radial).

(b) For a string being pulled through the hole with a constant force of magnitude F (not "slowly"), find the velocity at time t in terms of the initial velocity v_0 when the object is traveling in a circle of radius r_0.

22. (a) In Problem 1-15, the force must be $F(t) = m'(t)v$. Show that the total work done by this force from time 0 to t will be

$$\int_0^t vF(\tau)\,d\tau = \int_0^t v^2 m'(\tau)\,d\tau = v^2 m(t),$$

which is twice the kinetic energy of an object of mass $m(t)$ moving with speed v. Why isn't this a contradiction?

To put it another way, what does this say about this force?

(b) Suppose a rope is lying on the floor, and it is pulled up from one end with constant speed. Compare the work done just after the whole rope has been pulled up with the total potential plus kinetic energy that the rope has.

23. (a) Check that conservation of energy holds for the calculated solution to Problem 1-13 (a chain sliding off a table).
(b) Similarly, use conservation of energy to derive this solution.

24. For the folded chain problem of Addendum 3A, derive equations (B) on page 106 from conservation of energy.

25. Suppose that we solved the original Cayley problem in the same way, by assuming conservation of energy. Using the formula for ΔE on page 102, show that we would get

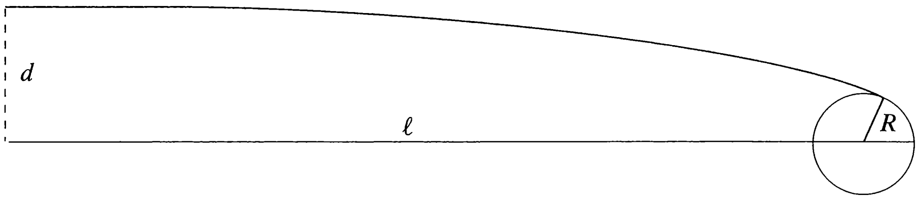
$$(x')^2 = g\left(x - \frac{a^2}{x}\right), \qquad x'' = \frac{g}{2}\left(1 + \frac{a^2}{x^2}\right),$$

which still has the disconcerting continuity when the chain leaves the table.

26. A ball of mass M having a coefficient of restitution nearly equal to 1 in collisions with the floor, is dropped on the floor with a ball of mass $m \ll M$ sitting on top of it. What happens? (To avoid the sorts of difficulties encountered in Addendum 3B, assume that the two balls are actually separated by a small

distance ε.) The Wham-O® SuperBall® (for which Wikipedia has an interesting entry) can be used for an instructive experiment, with something like a marble for the top ball.

174 **27.** A space ship of mass m, launched far away form a planet of mass M and radius R, has been traveling with constant speed v_0 along a path parallel to the line ℓ through the center of the planet, at distance d from this line. It is now



beginning to approach the planet, and we want to know how large d can be so that it will hit the planet. If r is the distance from the space ship to the planet, then the total energy of the space ship is

$$\frac{1}{2}mv^2 - \frac{mMG}{r},$$

where r is the distance from the center of the planet to the space ship (and G is the constant in the law of gravity [page 37]).

(a) Far from the planet, the angular momentum of the rocket ship with respect to the center of the planet has magnitude mdv_0, and the total energy has value $\frac{1}{2}m{v_0}^2$.

(b) Suppose that the ship just grazes the planet, with speed w. Show that the angular momentum at that time has magnitude mRw, and the total energy has the value

$$\frac{1}{2}mw^2 - \frac{mMG}{R}.$$

Explain why we can conclude that

$$d^2 = R^2\left(1 + \frac{2MG}{Rv^2}\right).$$

CHAPTER 4

THE ONE-BODY AND TWO-BODY PROBLEMS

After the bruising engagement with real-world problems of the previous chapter, we beat a hasty temporary retreat to the safer realm of more purely mathematical questions, restricting our consideration of "conservation of energy" to its most basic form for mechanics—conservation of kinetic plus potential energy. This chapter is a sort of companion to Chapter 2, giving a connected modern treatment of the material covered there, together with further developments, most of which Newton also treated in the Principia.

The one-body problem. Consider a particle

$$c(t) = r(t) \cdot (\cos\theta(t), \sin\theta(t))$$

moving under a radially symmetric central force, so that

$$m\mathbf{v}' = -(f \circ r) \cdot \frac{c}{r};$$

as in Chapter 3, we add the $-$ sign so that a positive f corresponds to an attractive force. A planet moving around the sun is the prototype of this problem, which is usually called the "one-body problem" because we are ignoring the force that the planet exerts on the sun, and thus really only considering a single object under the influence of some force that we do not specifically attribute to another particle.

As we noted in the previous chapter, our force is conservative, with a potential energy function V. Writing $\mathbf{v}$ in terms of r and θ, the conservation of energy formula

$$\tfrac{1}{2}mv^2 + V = E$$

becomes

$$\tfrac{1}{2}m(r'^2 + r^2\theta'^2) + V \circ c = E \tag{1}$$

We also have (page 56)

$$r^2\theta' = h \quad \text{for a constant } h. \tag{2}$$

Squaring (2), and substituting into (1), we obtain

$$\text{(A)} \qquad r'^2 = \frac{2}{m}\big(E - (V \circ c)\big) - \frac{h^2}{r^2} \qquad \text{or} \qquad r'^2 = -\frac{2}{m}(F \circ r) - \frac{h^2}{r^2},$$

where the second equation, previously obtained on page 75, has the potential function $V \circ c$ written in terms of F with $F' = f$, which is only defined up to a constant; in the first equation the constant is written as E, which amounts to a choice of the constant in V.

Note that in the second equation of (A) the mass m doesn't really play a role, because we are going to be considering forces where f is proportional to m. Similarly, m can be ignored in the first equation if we replace E by $\bar{E} = E/m$, the total "energy per unit mass"; in the same vein, the angular momentum is mh, so h is just the angular velocity, or "angular momentum per unit mass". Often it is convenient simply to assume that $m = 1$.

Taking the derivative of (A), and dividing by r' we obtain

$$\text{(B)} \qquad r'' = -\frac{f \circ r}{m} + \frac{h^2}{r^3}.$$

As on page 101, we technically need to be more careful, especially for a circular orbit, with r constant. Note, by the way, that an attractive force $f > 0$ always has circular orbits for any radius ρ,

$$c(t) = \rho(\cos at, \sin at),$$

since we just need to choose a so that $f(\rho)$ is m times the magnitude of the acceleration, or $f(\rho) = m\rho a^2$; in terms of $h = \rho^2 a$, we need

$$(\mathrm{B}_\rho) \qquad mh^2 = \rho^3 f(\rho),$$

which is what (B) then reduces to.

The solutions r of (B), together with (2), giving θ, theoretically provides all orbits, complete with parameterization, but to determine only the shape of the orbit we can combine (2) and (A) to find the derivative of $r \circ \theta^{-1}$. As in Problem 2-6, calculations are simplified by using Leibnizian notation. We have

$$\frac{dr}{d\theta} = \frac{dr}{dt}\Big/\frac{d\theta}{dt},$$

and obtain [with the obvious interpretation of $V(r)$]

$$\frac{dr}{d\theta} = \frac{r^2}{h}\sqrt{\frac{2}{m}\big(E - V(r)\big) - \frac{h^2}{r^2}} \qquad \text{or} \qquad \frac{dr}{d\theta} = \frac{r^2}{h}\sqrt{-\frac{2}{m}F \circ r - \frac{h^2}{r^2}},$$

and then

$$\theta = \int \frac{h\,dr}{r^2\sqrt{\frac{2}{m}\big(E - V(r)\big) - \frac{h^2}{r^2}}} \quad \text{or} \quad \theta = \int \frac{h\,dr}{r^2\sqrt{-\frac{2}{m}F \circ r - \frac{h^2}{r^2}}},$$

theoretically allowing us to express θ in terms of r, and thus r in terms of θ. Although Newton's first derivation of the orbits for an inverse square law was the geometric one given in Chapter 2, later in the Principia Newton also gave essentially these same equations, though stated entirely in geometric terms—see Addendum A.

It is usually more convenient to write our equations in terms of $u = 1/r$, with

$$\text{(C)}\quad \frac{du}{d\theta} = -\frac{1}{h}\sqrt{\frac{2}{m}\big(E - V(1/u)\big) - h^2u^2} \quad \text{or} \quad \frac{du}{d\theta} = -\frac{1}{h}\sqrt{-\frac{2}{m}F(1/u) - h^2u^2},$$

and thus

$$\text{(D)}\quad \theta = -\int \frac{h\,du}{\sqrt{\frac{2}{m}\big(E - V(1/u)\big) - h^2u^2}} \quad \text{or} \quad \theta = -\int \frac{h\,du}{\sqrt{-\frac{2}{m}F(1/u) - h^2u^2}}.$$

Squaring (C) gives

$$\text{(E)}\quad \left(\frac{du}{d\theta}\right)^2 + u^2 = \frac{2}{h^2m}\big(E - V(1/u)\big) \quad \text{or} \quad \left(\frac{du}{d\theta}\right)^2 + u^2 = \frac{-2F(1/u)}{h^2m},$$

and taking the derivative of (E) and dividing by $du/d\theta$ yields the equation

$$\text{(F)}\qquad \frac{d^2u}{d\theta^2} + u = \frac{f(1/u)}{mh^2u^2}.$$

These equations all involve the usual deviousness of Leibnizian notation; for example, the term u in (F) actually means "u as a function of θ" (i.e., $u \circ \theta^{-1}$). This facilitates a lot of manipulations of equations, but on at least one occasion (cf. page 130) it will necessitate a bit of extra care.

Although the case of an inverse square force is of most interest in terms of gravitation, other radially symmetric forces are important in physics, and we can get a good idea of the general nature of orbits under any such force. There are only a few examples where the equations for n^{th} power forces can be solved in terms of elementary functions, but they provide a good introduction to the general nature of orbits.

1. First we have the case $f(r) = mKr$, for a constant $K > 0$, treated previously in Problem 2-4. For the somewhat cumbersome solution in terms of our general equations, we have $V(r) = mKr^2/2$, so equation (D) becomes

$$\theta = -\int \left(\frac{2\bar{E}}{h^2} - \frac{K}{h^2 u^2} - u^2 \right)^{-\frac{1}{2}} du,$$

and the substitution $u = \sqrt{v}$ gives

$$\begin{aligned} \theta &= -\frac{1}{2} \int \left(\frac{2\bar{E}}{h^2} v - \frac{K}{h^2} - v^2 \right)^{-\frac{1}{2}} dv \\ &= \frac{1}{2} \cdot - \int \frac{dv}{\sqrt{A^2 - (v - B)^2}} \end{aligned}$$

for

$$A^2 = \frac{\bar{E}^2}{h^4} - \frac{K}{h^2} \quad \text{and} \quad B = \frac{\bar{E}}{h^2}.$$

Thus we have

$$2\theta = \arccos\left(\frac{v - B}{A} \right),$$

where we have replaced the more precise $\theta + \theta_0$ simply by θ, as we will continue to do for the other examples, since this just amounts to a rotation of our axes. We can write this as

$$\text{(a)} \qquad \frac{1}{r^2} = B + A \cos 2\theta,$$

which is an ellipse *centered at the origin* (Problem 2).

2. For the case $f(r) = mK/r^2$, for gravitational attraction, it is easiest to use equation (F), which simply becomes

$$\frac{d^2 u}{d\theta^2} + u = \frac{K}{h^2},$$

with the solution $u = \dfrac{K}{h^2} + (\text{constant}) \cdot \cos\theta$, which it will be convenient to write as $u = \dfrac{K}{h^2} + KA\cos\theta$. We can assume that $A > 0$, since this just amounts to replacing θ with $\theta + \pi$, and we can write our solution as

(b) $$r = \frac{h^2/K}{1 + h^2 A \cos\theta},$$

which is a conic section *with the origin as focus*, and eccentricity $\varepsilon = h^2 A$ (see Problem 3 for a review).

Substituting the solution $u = \dfrac{K}{h^2} + KA\cos\theta$ into (E) and simplifying, we obtain

$$A^2 = \frac{1}{h^4} + \frac{2\bar{E}}{h^2 K^2},$$

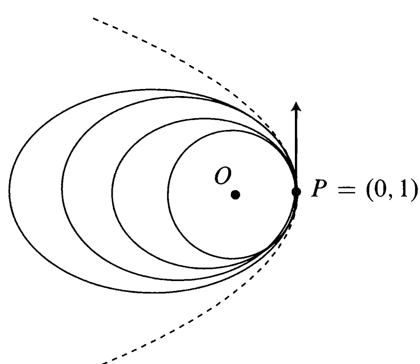
so we have

(b₁) $$\varepsilon = \sqrt{1 + 2h^2 \bar{E}/K^2}.$$

This shows that

$$\text{the orbit is} \begin{cases} \text{an ellipse} & \text{if } E < 0, \\ \text{a parabola} & \text{if } E = 0, \\ \text{a hyperbola} & \text{if } E > 0. \end{cases}$$

In the figure below, for a particle of mass $m = 1$ with $V(r) = -1/r$, on the circular orbit of radius 1 we have $V = -1$. Equation (B_ρ) gives $h^2 = 1$, so $\theta' = 1$ and the kinetic energy is $\frac{1}{2}$, so that $E = -\frac{1}{2}$, which we could also have obtained from (b_1), since the circle has eccentricity $\varepsilon = 0$. As we increase the initial velocity at P, with h increasing from 1 to $\sqrt{2}$, we get ellipses with energy approaching 0, and thus a parabola, with eccentricity 1; and for larger initial velocities we get hyperbolas.



When the conic section (b) is an ellipse, the semiaxes a and b are given (Problem 3) by

$$\text{(b}_2\text{)} \qquad a = \frac{h^2/K}{1-\varepsilon^2} \quad \text{and} \quad b = a\sqrt{1-\varepsilon^2},$$

while (b_1) can be written $1-\varepsilon^2 = -2h^2\bar{E}/K^2$, so that we have

$$\text{(b}_3\text{)} \qquad a = -\frac{K}{2\bar{E}}.$$

Consequently, the total area of the ellipse is

$$\text{(b}_4\text{)} \qquad \pi ab = \pi a^2\sqrt{1-\varepsilon^2} = \frac{\pi h K}{\sqrt{(-2\bar{E})^3}}.$$

Since the area of the graph of the function $r(\theta)$ between θ_0 and θ_1 is given by $\frac{1}{2}\int_{\theta_0}^{\theta_1} r^2\,d\theta$, the area of the ellipse swept out from time t_0 to t_1 is

$$\frac{1}{2}\int_{t_0}^{t_1} r(t)^2\theta'(t)\,dt = \frac{h}{2}(t_1 - t_0).$$

So the "period" of the orbit, the time τ for an orbit to be completely covered once, is given by

$$\tau = \frac{2}{h}\cdot \text{area of ellipse} = \frac{2\pi K}{\sqrt{(-2\bar{E})^3}} \qquad \text{by (b}_4\text{)},$$

depending only on the energy $\bar{E}$ of the orbit. Moreover, (b_3) also allows this to be written as

$$\tau = 2\pi\sqrt{\frac{a^3}{K}},$$

depending only on the length of the semimajor axis.

In the case of a force of magnitude GMm/r^2, with M being the mass of the sun, and G the "universal constant" in the law of gravitation, we have

$$\tau = 2\pi\sqrt{\frac{a^3}{GM}} \quad \text{or} \quad \frac{2\pi GM}{\sqrt{(-2\bar{E})^3}}.$$

From the first of these we have Kepler's Third Law: The squares of the periods of the plancts are proportional to the cubes of the major axes of their orbits;

conversely, the observational evidence for Kepler's Third Law shows that the forces on the various planets must involve the same constant G for all of them.

Hyperbolic orbits are of some importance in studying particles around the sun that come from or escape to outer space, but are most interesting in regard to Rutherford's early investigations of the structure of the atom, as described in Addendum C, where the appropriate formulas for such orbits are presented.

3. The case $f(r) = mK/r^3$ can also be solved explicitly. It is convenient to start directly with (A), which becomes

$$r'^2 = (K - h^2)\frac{1}{r^2} + C. \tag{c}$$

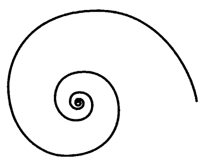
For $C = 0$, we must have $h^2 \le K$, and we can divide by $\theta' = h/r^2$ to get

$$\frac{dr}{d\theta} = r \cdot \frac{\sqrt{K-h^2}}{h},$$

giving

$$r = a \cdot e^{\gamma\theta} \qquad \text{for } \gamma = \sqrt{K-h^2}/h.$$

For $K = h^2$, or $\gamma = 0$, we get a circle. Otherwise, we obtain a *logarithmic spiral*, also called an *equiangular spiral* (Problem 6), which spirals around infinitely often as it approaches the origin, as well as when it approaches infinity, with r growing monotonically in each direction.



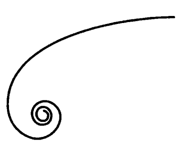
For the general case of our equation (c) with $C \neq 0$, we first have to consider yet another special case, namely $K = h^2$. Then $r' = C$ and we have

$$dr/d\theta = (dr/dt)\big/(d\theta/dt) = C/\theta' = \frac{C}{h}r^2,$$

and thus

$$d\theta = \frac{h}{C}\frac{dr}{r^2}, \qquad \theta = -\frac{h}{C}\frac{1}{r}.$$

The equation $r = \text{constant}/\theta$ is a *hyperbolic spiral*, also known as a *reciprocal spiral* (Problem 7).



On the other hand, if $K \neq h^2$, there will be r_0 with $r'(r_0) = 0$ (a point either at minimum or maximum distance from the origin). For $K > h^2$, we write (c) as

$$r'^2 = (K - h^2)\left(\frac{1}{r^2} - \frac{1}{r_0{}^2}\right) \qquad (r \leq r_0).$$

Dividing by $\theta' = h/r^2$ we obtain

$$\frac{d\theta}{dr} = -\frac{hr_0}{\sqrt{K - h^2}}\frac{r}{\sqrt{r_0{}^2 - r^2}},$$

with the solution

$$\theta = \frac{h}{\sqrt{K - h^2}}\cosh^{-1}\frac{r_0}{r},$$

spiraling around the origin as it approaches it. Part (a) of the figure below shows the orbit from $t = 0$ to $t = \infty$, while (b) adds the orbit from $t = -\infty$ to $t = 0$.

(a) (b)

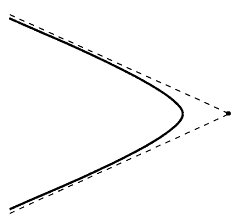
These curves are sometimes known as *Cotes' spirals* (Problem 8).

If $K < h^2$, we obtain instead

$$\frac{d\theta}{dr} = \frac{hr_0}{\sqrt{h^2 - K}}\frac{1}{r\sqrt{r^2 - r_0{}^2}},$$

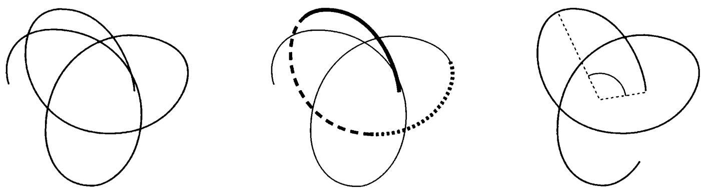
with a solution that goes off to infinity for θ in a finite interval,

$$\theta = \frac{h}{\sqrt{h^2 - K}}\arccos\frac{r_0}{r}.$$



Newton's investigation of inverse cube forces in the Principia is split into two parts. Early on, Newton determines geometrically that an inverse cube force is required to produce an equiangular spiral as an orbit, just before determining geometrically that a force proportional to distance is needed to produce an elliptical orbit centered at the origin (this is the result referred to in Problem 2-4). The general case of an inverse cube force was handled only later, after he had the equivalent of the equations in this chapter. Oddly enough, although Newton used those results to investigate inverse cube forces in general, he never bothered to use the general results to redo the case of the inverse square force—see Chandrasekhar [2; pp. 172–180] for an illuminating discussion.

The case $f(r) = 1/r$ (where $F = \log r$ isn't even a power function) cannot be solved explicitly, and it usually isn't even considered in physics texts, because forces of this sort normally arise from a line source, rather than a point source. But a graph of the solution gives a good illustration of the general nature of orbits. The orbit, shown in (a), can be constructed, as shown in (b), from a single piece that is reflected over and over again. This basic piece goes between



(a) (b) (c)

two apsides, an *apsis* being a point of maximum or minimum distance from the center—a point at minimum distance is a *periapsis* or *pericenter*, a point at maximum distance an *apoapsis* or *apocenter*.[1] The angle shown in (c) between two apsides is called the *apsidal angle*.

For the case $f(r) = mr$, the apsides are the ends of the two axes, with apsidal angle $\pi/2$; for the case $f(r) = m/r^2$, hyperbolic and parabolic orbits have only a pericenter, while an ellipse with focus at the origin has the two ends of the major axis as the apsides, with apsidal angle π; for the case $f(r) = m/r^3$, our first two solutions have no apsides, our third solution has only an apocenter, and the fourth has only a pericenter. Many of these features are often discussed in terms of a "reduction to a one-dimensional problem", as in Addendum B.

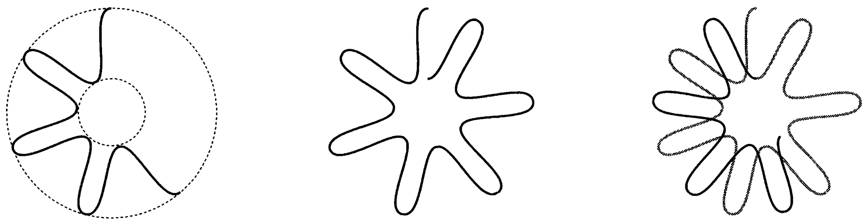
For the general case, note that equation (F) has the form

$$\frac{d^2u}{d\theta^2} = g(u)$$

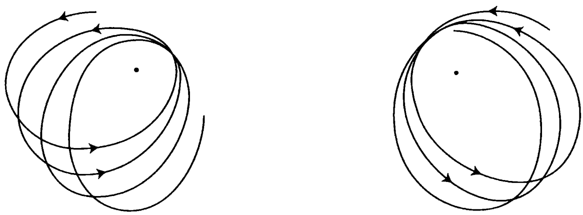
and we have $du/d\theta = 0$ at an apsis point $\theta = \theta_0$. Changing $\theta_0 + \theta$ to $\theta_0 - \theta$ produces the same equation, with the same initial condition $du/d\theta = 0$ at θ_0,

[1] The terms *perigee* and *apogee* are used for the moon revolving around the earth, *perihelion* and *aphelion* for the planets around the sun, and a flock of analogous terms are used for other astronomical bodies.

which shows that the orbit is symmetric with respect to the line drawn from the origin to an apsis, so every orbit with two apsides can be obtained by this general method of reflecting a basic piece. The basic portion of the orbit need not be concave with respect to the origin, so a more complicated figure is needed to



give some idea of the general orbit. On the other hand, the figure below shows an orbit that is concave with respect to the origin, for the force $f(r) = r^{-2.1}$, looking vaguely like an ellipse revolving around its focus, together with the orbit for the force $f(r) = r^{-1.9}$, revolving in the other direction.



"The motion of bodies in mobile orbits, and the motion of the apsides". That is the title of the section that Newton presents after deriving the basic equations for central force motion and applying it to the inverse cube law. Consider a particle moving in an orbit around the origin of a central force, and now suppose that the orbit itself is revolved around that origin in some way. If our original orbit is

$$c(t) = r(t) \cdot (\cos\theta(t), \sin\theta(t)),$$

with constant $h = r^2\theta'$, our "revolving orbit" can be described as

$$\tilde{c}(t) = r(t) \cdot (\cos\tilde{\theta}(t), \sin\tilde{\theta}(t))$$

for some function $\tilde{\theta}$; for simplicity we assume that $\tilde{\theta}(0) = \theta(0)$. In order for $r^2\tilde{\theta}'$ also to be constant, so that the revolving orbit is also due to a central force, we must have $\tilde{\theta}' = \alpha\theta'$ for some constant α, and thus $\tilde{\theta} = \alpha\theta$, so that we can write

$$\tilde{c}(t) = r(t) \cdot (\cos\tilde{\theta}(t), \sin\tilde{\theta}(t)) \qquad \tilde{\theta}(t) = \alpha\theta(t),$$

where now $r^2\tilde{\theta}' = \tilde{h} = \alpha h$.

From equation (F) for the original orbit c,

$$(\mathrm{F}_c) \qquad \frac{d^2u}{d\theta^2} + u = \frac{f(1/u)}{mh^2u^2}$$

we can derive a corresponding equation for the revolving orbit $\tilde{c}$, and thus determine the force $\tilde{f}$ needed to produce this orbit. We have[1]

$$\text{(a)} \qquad \frac{du}{d\tilde{\theta}}(\phi) = \frac{1}{\alpha}\frac{du}{d\theta}(\alpha\phi) \quad \text{and then} \quad \frac{d^2u}{d\tilde{\theta}^2}(\phi) = \frac{1}{\alpha^2}\frac{d^2u}{d\theta^2}(\alpha\phi).$$

In the desired equation for $\tilde{c}$

$$\frac{d^2u}{d\tilde{\theta}^2} + u = \frac{\tilde{f}(1/u)}{m\tilde{h}^2u^2},$$

the term u really stands for "u as a function of $\tilde{\theta}$", so that the equation means

$$\frac{d^2u}{d\tilde{\theta}^2}(\phi) + u(\alpha\phi) = \frac{\tilde{f}\big(1/u(\alpha\phi)\big)}{m\tilde{h}^2u^2(\alpha\phi)}.$$

Substituting from the second equation in (a), and multiplying by α^2, we obtain

$$(\mathrm{F}_{\tilde{c}}) \qquad \frac{d^2u}{d\theta^2} + \alpha^2 u = \frac{\tilde{f}\big(1/u\big)}{mh^2u^2},$$

where the common argument $\alpha\phi$ has now be dropped without ambiguity. Substituting from (F_c) then gives

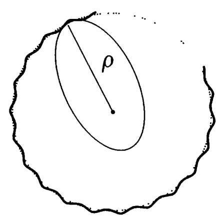
$$\begin{aligned}\tilde{f}(1/u) &= mh^2u^2\left(\alpha^2 u + \frac{f(1/u)}{mh^2u^2} - u\right)\\ &= f(1/u) + \frac{mh^2(\alpha^2-1)}{r^3},\end{aligned}$$

so that the force for a revolving version of an orbit differs from the force for the orbit itself by an inverse cube force. As usual, the m on the right side is essentially irrelevant, since f and $\tilde{f}$ will normally be taken proportional to m. This result is known, or at any rate was once known, as *Newton's theorem of revolving orbits*.

[1] To obtain (a) we can write $du/d\tilde{\theta} = (du/d\theta)\big/(d\tilde{\theta}/d\theta)$, or if necessary explicitly write $du/d\tilde{\theta} = (u \circ \tilde{\theta}^{-1})'$, etc.

Revolving orbits were introduced in the Principia in order to investigate orbits that aren't precisely elliptical, but whose apsides rotate in time, the prime example for Newton being the orbit of the moon, which is close to circular, just like all the planetary orbits. Newton used the formula just derived to approximate an almost circular orbit under any central force by a revolving orbit, starting with an elliptical orbit under an inverse square force, and then adding an appropriate inverse cube force to get a revolving elliptical orbit close to the given orbit.

To solve this problem, Newton works backwards, finding a revolving elliptical orbit for which the requisite force would be proportional to the given central force. So, for a given central force f and a particle, assumed to have mass $m = 1$ for simplicity, moving under this force in an orbit that is close to a circle of radius ρ we first choose an ellipse having the center of the force as a focus, with the end of the major axis at a point on the orbit at distance ρ from the center. This is the orbit under an inverse square force, and for simplicity we will choose our units so that it is an orbit for the force $g(r) = 1/r^2$; equation (b) on page 124 then shows that $\rho = h^2$.



Now consider the revolving orbit we obtain from this ellipse for some α, so that it is the orbit under the force

$$\tilde{g}(r) = \frac{1}{r^2} + \frac{h^2(\alpha^2 - 1)}{r^3} = \frac{1}{r^2} + \frac{\rho(\alpha^2 - 1)}{r^3}.$$

Letting k be the ratio of $f(\rho)$ to $\tilde{g}(\rho)$,

$$f(\rho) = k \cdot \tilde{g}(\rho) = k \cdot \frac{\alpha^2}{\rho^2},$$

we are now going to choose α so that $f = k \cdot \tilde{g}$ up to first order, i.e., so that we also have

$$f'(\rho) = k \cdot \tilde{g}'(\rho).$$

This means that we want

$$f'(\rho) = k \cdot \frac{1 - 3\alpha^2}{\rho^3}$$

hence

$$(*) \qquad \alpha^2 = \frac{f(\rho)}{3f(\rho) + \rho f'(\rho)}.$$

Newton introduced this calculation with the explanation that "Orbits will acquire the same shape if the centripetal forces with which those orbits are described, when compared with each other, are made proportional at equal heights." I.e., given two central forces f_1 and f_2 with $f_1 = k \cdot f_2$, if c is an orbit for f_2, then $\gamma(t) = c(\sqrt{k} \cdot t)$, with only a multiplicative change of parameter, will be an orbit for f_1. Since $f = k \cdot \tilde{g}$ up to first order, Newton concludes that

> An orbit close to the circular orbit of radius ρ will be close to the revolving orbit obtained for this choice of ellipse and α.

This actually requires continuity with respect to the *defining equation*, the subject of Problem 12, rather than continuity with respect to initial conditions, which is what differential equation texts customarily prove.

Moreover, since the ellipse has the apsidal angle π and thus the apsidal angle of the rotating orbit is $\pi\alpha$, Newton further concludes that

> The apsidal angle of our given orbit must be close to $\pi\alpha$ for this α.

This is the statement most often found in modern texts, with an argument not relying on revolving orbits at all. We begin with equation (B)

$$r'' = -f \circ r + \frac{h^2}{r^3},$$

together with equation (B_ρ), that for the circular orbit of radius ρ we have

$$h^2 = \rho^3 f(\rho).$$

Writing $r(t) = \rho + x(t)$ for a small "perturbation" $x(t)$, we get the equation

$$(\mathrm{P}) \qquad x''(t) + g(x(t)) = 0,$$

where

$$g(y) = f(\rho + y) - \frac{\rho^3 f(\rho)}{(\rho + y)^3},$$

with

$$(\mathrm{P}_0) \qquad g(0) = 0, \qquad g'(0) = \frac{3f(\rho) + \rho f'(\rho)}{\rho}.$$

If g is linear, $g(y) = g'(0)y$, then the solutions of (P) with $x(0) = 0$ are just multiples of $x(t) = \sin\big(\sqrt{g'(0)} \cdot t\big)$, with a semiperiod of $\pi \big/ \sqrt{g'(0)}$ (we're assuming $g'(0) > 0$; the case $g'(0) \leq 0$ will be discussed later). So it would seem that the semiperiod σ of a small solution x of (P) ought to satisfy

$$(**) \qquad \sigma \approx \pi/\sqrt{g'(0)} = \pi/\sqrt{(3f(\rho)+\rho f'(\rho))/\rho}.$$

Taking into account the fact that the radial speed θ' satisfies

$$\theta' = \frac{h}{\rho^2} = \frac{f(\rho)}{\rho},$$

the approximation $(**)$ for σ is consistent with equation $(*)$ for α. To complete the argument rigorously we need the following result, which is hardly ever explicitly stated, let alone proved.

PERIOD LEMMA. For small x satisfying

$$x''(t) + g(x(t)) = 0 \qquad\qquad g(0)=0, \quad g'(0)>0$$

the semiperiod of x various continuously with x [i.e., varies continuously with $x'(0)$ for solutions with $x(0)=0$], and approaches $\pi/\sqrt{g'(0)}$ as x approaches 0.

PROOF.[1] Without loss of generality we can assume that $g'(0)=1$, by considering $x\left(\sqrt{g'(0)}\cdot t\right)$.

Let G be the function $G(x) = \int_0^x g$, i.e., the function with $G' = g$ and $G(0)=0$, and define the function η by $\eta(t) = (\operatorname{sgn} t)\cdot\sqrt{2G(t)}$. Then η is differentiable at 0; in fact, for the left- and right-hand derivatives at 0 we have

$$[\eta'_{\pm}(0)]^2 = \lim_{h\to 0}\frac{2G(h)}{h^2} = \lim_{h\to 0}\frac{g(h)}{h} = g'(0) = 1,$$

and we can easily conclude that $\eta'(0)=1$. Thus η^{-1} is differentiable on some interval around 0, with $(\eta^{-1})'(0)=1$.

For a solution x, suppose that $x'(t_0)=0$ so that $x(t_0)$ is a relative maximum or minimum point, and let $t_1 > t_0$ be the next point where $x'(t_1)=0$, so that $x(t_1)$ is the next relative maximum or minimum point, and the semiperiod is thus $t_1 - t_0$.

For any solution x, the "energy" $\frac{1}{2}(x')^2 + G\circ x$ has derivative 0, so

$$\tfrac{1}{2}(x')^2 + G\circ x = E_x$$

[1] This proof comes from C. Chicone and M. Jacobs [1]. According to Prof. Chicone, the result, being already known, was merely given for completeness.

for a constant E_x, depending continuously on x, with $E_0 = 0$. Since

$$x' = \sqrt{2}\sqrt{E_x - G \circ x},$$

the substitution $dx = x'\,dt$ gives

$$\text{semiperiod}\, x = \int_{t_0}^{t_1} dt = \int_{x(t_0)}^{x(t_1)} \frac{dx}{\sqrt{2}\sqrt{E_x - G \circ x}}.$$

Now we use the substitution $X = \eta \circ x = \pm\sqrt{2G \circ x}$ [essentially replacing the solution x with a sine curve], with

$$dX = \frac{g}{\pm\sqrt{2G \circ x}}\, dx$$

$$dx = \pm\frac{X\, dX}{g}.$$

The limits of integration become $\pm\sqrt{2G(x(t_i))}$, and since $x'(t_i) = 0$, we have $G(x(t_i)) = E_x$, so we obtain

$$\text{semiperiod}\, x = \frac{1}{\sqrt{2}} \int_{-\sqrt{2E_x}}^{\sqrt{2E_x}} \frac{X\, dX}{g(\eta^{-1}(X))\sqrt{E_x - X^2/2}}.$$

For the limit as $x \to 0$, we use the substitution $X = \sqrt{2E_x}\sin\theta$ [essentially "blowing up" the singular solution $u = 0$] to express the semiperiod of x as

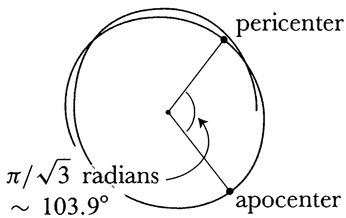
$$\int_{-\pi/2}^{\pi/2} \frac{\sqrt{2E_x}\sin\theta}{g\big(\eta^{-1}\big(\sqrt{2E_x}\sin\theta\big)\big)}\, d\theta = \int_{-\pi/2}^{\pi/2} \frac{1}{\eta'\big(\eta^{-1}\big(\sqrt{2E_x}\sin\theta\big)\big)}\, d\theta$$

$$= \int_{-\pi/2}^{\pi/2} (\eta^{-1})'\big(\sqrt{2E_x}\sin\theta\big)\, d\theta.$$

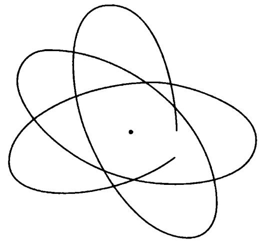
Since $E_0 = 0$, the semiperiod of x thus approaches

$$\int_{-\pi/2}^{\pi/2} (\eta^{-1})'(0)\, d\theta = \int_{-\pi/2}^{\pi/2} 1\, d\theta = \pi. \quad ❖$$

As a very simple example of these results, suppose that we have the constant force $f(r) = 1$. Then $(*)$ gives $\alpha = 1/\sqrt{3}$, so the revolving orbit must have an apsidal angle of $\pi/\sqrt{3}$. Consequently, a nearly circular orbit (a) under the constant force $f(r) = 1$ must have an apsidal angle close to this value. It



(a)



(b)

appears from some graphing experiments that the apsidal angle is fairly close even for orbits (b) not so close to circular.

More generally, for any power force $f(r) = r^n$ the formula for α is independent of ρ,

$$\alpha^2 = \frac{1}{3+n}.$$

This answer makes no sense for $n \leq -3$, and, correspondingly, the value $g'(0)$ in (P_0),

$$g'(0) = (3+n)r^{n-1}$$

is no longer positive. This is not surprising in light of our analysis of inverse cube forces: We found that all orbits were either circles or curves that spiraled into the origin or escaped to infinity; in other words, the only orbits close to circles are circles, or to put it another way, circular orbits are not stable for $f(r) = r^{-3}$, and this is in fact true for $f(r) = r^n$ whenever $n \leq -3$ (Problem 13).

By contrast, consider a force close to an inverse square force, $f(r) = r^{-(2+\varepsilon)}$ for small ε. Then the apsidal angle of a nearly circular orbit, measured in degrees, must be close to

$$\frac{180}{\sqrt{1-\varepsilon}} = 180(1-\varepsilon)^{-1/2} = 180(1 + \tfrac{1}{2}\varepsilon + \cdots) \approx 180 + 90\varepsilon.$$

So if $\varepsilon = 10^{-3}$, the apocenter would advance by about .09°, which is more than 5′ [where ′ denotes a minute, or 1/60 of a degree], easily measurable by astronomers. At the beginning of Book 3 of the Principia, Newton refers to these considerations to point out that observations show that the inverse square law must be true "with the greatest exactness. ... For the slightest departure from the ratio of the square would ... necessarily result in a noticeable motion of the apsides in a single revolution and an immense such motion in many revolutions."

The two-body problem. Newton easily disposed of the two-body problem, involving two particles c_1 and c_2 with masses m_1 and m_2, each acting on the other by a radially symmetric central force:

> *Two bodies that attract each other describe similar figures about their common center of gravity and also about each other.*

It might be entertaining to read Newton's explanation, Newton [2; pg. 561], but we will simply resort to a few formulas. Instead of writing the magnitude of the force that c_1 exerts on c_2 as $-m_2 f$, let us now simply write the equations for this force, and the opposite force that c_2 exerts on c_1, in terms of the unit vector $\mathbf{u} = (c_2 - c_1)/|c_2 - c_1|$ as

$$\begin{aligned} m_1 c_1'' &= -f(|c_1 - c_2|) \cdot \mathbf{u} \\ m_2 c_2'' &= -f(|c_1 - c_2|) \cdot -\mathbf{u} = +f(|c_1 - c_2|) \cdot \mathbf{u}, \end{aligned}$$

where f is assumed symmetric in m_1 and m_2 (normally simply involving the factor $m_1 m_2$). We immediately obtain

$$(c_1 - c_2)'' = \left(\frac{1}{m_1} + \frac{1}{m_2}\right) \cdot -f(|c_1 - c_2|) \cdot \mathbf{u} = -\frac{m_1 + m_2}{m_1 m_2} f(|c_1 - c_2|) \cdot \mathbf{u},$$

which shows that $c_1 - c_2$ (the motion of c_1 about c_2 in Newton's statement) is given by an equation of the exact same form, for a particle with the *reduced mass* $\mu = m_1 m_2/(m_1 + m_2)$—see Problem 10 for a specific example of how this works out.

Moreover, taking the center of mass $C = (m_1 c_1 + m_2 c_2)/(m_1 + m_2)$ as the origin of an inertial system, we have

$$\begin{aligned} c_1 - C &= \frac{m_2}{m_1 + m_2}(c_1 - c_2), \\ c_2 - C &= -\frac{m_1}{m_1 + m_2}(c_1 - c_2), \end{aligned}$$

so $\gamma_i = c_i - C$ satisfy

$$\frac{m_1 + m_2}{m_2}\gamma_1 = -\frac{m_1 + m_2}{m_1}\gamma_2 = (c_1 - c_2),$$

and thus the path of each particle with respect to the center of mass is a similar orbit. Note that the two paths will generally lie in different planes through C.

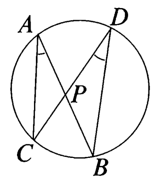
In the case of the moon and the earth, the center of mass lies a little below the earth's surface on the side facing the moon, and the earth is rotating about this point as the moon rotates about this same center of mass; for the period of the earth's rotation about the center of mass see Problem 10(d).

"The attractive forces of spherical bodies". This is the title of the section of the Principia that follows this initial analysis of the two-body problem. Although that analysis addresses one problem with the purely theoretical analysis for the one-body problem, it still requires that we consider only *particles*, bodies which can essentially be regarded as point masses. But the attraction of an object toward the earth, for example, is the result of its attraction toward myriad particles within the earth, not only at varying distances, but of varying density. In fact, the section of the Principia that treats the two-body problem ends with the words "Let us see, therefore, what the forces are by which spherical bodies, consisting of particles that attract in the way already set forth, must act upon one another, and what sorts of motions result from such forces."

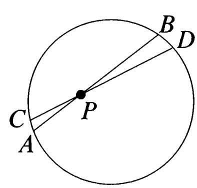
In essence, Newton showed that the inverse square force attraction for particles holds just as well for spherical bodies. Moreover, these bodies need not have uniform density; it is only necessary that their densities are spherically symmetrical around their centers, which is a good rough approximation even for complicated bodies like the earth. These results were apparently a pleasant surprise for Newton, who at first suspected that for spherical bodies there would only be a close approximation to an inverse square force at large separations.

Newton begins by considering a (2-dimensional) sphere whose mass m is uniformly distributed over its surface, and a particle P that is attracted toward the various points of the sphere by a force inversely proportional to the square of its distance from P to that point. Newton first proves that the total force on P is 0 when P is inside the sphere, and his geometric proof is both so simple and so alluring that it is was once the proof of choice.

Recall that for the two intersecting segments of a circle shown below, triangles



APC and BPD are similar, for $\angle A = \angle D$ since they subtend the same arc $\widehat{CB}$. Now consider a point P inside a sphere, and draw AB and CD through P intersecting this sphere in very small arcs AC and D. Since triangles PBD and

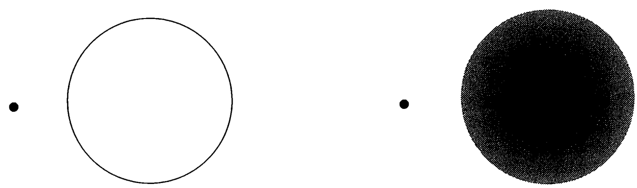


PAC are similar, we have

$$\frac{BD}{PB} = \frac{AC}{PA} \implies \frac{\overline{BD}^2}{\overline{PB}^2} = \frac{\overline{AC}^2}{\overline{PA}^2}.$$

When we rotate our lines around the angle bisector of $\angle P$, we get a corresponding 3-dimensional picture in which $\overline{BD}^2$ is close to the area of the portion of the sphere that is cut off on one side, while $1/\overline{PB}^2$ is the factor by which points in this portion attract P; and $\overline{AC}^2/\overline{PA}^2$ similarly represents the force by which P is attracted in the opposite direction. So these forces cancel out, and, in Newton's words "by a similar argument, all the attractions throughout the whole spherical surface are annulled by opposite attractions."

Newton next considers the case where P lies outside the sphere, as in (a) of the figure below, and shows that it acts as if it were attracted by a particle at the center of the sphere with the total mass m of the sphere. Newton analyses this case with an even more ingenious geometric argument that textbooks



(a) (b)

never use (see Chandrasekhar [2; pp. 270–273] for Newton's proof and further discussion), usually resorting instead to a fairly straightforward integration, as in Problem 16. From there it is straightforward to extend the result to the case (b) of the gravitational force exerted by any 3-ball whose density various only with the distance from the center. But a more elegant treatment is available using the Divergence Theorem, a.k.a. Gauss' Theorem, Ostrogradsky's Theorem, Green's Theorem, which is more frequently mentioned in connection with electric fields.

Recall that for the vector field $X = \sum_{i=1}^{n} a^i \partial/\partial x^i$ (where $(x^1, \dots, x^n)$ is the standard coordinate system on $\mathbb{R}^n$), the **divergence** of X is defined by

$$\operatorname{div} X = \sum_{i=1}^{n} \frac{\partial a^i}{\partial x^i}$$

and for a compact n-dimensional manifold-with-boundary $B \subset \mathbb{R}^n$, with outward pointing unit normal vector ν on ∂B, we have the Divergence Theorem[1]

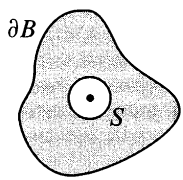
$$\int_B \operatorname{div} X \, dV_n = \int_{\partial B} \langle X, \nu \rangle \, dV_{n-1},$$

where dV_n is the n-dimensional volume element on B, and dV_{n-1} is the $(n-1)$-dimensional volume element on ∂B.

We compute that $\operatorname{div} X = 0$ for the vector field $X(p) = p/|p|^n$ in $\mathbb{R}^n$, and in particular, $\operatorname{div} X = 0$ for the vector field X in $\mathbb{R}^3$

$$X(p) = \frac{p}{|p|^3},$$

which is just a radial vector field whose length is inversely proportional to $1/|p|^2$. So if $B \subset \mathbb{R}^3$ contains a sphere S around the origin, applying the divergence theorem to $B -$ (interior of S), and noting that on S the "outward pointing



normal" ν is actually inward pointing, we obtain

$$\int_{\partial B} \langle X, \nu \rangle \, dA = -\int_S \langle X, \nu \rangle \, dA = 4\pi,$$

where the same dA is being used to denote the 2-dimensional volume element on both ∂B and the sphere S.

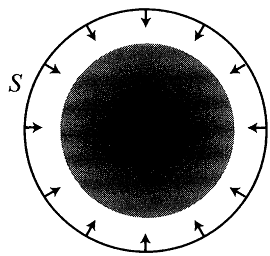
Applying this to the vector field

$$X(p) = -Gm \frac{p}{|p|^3}$$

giving the gravitational force exerted by a particle of mass m at the origin, we see that the *flux* $\int_{\partial B} \langle X, \nu \rangle \, dA$ of X through ∂B is $-4\pi Gm$. In physics texts this special case is often called Gauss' Law, and an elementary proof is often provided—see Problem 19. The same result holds for the field produced by any collection of points surrounded by ∂B, and even for a 3-dimensional collection of points surrounded by ∂B, where we specify a density rather than individual masses: the flux through ∂B is always $-4\pi GM$, where M is the total mass.

[1] See, e.g., DG, Vol. 1.

Now consider a 3-ball whose density various only with the distance from the center, and a sphere $S = \partial B$ of radius R, with the same center, surrounding it.



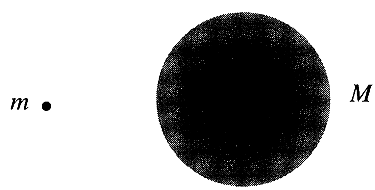
By symmetry, the gravitational field X produced by the ball must always point toward the origin and have the same magnitude μ at all points of S. So

$$-4\pi GM = \int_S -\mu \, dA = -4\pi R^2 \mu,$$

and thus the gravitational force of the ball on a particle of mass m at distance R from the center of the ball must have magnitude

$$\frac{GmM}{R^2}.$$

Finally, we leave it to the reader to conclude first, that for a particle of mass m, and a radially symmetric 3-ball of mass M, the total force of the particle on the



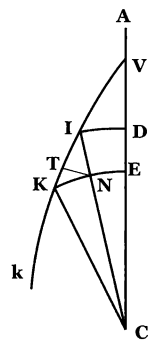
ball (the sum of the forces of the particle on each of the particles of the ball) also has magnitude GmM/R^2, where R is the distance from the particle to the center of the ball; and second, that for any two such balls, of masses M_1 and M_2, the total force of each on the other has magnitude GM_1M_2/R^2, where R is the distance between their centers.

ADDENDUM 4A
À LA *PRINCIPIA*

Although Newton never stated "conservation of energy" or even gave a name to the quantity we call kinetic energy, he essentially recognized it in a pair of Propositions. The second of these corresponds to the calculation on page 88:

> Proposition 40. *If a body, under the action of any centripetal force, moves in any way whatever and another body ascends straight up or descends straight down, and if their velocities are equal in some one instance in which their distances from the center are equal, their velocities will be equal at all equal distances from the center.*

As usual, the proof (Newton [2; pg. 528]), with a geometric diagram, contains



no equations and almost no symbols, but one can see that it is equivalent to the few lines on page 88 (in comparing the force DE on the body descending straight down with the force IN on the body moving on the path VITKk, Newton decomposes IN as IT plus TN, where TN is perpendicular to the path, and notes that it doesn't affect the motion, corresponding to the next-to-last step of the equation on pages 88).

Newton also points out the result that we have discussed on pages 92–93:

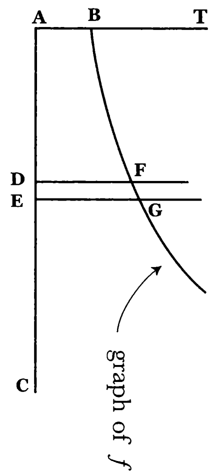
> Corollary 1. Hence if a body either oscillates while hanging by a thread or is compelled by any very smooth and perfectly slippery impediment to move in a curved line, and another body ascends straight up or descends straight down, and their velocities are equal at any identical height, their velocities at any other equal heights will be equal. For the thread of the pendent body or the impediment of an absolutely slippery vessel produces the same effect as the transverse force NT. The body is neither retarded nor accelerated by these, but only compelled to depart from a rectilinear course.

Moreover, Newton had essentially computed the potential energy function V for an arbitrary radially symmetric central force in his previous

> Proposition 39. *Suppose a centripetal force of any kind, and grant the quadratures of curvilinear figures; it is required to find, for a body ascending straight up or descending straight down, the velocity in any of its positions* ...

Here "grant the quadratures of curvilinear figures" means that the answer is allowed to be expressed in terms of integrals, and Newton's answer amounts to the term $-(F \circ r)$ on page 89.

Newton doesn't actually write down an integral, of course. His proof involves another complicated geometric diagram, of which we reproduce only a part, showing two positions D and E of the falling body, with the length of DF being $f(\mathrm{D})$, and similarly for EG, and the curved line being the locus of all such

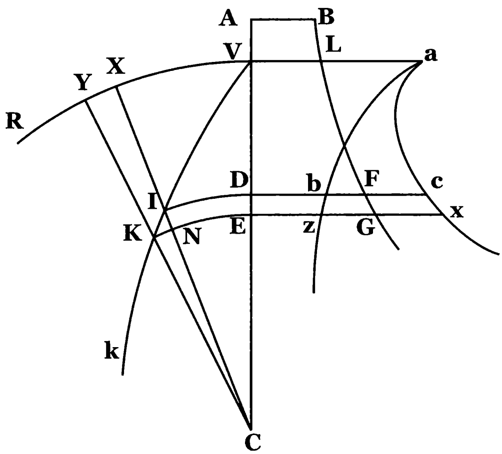


points. In other words, this is basically just the graph of f turned on its side, and Newton gives his answer in terms of the area under (i.e., to the left of) this graph. Details may be found in Chandrasekhar [2; pp. 161–163].

And finally we have

> Proposition 41. *Supposing a centripetal force of any kind and granting the quadratures of curvilinear figures, it is required to find the trajectories in which bodies will move and also the times of their motions in the trajectories so found.*

As discussed in Cohen and Whitman [1; pp. 141–142], Newton's contemporaries generally failed to appreciate the significance of this Proposition, hardly surprising, since the proof comes with a terrifying diagram—basically a combination of the diagrams for Proposition 39 and Proposition 40—and the demon-



stration is given totally in terms of complicated geometric constructions. These correspond, step-by-step, to a modern proof using integrals; a detailed account is given in Chandrasekhar [2; pp. 168–171] or Cohen and Whitman [1; pp. 334-345].

ADDENDUM 4B

REDUCTION TO A ONE-DIMENSIONAL PROBLEM

Equation (B) on page 121 can be written (taking $m = 1$ for simplicity)

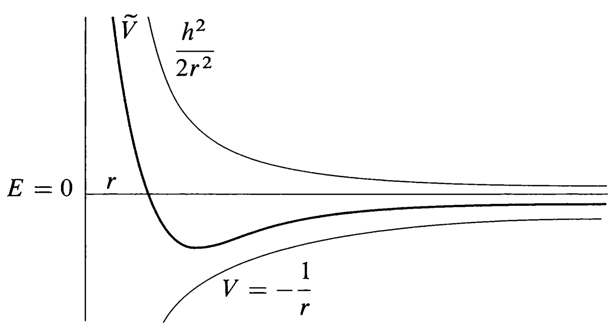
$$\begin{aligned} r'' &= -\frac{\partial V(r)}{\partial r} + \frac{h^2}{r^3} = -\frac{\partial}{\partial r}\left(V(r) + \frac{h^2}{2r^2}\right) \\ &= -\frac{\partial}{\partial r}\widetilde{V}(r) \end{aligned}$$

for the "effective potential energy" $\widetilde{V}(r) = V(r) + h^2/2r^2$, which is just the equation for a one-dimensional problem with potential energy $\widetilde{V}$. [Moreover, the energy $\widetilde{E}$ for this problem is

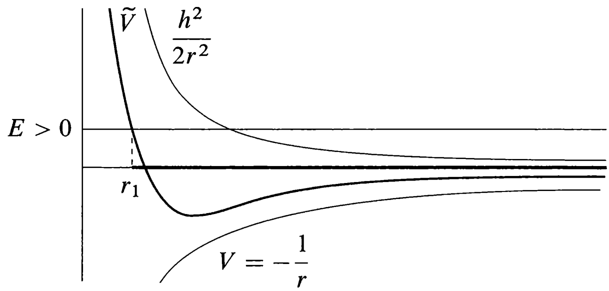
$$\begin{aligned} \widetilde{E} &= \tfrac{1}{2}r'^2 + \widetilde{V}(r) \\ &= \tfrac{1}{2}r'^2 + \frac{h^2}{2r^2} + V(r) \\ &= \tfrac{1}{2}(r'^2 + r^2\theta'^2) + V(r) \end{aligned}$$

by equation (2) on page 120, which is the same as the energy E for the original problem, by equation (1) on that page.]

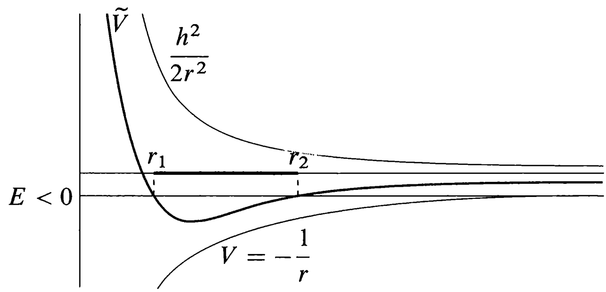
Many aspects of the general nature of orbits can be interpreted quite simply in terms of this one dimensional problem, by considering the graph of $\widetilde{V}$, as in the case of an inverse square law, shown below, with the graph of $\widetilde{V}$ decreasing from ∞ to a minimum, and then increasing asymptotically to 0. Since kinetic



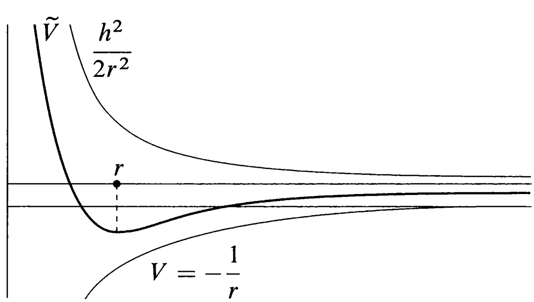
energy is non-negative, we must have $E \geq \widetilde{V}$, so if $E \geq 0$, then r must lie in an interval $[r_1, \infty)$ for some r_1, so the particle comes in from infinity to $r = r_1$, and then moves back out to infinity (hyperbolic and parabolic orbits). On the



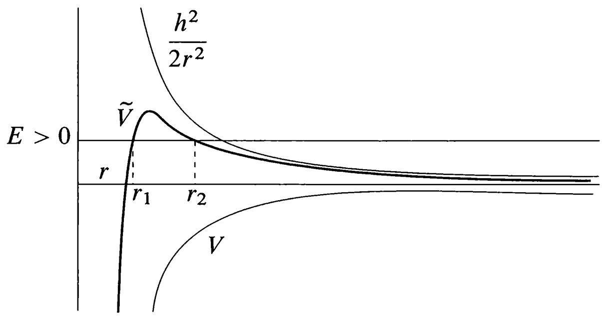
other hand, if $E < 0$, then $\widetilde{V}$ will eventually exceed E, so r will have to be in some interval $[r_1, r_2]$ (elliptical orbits). In this case the function r oscillates



between the values r_1 and r_2, the sort of motion considered in Chapter 8, see Problem 8-4. If the value of E is the minimum possible value of $\widetilde{V}$, then r can have only a single value (circular orbit).



In the diagram for an inverse fourth law, after $\tilde{V}$ reaches its maximum it decrease asymptotically to 0. For a given value of $E > 0$, the particle can never



have $r_1 < r < r_2$. If it starts with an initial value $r_0 < r_1$, it will stay in the region $r < r_1$ and eventually "fall into the center", even faster than with an inverse cube force (compare Problem 6). If it starts with $r_0 > r_2$, it stays in the region $r > r_2$ and eventually goes to infinity.

ADDENDUM 4C

RUTHERFORD SCATTERING

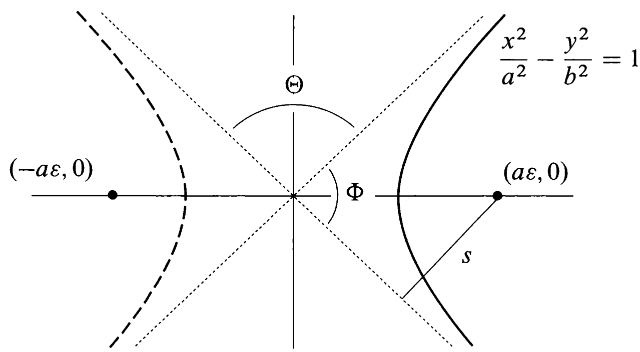
If the conic section for an inverse square force is a hyperbola, then instead of equation (b_2) on page 125 we have, by Problem 3,

(b_2')
$$a = \frac{h^2/K}{\varepsilon^2 - 1} \quad \text{and} \quad b = a\sqrt{\varepsilon^2 - 1},$$

and (b_1) on page 124 then gives

(b_3')
$$a = \frac{K}{2\overline{E}}.$$

For reasons that will appear shortly, the distance from a focus of a hyperbola to an asymptote is called the *impact parameter* s, and the angle Θ in the figure below is called the *scattering angle*; its complement Φ is the total angle through



which a particle moving on the orbit turns as it comes in from infinity and then moves out to infinity. Problem 3 shows that $\cos \frac{1}{2}\Phi = 1/\varepsilon$, so

$$\begin{aligned} b^2 = a^2(\varepsilon^2 - 1) &= a^2\left[\sec^2 \tfrac{1}{2}\Phi - 1\right] \\ &= a^2 \tan^2 \tfrac{1}{2}\Phi \\ &= a^2 \cot^2 \tfrac{1}{2}\Theta, \end{aligned}$$

and thus

$$b = \frac{K}{2\overline{E}} \cot \tfrac{1}{2}\Theta.$$

On the other hand, a direct calculation shows that the impact parameter s is simply b. So

$$s = \frac{K}{2\bar{E}} \cot \tfrac{1}{2}\Theta,$$

giving the scattering angle Θ in terms of the impact parameter s and the energy per unit mass $\bar{E}$.

These formulas all hold even if $K < 0$, the only difference being that, as indicated in Problem 3, the solution

$$\text{(b)} \qquad r = \frac{h^2/K}{1 + h^2 A \cos\theta}$$

is now always a hyperbola, though it is now the dashed branch in the figure on the previous page.

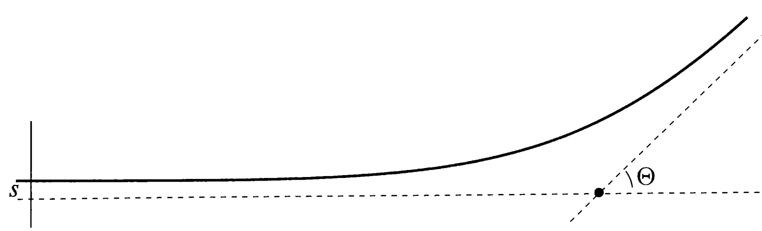
In Rutherford's experiments, a uniform beam of α-particles (helium nuclei) of known energy was directed at the (much heavier) atoms in a stationary piece of gold foil. If q_2 is the charge on the nucleus of the gold atom, and q_1 the charge on one of the particles, then the repulsive force between them is $q_1 q_2 / r^2$ (up to a constant depending on the units of charge and force). In this formula, the force on an α-particle of mass m *isn't* proportional to m, so we should really write it as

$$f(r) = \frac{mK}{r^2} = \frac{m(q_1 q_2/m)}{r^2},$$

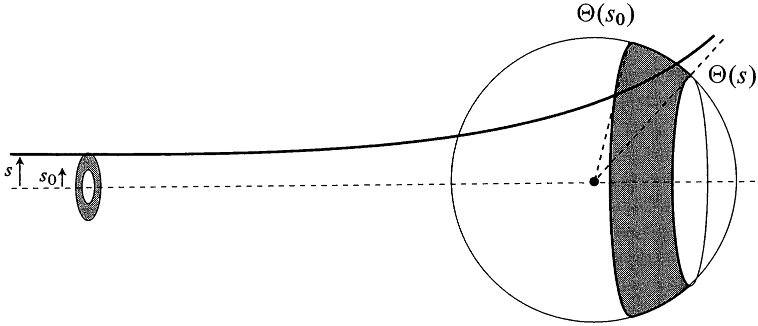
and the above formula for the impact parameter s becomes

$$\text{(S)} \qquad s = \frac{q_1 q_2}{2E} \cot \tfrac{1}{2}\Theta.$$

The figure below shows an α-particle approaching the nucleus from a distance, with impact parameter s. Its path is the hyperbola with one asymptote being the horizontal dashed line, and the other asymptote determined by the energy of the α-particle, with Θ being the angle through which the α-particle is scattered. Given α-particles of a particular energy E, we can com-



pute $\Theta(s)$ for each s, and thus determine how the number of particles varies as we vary Θ. However, we need a more realistic picture to correspond to actual experiments, where we want to measure the density of particles per unit area of a sphere around the nucleus for the scattered α-particles from a beam initially moving in the direction of the horizontal line.



A stream of α-particles with impact parameters between s_0 and $s > s_0$ will have scattering angles between $\Theta(s)$ and $\Theta(s_0)$, with $\Theta(s) < \Theta(s_0)$, since a larger impact parameter implies a smaller scattering angle, as shown by (S). If there are N particles per unit area, then the number with impact parameter between s_0 and s is $2\pi N(s - s_0)$, which must equal the number with scattering angles between $\Theta(s)$ and $\Theta(s_0)$. If $\sigma(\theta)$ is the density of scattered particles at points of the sphere with a given θ, so that $N\sigma(\theta)$ is the number per unit area at these points, then calculating the integral of $\sigma\, da$ over the corresponding region of the sphere by Fubini's theorem (compare Problem 16), we have

$$2\pi N(s - s_0) = \int_{\Theta(s)}^{\Theta(s_0)} N\sigma(\theta) \cdot 2\pi \sin\theta\, d\theta$$

and thus, taking derivatives, we have

(1) $$s = -\sigma(\Theta) \sin\Theta \frac{d\Theta}{ds}$$

(as usual with Leibnizian notation, Θ really denotes $\Theta(s)$, etc.). Similarly, writing (S) as

(2) $$\cot \tfrac{1}{2}\Theta = \frac{2Es}{q_1 q_2}$$

and differentiating, we get

(3) $$\frac{d\Theta}{ds} = \frac{-4E}{q_1 q_2} \sin^2 \tfrac{1}{2}\Theta.$$

Substituting (S), (2), and (3) back into (1) then gives

$$\frac{q_1q_2}{2E}\cot\tfrac{1}{2}\Theta = \frac{4E}{q_1q_2}\cdot\sigma(\Theta)\sin\Theta\sin^2\tfrac{1}{2}\Theta,$$

leading to the famous Rutherford scattering formula

$$\sigma(\Theta) = \frac{1}{4}\left(\frac{q_1q_2}{2E}\right)^2\frac{1}{\sin^4\frac{1}{2}\Theta}.$$

For α-particles of energy E approaching the nucleus of a gold atom head on, the velocity is 0 at the closest approach, which must therefore be

$$\frac{q_1q_2}{E},$$

where the potential energy is E. This is pretty clearly the smallest possible value for this E for any impact parameter s, as shown explicitly by the more involved formula for arbitrary s given in Problem 24. So if r_0 is the radius of the nucleus, then the scattering formula should hold as long as

$$\frac{q_1q_2}{E} > r_0 \qquad\text{or}\qquad E > \frac{q_1q_2}{r_0}.$$

By examining the scattering results for high values of E, Rutherford was able to conclude that the radius of the nucleus must be on the order of 10^{-12} cm.

It is often pointed out that the integral of $\sigma\,da$ over the whole sphere is ∞, and the scattering formula also gives $\sigma(0) = \infty$. To make sense of this, we note that to get scattering angles Θ arbitrarily close to 0, we must have α-particles with arbitrarily large impact parameters. Thus we would have to have a beam of infinite extent; moreover, although the particles with large impact parameters have small scattering angles, most of them will completely miss any particular sphere around the nucleus, and would only be detected if we made measurements infinitely far from the nucleus.

ADDENDUM 4D

BERTRAND'S THEOREM

Bertrand's Theorem is a result of the sort that endlessly fascinates because of its elegance, simplicity, and uselessness—the paper Bertrand [1] announcing the theorem concludes: "Our illustrious Corresponding Member Mr. Tchebychef, to whom I had communicated the preceding proof, sent me the judicious observation that the theorem, although useless nowadays for the already so perfect theory of the planets, will have a useful application in extending Newton's laws of gravitation to the case of double stars."

Bertrand's Theorem states that the only central forces for which all bounded orbits are closed are multiples of either $f(r) = r$ or $f(r) = r^{-2}$; in both cases the bounded orbits are ellipses, centered at the origin in the first case, and with focus at the origin in the second. Many proofs have been given, almost all of which first show that the force law must be a power law, and then restrict the possible powers to 1 or -2, and almost all of these proofs proceed by taking only a convenient number of terms in various Taylor series, and/or Fourier series, without being overly concerned about the validity of the approximation. In fact, in physics books the usual analysis of the apsidal angle carries this to the extreme, simply replacing $x''(t) + g(x(t)) = 0$ with the equation $x''(t) + g'(0)x(t) = 0$, totally dispensing with the Period Lemma.

The following argument uses our results about the apsidal angle to carry out the first part of the proof rigorously, and then relies on an argument of Arnold [2] for the second part of the proof. We assume as hypothesis that our central force has stable circular orbits of any radius, and for simplicity, we consider particles of mass $m = 1$ in all equations.

For orbits near the circular orbit of radius ρ, the apsidal angle varies continuously, and approaches α with

$$\frac{3f(\rho) + \rho f'(\rho)}{f(\rho)} = \frac{1}{\alpha^2}.$$

But these orbits are closed only when the apsidal angle is a rational multiple of π, so by continuity the apsidal angle must actually be this α for nearby orbits. Moreover, it also follows that this apsidal angle must be the same for all ρ. Thus, for $A = 1/\alpha^2 > 0$ we have

$$\rho\frac{f'(\rho)}{f(\rho)} = A - 3$$

for all ρ, or in usual differential equation notation, with $y = f(\rho)$,

$$\frac{dy}{d\rho} = (A - 3)\frac{y}{\rho},$$

with solutions $y = k \cdot \rho^{A-3}$ for some constant k, or $f(r) = k \cdot r^{A-3}$, and we have already reduced the possibilities to multiples of power functions

$$f(r) = r^n \qquad \text{with } n > -3.$$

The case $n = -1$ can be discarded because it has an apsidal angle of $\pi/\sqrt{2}$, which is not commensurable with π. We now show that for $n > -1$ the only possibility is $n = 1$, while for $-1 > n > -3$ the only possibility is $n = -2$; to do this, we need to look at the apsidal angle for orbits that are not close to circular.

If u_{min} and u_{max} are consecutive minimum and maximum values of $u = 1/r$, we can use equation (D) on page 122 to write the apsidal angle as

$$\int_{u_{\text{min}}}^{u_{\text{max}}} \frac{h\,du}{\sqrt{2\big(E - V(1/u)\big) - h^2u^2}},$$

the sign of this quantity being irrelevant, since the apsidal angle really refers to the absolute value of the angle between the pericenter and apocenter. The substitution $v = u/u_{\text{max}}$ then changes this to

$$(*) \qquad \int_{v_{\text{min}}}^{1} \frac{h\,dv}{\dfrac{1}{u_{\text{max}}}\sqrt{2\big(E - V(1/vu_{\text{max}})\big) - h^2v^2{u_{\text{max}}}^2}}.$$

In addition, since $du/d\theta = 0$ at u_{max}, equation (C) on page 122 gives

$$(**) \qquad 2E = h^2{u_{\text{max}}}^2 + 2V(1/u_{\text{max}}).$$

Suppose first that $n > -1$, so $n + 1 > 0$. Since $V(r)$ is a positive constant times r^{n+1}, we have $V(r) \to \infty$ as $r \to \infty$, so by conservation of energy all orbits are bounded, and thus closed. We now consider orbits with $E \to \infty$, but with bounded h, obtained by starting at an initial point, say with $r = 1$, and choosing large initial velocities pointing almost completely outward. Then $1/u_{\text{max}} = r_{\text{min}} \le 1$, so equation $(**)$ implies that we must have $u_{\text{max}} \to \infty$, and thus also $v_{\text{min}} \to 0$.

We can also use $(**)$ to write the integral $(*)$ as

$$\int_{v_{\text{min}}}^{1} \frac{h\,dv}{\sqrt{h^2(1 - v^2) + \dfrac{2}{{u_{\text{max}}}^2}\big[V(1/u_{\text{max}}) + V(1/vu_{\text{max}})\big]}},$$

so by choosing large E we can get the integral arbitrarily close to

$$\int_0^1 \frac{dv}{\sqrt{1-v^2}} = \frac{\pi}{2}.$$

So the apsidal angle must be $\pi/2$. But it is also $\pi/\sqrt{3+n}$, so $n = 1$.

For $-1 > n > -3$, with $V(r) = r^{n+1}/(n+1)$, we look at orbits with negative energy E approaching 0, like the situation shown on page 124 for $n = -2$; since $-1 > n$, as long as we keep E negative our orbit will be bounded, and thus closed. For E close to 0 the quantity inside the square root sign of $(*)$ is close to

$$\begin{aligned}
-h^2v^2 - \frac{2}{{u_{\max}}^2}V(1/vu_{\max}) &= -h^2v^2 - \frac{2}{{u_{\max}}^2}\cdot\frac{1}{n+1}\left(\frac{1}{vu_{\max}}\right)^{n+1} \\
&= -h^2v^2 - v^{-(n+1)}\left[\frac{2}{{u_{\max}}^2}\cdot\frac{1}{n+1}\left(\frac{1}{u_{\max}}\right)^{n+1}\right] \\
&= -h^2v^2 - v^{-(n+1)}\left[\frac{2}{{u_{\max}}^2}V(1/u_{\max})\right],
\end{aligned}$$

and $(**)$ also shows that the quantity in brackets is close to $-h^2$, so the whole integral is close to

$$\int_0^1 \frac{dv}{\sqrt{v^{-(n+1)} - v^2}}.$$

This has the value $\pi/(3+n)$ (Problem 21), so the apsidal angle must be $\pi/(3+n)$. But it is also $\pi/\sqrt{3+n}$, so $n = -2$.

ADDENDUM 4E

POWER FORCE LAWS AND DUALITY

As we noted in Problem 2-4, Newton was able to relate elliptical orbits under a central force $f(r) = r^{-2}$, with the center being at a focal point of the ellipse, to elliptical orbits under the force $f(r) = r$, with the center of the force now being at the center of the ellipse, a sort of "duality" between the two forces. Moreover, Newton also showed (cf. page 70) that the force $f(r) = r^{-5}$ is "self-dual": the orbits for this one force with two different points as center can be the same; in fact, a circle is the orbit for this force with any point on the circle as center.

It turns out that these two results are part of a more general result, which can be formulated by considering the orbits as curves in the complex plane. We will follow the treatment in Arnold [3; pp. 95–100], which is an exposition and extension of a paper by Bohlin [1].

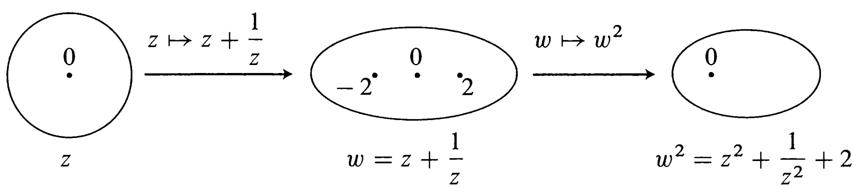
We begin by considering the map $z \mapsto z + 1/z$ of the complex plane to itself. For a point z on a circle of radius r, we have

$$z = r\cos\theta + i\,r\sin\theta$$
$$z^{-1} = r^{-1}\cos\theta - i\,r^{-1}\sin\theta,$$

so the points

$$z + z^{-1} = \left(r + r^{-1}\right)\cos\theta + i\left(r - r^{-1}\right)\sin\theta$$

are on an ellipse with semiaxes $a = r + r^{-1}$ and $b = r - r^{-1}$. The foci are the



points $(c, 0)$ for $c^2 = a^2 - b^2 = 4$, in other words the complex numbers ± 2. We can clearly get ellipses of any shape by this construction.

As indicated in the figure, squaring such an ellipse takes it into a set of exactly the same type, except moved over by 2, so that it has a focus at 0. This argument thus proves: Every ellipse with focus at the origin is the square of an ellipse centered at the origin.

We can then provide an analytic version of Newton's analysis as follows. Consider a curve $t \mapsto w(t)$ in the complex plane that moves along an ellipse centered at the origin, under the force $f(r) = r$, so that

$$\text{(a)} \qquad \frac{d^2 w}{dt^2} = -w,$$

with

$$|w|^2 \cdot \frac{d\theta}{dt} = h$$

for a constant h, where θ is the argument for w.

We consider the new curve $W(t) = w(t)^2$, and want to find a reparameterization τ so that

$$|W|^2 \cdot \frac{d\Theta}{d\tau}$$

is constant, with $\Theta = 2\theta$ being the argument of $W(t)$. So we want the quantity

$$|W|^2 \cdot \frac{d\Theta}{d\tau} = 2|w|^4 \cdot \frac{d\theta}{dt} \Big/ \frac{d\tau}{dt} = 2|w|^2 \cdot h \Big/ \frac{d\tau}{dt}$$

to be constant; we can simply take $d\tau/dt = |w|^2$, so that

$$\frac{d}{d\tau} = \frac{1}{|w|^2} \frac{d}{dt}.$$

Then we have

$$\begin{aligned}
\frac{d^2 W}{d\tau^2} &= \frac{1}{|w|^2} \frac{d}{dt} \left(\frac{1}{|w|^2} \frac{dw^2}{dt} \right) \\
&= \frac{2}{|w|^2} \frac{d}{dt} \left(\frac{1}{\bar{w}} \frac{dw}{dt} \right) \\
&= \frac{2}{|w|^2} \left(\frac{1}{\bar{w}} \frac{d^2 w}{dt^2} - \frac{1}{\bar{w}^2} \frac{d\bar{w}}{dt} \frac{dw}{dt} \right) \\
&= -\frac{2}{|w|^2} \left(\frac{1}{\bar{w}} w + \frac{1}{\bar{w}^2} \frac{d\bar{w}}{dt} \frac{dw}{dt} \right) && \text{using (a)} \\
&= -\frac{2}{w\bar{w}^3} \left(|w|^2 + \left| \frac{dw}{dt} \right|^2 \right).
\end{aligned}$$

But the term in parentheses is the constant $2E$, by conservation of energy (for the original force $f(r) = r$), so we have

$$\frac{d^2 W}{d\tau^2} = -\frac{4E}{w\bar{w}^3} = -\frac{(4E)W}{|W|^3},$$

so that W is an orbit for an inverse square force.

More generally, it is possible for the orbit under a force proportional to r^a to become an orbit for a force proportional to $r^{\bar{a}}$ under a map $w \mapsto w^\beta$: This will happen whenever

$$(*) \qquad (a+3)(\bar{a}+3) = 4, \qquad \beta = \frac{a+3}{2}.$$

For the calculations, similar to ones on the previous page, we are considering a curve w satisfying

$$(\flat) \qquad \frac{d^2 w}{dt^2} = -w|w|^{a-1},$$

with $|w|^2 \cdot d\theta/dt$ constant, as before, and the conservation of energy equation

$$2E = \left(\frac{2}{a+1}|w|^{a+1} + \left|\frac{dw}{dt}\right|^2\right).$$

We consider the curve $W(t) = w(t)^\beta$, having corresponding argument $\Theta = \beta\theta$. To make $|W|^2 \cdot d\Theta/d\tau$ constant, we take $d\tau/dt = |w|^{2(\beta-1)} = |w|^{a+1}$, or

$$\frac{d}{d\tau} = \frac{1}{|w|^{a+1}}\frac{d}{dt}.$$

Then we have

$$\begin{aligned}
\frac{d^2 W}{d\tau^2} &= \frac{1}{|w|^{a+1}}\frac{d}{dt}\left(\frac{1}{|w|^{a+1}}\frac{dW}{dt}\right) \\
&= \frac{\beta}{|w|^{a+1}}\frac{d}{dt}\left(\frac{w^{\beta-1}}{|w|^{a+1}}\frac{dw}{dt}\right) \\
&= \frac{\beta}{|w|^{a+1}}\frac{d}{dt}\left(\frac{1}{\bar{w}^{(a+1)/2}}\frac{dw}{dt}\right) \\
&= \frac{\beta}{|w|^{\beta+1}}\left(\frac{1}{\bar{w}^{(a+1)/2}}\frac{d^2 w}{dt^2} - \frac{a+1}{2}\frac{1}{\bar{w}^{-(a+3)/2}}\frac{d\bar{w}}{dt}\frac{dw}{dt}\right)
\end{aligned}$$

which, finally,

$$= -\frac{\beta(\beta-1)w^2}{|w|^{4\beta-2}}\left(\frac{2}{a+1}|w|^{a+1} + \left|\frac{dw}{dt}\right|\right),$$

after taking into account both ($\flat$) and ($*$), which gives $\frac{1}{2}(a+1) = \beta - 1$. Using the conservation of energy equation, together with ($*$) again, we finally obtain the desired result

$$\frac{d^2 W}{d\tau^2} = -2E\beta(\beta-1)W|W|^{\bar{a}-1}.$$

Our initial result relating orbits under an inverse square force to those under a direct first order force corresponds to $a = 1, \bar{a} = -2$, while Newton's result about inverse fifth powers corresponds to $a = \bar{a} = -5$ (see also Problem 9).

It should also be noted that when our calculation involves orbits with energy $E = 0$, we obtain $d^2W/d\tau^2 = 0$, which are simply straight lines. Or, working backwards, orbits of energy $E = 0$ are the images of straight lines under the map $w \mapsto w^{1/\beta}$, where

$$\frac{1}{\beta} = \frac{2}{a+3}.$$

In particular, for an inverse square force, $a = -2$, the parabolas are the images of straight lines under the map $z \mapsto z^2$.

PROBLEMS

1. (a) If the acceleration $\mathbf{a} = c''$ of a curve

$$c(t) = r(t)(\cos\theta(t), \sin\theta(t))$$

is decomposed as $\mathbf{a}_r + \boldsymbol{a}_\theta$, where $\mathbf{a}_r$ points towards the origin and $\mathbf{a}_\theta$ is perpendicular to it, show that

$$\mathbf{a}_r = r'' - r\theta'^2.$$

(b) Let $-P$ be the magnitude of a force that is not necessarily radially symmetric but is a central force (with $P > 0$ for attractive forces), so that we still have $r^2\theta' = h$ for a constant h. Show that the equation of motion can be written in terms of $u = 1/r$ as

$$\frac{d^2u}{d\theta^2} + u = \frac{P(1/u)}{mh^2u^2},$$

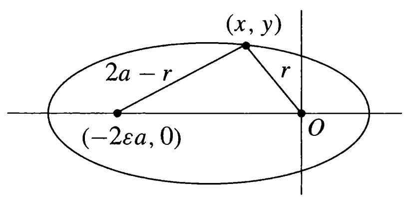
generalizing equation (F).

2. (a) The equation

$$\frac{1}{r^2} = B + A\cos 2\theta$$

is an ellipse centered at the origin when $A < B$ and a hyperbola when $A > B$.
(b) For the attractive force $f(r) = mKr$ all orbits have the same period.
(c) The orbits for the repulsive force $f(r) = -mKr$ are hyperbolas. (A modification of the argument in Problem 2-4 (a) is also possible.)

3. Consider an ellipse with one focus at the origin O, and the other at $(-2\varepsilon a, 0)$, for which the sum of the distances r and $2a - r$ from the foci to any point (x, y)



is the constant $2a$ (which requires $0 \le \varepsilon < 1$).

(a) The distance r satisfies

$$\begin{aligned} r &= (1 - \varepsilon^2)a - \varepsilon x \\ &= \Lambda - \varepsilon x, \qquad \text{for } \Lambda = (1 - \varepsilon^2)a > 0, \end{aligned}$$

and thus

$$r = \frac{\Lambda}{1 + \varepsilon\cos\theta}.$$

(b) One semiaxis of the ellipse is obviously $a = \Lambda/(1-\varepsilon^2)$. The maximum and minimum values $\pm b$ of $r\sin\theta$ occur when $\cos\theta = -\varepsilon$, so the other semiaxis b satisfies

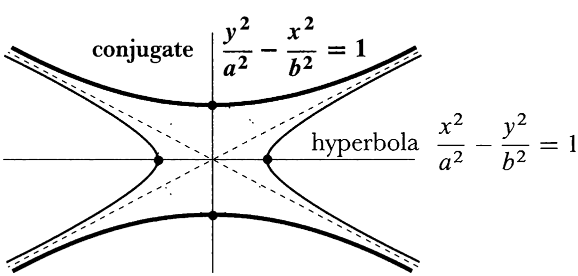
$$b = \frac{\Lambda}{\sqrt{1-\varepsilon^2}} = a\sqrt{1-\varepsilon^2}.$$

Thus, the ellipse $(x/a)^2 + (y/b)^2 = 1$ has eccentricity $\varepsilon = \sqrt{1-(b/a)^2}$.
(c) If we choose our origin at the center of the ellipse, $(-\varepsilon a, 0)$, and let $\bar{x}, y$ denote coordinates with respect to this origin, then $\bar{x} = x + \varepsilon a$. Conclude that we have

$$r = a - \varepsilon\bar{x}.$$

(d) Now consider a hyperbola with one focus at the origin, with the difference of the distances being $2a$, and the other focus at $(-2\varepsilon a, 0)$ (which requires $\varepsilon > 1$). For the branch of the hyperbola consisting of points where the distance to the origin minus the distance to the other focus is a, show that we obtain the same equation as for the ellipse, except that $\Lambda = (1-\varepsilon^2)a < 0$. On the other hand, for the other branch we get the same equation, but with $\Lambda = (\varepsilon^2 - 1)a > 0$, so that $a = h^2/(1-\varepsilon^2)$.
(e) For this branch, since we must have $r = \Lambda/(1+\varepsilon\cos\theta) > 0$, and thus $1 + \varepsilon\cos\theta > 0$, conclude that the positive angle θ that one of the asymptotes of the hyperbola makes with the x-axis satisfies $\cos\theta = 1/\varepsilon$. Comparing with the asymptotes of the hyperbola $(x/a)^2 - (y/b)^2 = 1$, conclude that this has the same shape when $b = a\sqrt{\varepsilon^2 - 1}$, so that the eccentricity is given by $\varepsilon = \sqrt{1 + (b/a)^2}$.
(f) Consider the parabola consisting of all points (x, y) whose distance from the origin is equal to the distance from the line $y = a$. Show that we obtain the equation $a = r(1+\cos\theta)$, again of the same form.

4. Every hyperbola has a *conjugate* hyperbola with the same asymptotes, the same distance between its vertices, and the same distance between its foci. In part (c) of Problem 2, show that for a given $\bar{E}$ and h, the orbit has two possible shapes, a hyperbola and its conjugate.



5. (a) For an inverse square force $f(r) = mK/r^2$, use equation (D) to get

$$\theta = \int \frac{-h\,du}{\sqrt{2\bar{E} + 2Ku - h^2u^2}}.$$

A table of integrals shows that for $a < 0$ we have

$$\int \frac{dx}{\sqrt{ax^2 + bx + c}} = -\frac{1}{\sqrt{-a}} \arcsin\left(\frac{2ax + b}{\sqrt{b^2 - 4ac}}\right).$$

Use this to obtain the orbits.
(b) Use equation (D) to show that the orbits for a repulsive inverse square force are hyperbolas, each of which is an orbit for the attractive inverse square force with the center at the focus of the conjugate hyperbola.

For the next few problems, recall that for the graph $r = f(\theta)$ in polar coordinates, the length on $[\theta_0, \theta_1]$ is given by

$$\int_{\theta_0}^{\theta_1} \sqrt{f^2 + f'^2}.$$

For simplicity we will examine the inverse cube force with $K = 1$.

6. (a) Consider the logarithmic spiral $r = ae^{\gamma\theta}$, that is, the path $c(\theta) = (ae^{\gamma\theta}\cos\theta, ae^{\gamma\theta}\sin\theta)$. Show that the radius vector from the origin to $c(\theta)$ makes a constant angle ϕ with the tangent vector at $c(\theta)$, where $\cos\phi = \gamma^2/(1+\gamma^2)$.
(b) Show that the length of the spiral from any point to the origin, i.e.,

$$\int_{-\infty}^{\theta_1} \sqrt{f^2 + f'^2}, \qquad f(\theta) = ae^{\gamma\theta}$$

is finite. It is therefore hardly surprising that the orbit "reaches" the origin (or "falls into the center") in finite time. The equation $r'^2 = (1 - h^2)/r^2$ [equation (c) on page 126, for $C = 0$] gives

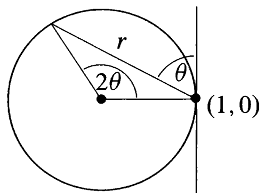
$$\frac{dr}{dt} = -\frac{\sqrt{1 - h^2}}{r}.$$

Solve for $r(t)$ and conclude that if the particle begins at distance r_0 from the center, it will fall into the center at time $t = {r_0}^2/2\sqrt{1 - h^2}$.

7. (a) The hyperbolic spiral $r = c/\theta$ does not spiral infinitely often as we go to infinity; instead it approaches the line $y = c$ asymptotically.
(b) The length of the spiral from any point to the origin is infinite.
(c) Nevertheless, the orbit on page 126 falls into the center in finite time.

8. The length of the Cotes' spirals from any point to the origin is also infinite, but the orbits again fall into the center in finite time.

9. (a) Problem 2-3 presented Newton's proof that if a particle moves in a circle under a central force directed to a point of the circle, then the force must be an inverse fifth power. To derive this from our equations, we consider a particle of mass $m = 1$ travelling on a circle of radius 1 around the origin, starting at the



point $(1,0)$. Letting θ be the angle from the vertical to the radius vector, show that

$$\begin{aligned} r(t) &= 2\sin\theta(t) \\ r'(t) &= 2\cos\theta(t)\theta'(t) \end{aligned}$$

and use the second form of equation (A) to obtain

$$\frac{4h^2}{r(t)^4} = -2F(r(t)).$$

(b) If, conversely, we start with the force $f(r) = r^{-5}$, use equation (E) to obtain

$$\left(\frac{du}{d\theta}\right)^2 = -\frac{u^4}{2h^2} - u^2 + \frac{2E}{h^2},$$

and for $r = 1/u$ obtain

$$\left(\frac{dr}{d\theta}\right)^2 = \frac{2E}{h^2}r^4 - r^2 - \frac{1}{2h^2}$$

(this has essentially already been done in the equation at the bottom of page 121).
(c) Conclude that the inversion through the origin of an orbit is also an orbit, and use this fact to give an immediate proof that the unit circle through $(1,0)$ is an orbit.

396 **10.** (a) For two particles c_1 and c_2 of masses m_1 and m_2 let $f(r) = Gm_1m_2/r^2$ be the gravitational force between them, where $r = |c_1 - c_2|$. Then the equation for $c'' = (c_1 - c_2)''$ derived on page 136 can be regarded as the equation for a particle of unit mass under an inverse square force of magnitude $G \cdot (m_1 + m_2)$. Conclude that the distance a between c_1 and c_2 and its period τ are related by

$$\tau = 2\pi \sqrt{\frac{a^3}{G \cdot (m_1 + m_2)}}.$$

(b) If M is the mass of the sun, and m_i $(i = 1, 2)$ are the masses of two planets, whose orbits have semimajor axes of length a_i and periods τ_i, then we have the more exact form of Kepler's third law

$$\frac{M + m_1}{M + m_2} = \left(\frac{a_1}{a_2}\right)^3 \left(\frac{\tau_2}{\tau_1}\right)^2.$$

(c) Let m be the mass of a planet, whose orbit around the sun has a semimajor axis of length a and period τ, and let m' be the mass of a moon of that planet, whose orbit around the planet has a semimajor axis of length a' and period τ'. Then

$$\frac{m + m'}{M + m'} = \left(\frac{a'}{a}\right)^3 \left(\frac{\tau}{\tau'}\right)^2,$$

and hence

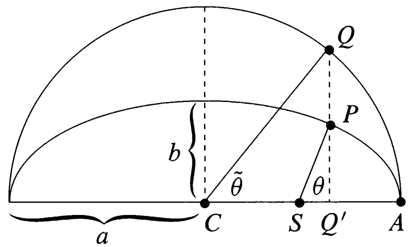
$$\frac{m}{M} \approx \left(\frac{a'}{a}\right)^3 \left(\frac{\tau}{\tau'}\right)^2.$$

This provided an accurate way to determine the mass, relative to the mass of the sun, of all planets except Mercury, Venus, and Pluto [which actually turns out to have moons]; their relative masses have now been measured by observing their effect on spacecraft.

(d) Using the equations for $\gamma_i = c_i - C$ derived on page 136, find the ratio of the periods of the two body's orbits around their center of mass in terms of the ratio of their masses and the semimajor axes of those orbits. The ratio of the mass of the moon to that of the earth is .0123, and the period of the moon's orbit is very close to 39,360 minutes. Find the period of the earth's orbit around the center of mass of the earth-moon system.

643 **11.** Determining the orbit of a planet or a comet from only a few observations requires the solution of a problem considered by Kepler, to find the time $t(\theta)$

required for a planet P to go from its perihelion A around the sun S to its



position with a particular angle $\angle ASP = \theta$. However, Kepler didn't work directly with θ. Drawing a circle whose diameter is the diameter of the elliptical orbit, with center C, and letting Q be the point of the circle on the perpendicular to the diameter through P, we consider the angle $\tilde{\theta} = \angle ACQ$ at the center C, the *eccentric anomaly*, as opposed to θ, the *true anomaly*. We will first find the connection between θ and $\tilde{\theta}$, and then derive *Kepler's equation*, which gives a formula for $t(\tilde{\theta})$.

(a) By Problem 3 we have

$$r = \frac{a(1-\varepsilon^2)}{1+\varepsilon\cos\theta}.$$

In addition, part (c) of the problem shows that for $\bar{x} = \overline{CQ'}$ we have $r = a - \varepsilon\bar{x}$. Conclude that

$$r = a(1 - \varepsilon\cos\tilde{\theta}),$$

and then that

$$(1 - \varepsilon\cos\tilde{\theta})(1 + \varepsilon\cos\theta) = 1 - \varepsilon^2,$$

which can also be written as

$$\tan\frac{\tilde{\theta}}{2} = \sqrt{\frac{1-\varepsilon}{1+\varepsilon}}\tan\frac{\theta}{2},$$

or

$$\sin\tilde{\theta} = \frac{\sqrt{1-\varepsilon^2}\sin\theta}{1+\varepsilon\cos\theta}.$$

(b) Since the area swept out by the planet is proportional to the time, and the total area πab is covered in the period τ, for the sector ASP we have

$$\begin{aligned} t(\tilde{\theta}) &= \frac{\tau}{\pi ab} \text{ area } ASP \\ &= \frac{\tau}{\pi ab} \cdot \frac{b}{a} \text{ area } ASQ \\ &= \frac{\tau}{\pi a^2} [\text{area } ACQ \ - \ \text{area } SCQ] \\ &= \frac{\tau}{\pi a^2} \left[\frac{a^2}{2} \tilde{\theta} - \frac{a^2 \varepsilon}{2} \sin \tilde{\theta} \right]. \end{aligned}$$

Recalling the formula for τ on page 125, conclude that we have Kepler's equation

$$t(\tilde{\theta}) = \sqrt{\frac{a^3}{K}} (\tilde{\theta} - \varepsilon \sin \tilde{\theta}).$$

Thus we have to solve a transcendental equation $\phi - \varepsilon \sin \phi = \text{constant}$, for which numerous numerical methods have been devised.

(c) For an analytic derivation of Kepler's equation, we start with equation (A) on page 121, and separate variables to obtain

$$t = \int \frac{dr}{\sqrt{2\bar{E} + \frac{2K}{r} - \frac{h^2}{r^2}}} = \int \frac{r\, dr}{\sqrt{2\bar{E} r^2 + 2Kr - h^2}}.$$

Use (b_1) on page 124 and (b_3) on page 125 to get

$$t = \int \frac{r\, dr}{\sqrt{\frac{-Kr^2}{a} + 2Kr - Ka(1 - \varepsilon^2)}}.$$

Now use $r = a(1 - \varepsilon \cos \tilde{\theta})$, derived in part (a), for a substitution in the integral. The denominator should simplify to $\sqrt{a}\, \varepsilon \sin \tilde{\theta}$, and you should end up with

$$t = \sqrt{\frac{a^3}{K}} \int (1 - \varepsilon \cos \tilde{\theta})\, d\tilde{\theta}.$$

12. This problem, taken from Palais and Palais [1], proves continuity of solutions of differential equations with respect to the defining equations.

(a) Let f and g be continuous functions on $[a,b]$ with g nonnegative. Show that on $[a,b]$ we have *Gronwall's inequality*:

$$\text{If } f(x) \le C + \int_a^x fg, \text{ then } f(x) \le Ce^{\int_a^x g}$$

(usually stated, and used, only for $C \ge 0$). Hint: Consider the derivative of $h(x) = \left(C + \int_a^x fg\right)e^{-\int_a^x g}$.

In particular, for $K \ge 0$

$$\text{If } f(x) \le C + K\int_0^x f, \text{ then } f(x) \le Ce^{Kx}.$$

(b) For simplicity, we will consider a system of differential equations determined by a function $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ without worrying about details concerning the actual domain of ϕ, etc. We will use $|\ |$ for the norm on $\mathbb{R}^n$ and assume that ϕ satisfies a Lipschitz condition $|\phi(x)-\phi(y)| \le K|x-y|$; solutions c_x with $c_x(0) = x$, $c_x'(t) = \phi(c_x(t))$ will be assumed extended to a maximal interval. All the considerations for these time-independent equations will also work just as well for the time-dependent case.

To make the role of the "defining equation" ϕ explicit, we will use c_x^ϕ for the curve with

$$c_x^\phi(0) = x, \qquad c_x^{\phi\prime}(t) = \phi(c_x^\phi(t)).$$

Now suppose we have $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ with

$$|\phi(x) - \psi(x)| \le \varepsilon \qquad \text{for all } x.$$

Prove that

$$|c_x^\phi(t) - c_x^\psi(t)| \le \frac{\varepsilon}{K}(e^{Kt} - 1).$$

Hint: Writing $u(t) = |c_x^\phi(t) - c_x^\psi(t)| + \frac{\varepsilon}{K}$, the conclusion may be written as $u(t) \le (\varepsilon/K)e^{Kt}$, which follows from Gronwall's inequality provided that $u(t) \le \frac{\varepsilon}{K} + K\int_0^t u(s)\,ds$. Note that

$$u(t) - \frac{\varepsilon}{K} = |c_x^\phi(t) - c_x^\psi(t)| \le \int_0^t |\phi(c_x^\phi(s)) - \psi(c_x^\psi(s))|\,ds,$$

and write $\phi(c_x^\phi(s)) - \psi(c_x^\psi(s))$ as a sum of two terms that can be estimated.

13. For the equation $x''(t) + g(x(t)) = 0$ with $g(0) = 0$ but $g'(0) \le 0$, use Problem 12 to prove that solutions x with $x(0) = 0$ will not remain small no matter how small we choose our initial value $x'(0)$.

14. Consider a central force with $f(r) = r^3 - r$, with corresponding potential determined by $F(r) = \frac{1}{4}(1 - r^2)^2$, a "Higgs potential".

(a) Show that for ρ close to 1, orbits close to the circular orbit at radius ρ have an apsidal angle α that is very small, so that the orbit oscillates around the circular orbit many times (as in the picture on page 131).

(b) But the semiperiod σ approaches $\pi/\sqrt{2}$ as ρ approaches 1; this apparent discrepancy is explained by the fact that θ' approaches 0 as ρ approaches 1, so that the orbits are being transversed more and more slowly.

That might lead us to conclude that motion on the orbit of radius 1 is infinitely slow, which is indeed true in a sense: *There is no circular orbit of radius* 1; since the force along the unit circle is 0, the only orbits that stay on the unit circle are the ones that stay at a single point of the unit circle (note that the argument on page 121 doesn't apply in this case!).

(c) But there are still orbits that oscillate around the unit circle. Such orbits can be found with arbitrarily small apsidal angles α, though always with semiperiods σ close to $\pi/\sqrt{2}$.

15. For the force

$$f(r) = \frac{1}{r^2} + \frac{C}{r^3}$$

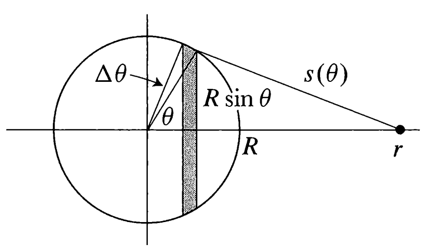
for a constant C, solve equation (E) explicitly to get

$$r = \frac{h^2 - C}{1 + A(h^2 - C)\cos\gamma\theta} \qquad \gamma = \sqrt{1 - \frac{C}{h^2}}$$

$$= \frac{a(1 - \varepsilon^2)}{1 + \varepsilon\cos\gamma\theta} \qquad \text{for an "eccentricity" } \varepsilon \text{ and "semimajor axis" } a.$$

Since the maximum of r occurs for $\gamma\theta$ a multiple of π, this is a precessing ellipse with apsidal angle of π/γ. Check that as $\varepsilon \to 0$, this approaches $\pi\alpha$ for the α given by $(*)$ on page 132.

16. Consider a particle of mass m at distance r from the center of a sphere of radius R having total mass M, and thus density $M/4\pi R^2$. Allowing ourselves



the luxury of a little rigor slippage, the mass of the shaded sector in the figure is approximately

$$\frac{M}{4\pi R^2}(2\pi R\sin\theta)(R\Delta\theta) = \frac{M}{2}\sin\theta\Delta\theta,$$

so the potential function V for the force on the particle is

$$V(r) = -\frac{GmM}{2}\int_0^{\pi}\frac{\sin\theta}{s(\theta)}\,d\theta, \qquad s(\theta)^2 = R^2 + r^2 - 2rR\cos\theta.$$

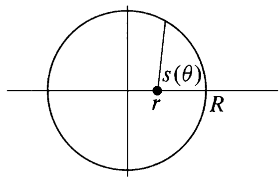
(a) Using the substitution $s^2 = R^2 + r^2 - 2rR\cos\theta$ in the integral, show that

$$V(r) = -\frac{GmM}{2rR}\int_{r-R}^{r+R} ds$$

to prove that the total force on the particle is GmM/r^2.

(b) For $r < R$ we have

$$V(r) = \frac{GmM}{2rR}\int_{R-r}^{R+r} ds,$$



so V is constant inside the sphere, and the force is 0.

17. (a) For an arbitrary potential V, written as

$$V(r) = \frac{f'(r)}{r},$$

at a point at distance r from the center of the shell the value of V will be a constant times

$$\frac{1}{2rR}[f(r+R) - f(r-R)].$$

(b) Suppose this is a constant C for $r \leq R$. Conclude that

$$C = \frac{1}{2r^2}f(2r)$$

and then that

$$Rf(2r) - rf(R+r) - rf(r-R) = 0.$$

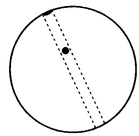
(c) Then show that

$$f''(r+R) = f''(r-R)$$

and conclude that f'' is a constant, so that V is a multiple of $1/r$.

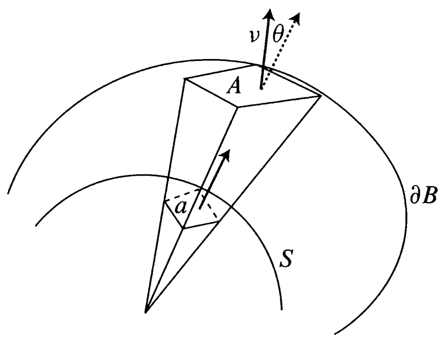
This proof, attributed to Laplace, is quoted in Maxwell [1; pg. 422].

18. A ball is dropped into a hole drilled straight through the earth. Assuming that the earth has constant density, determine how the force on the ball varies



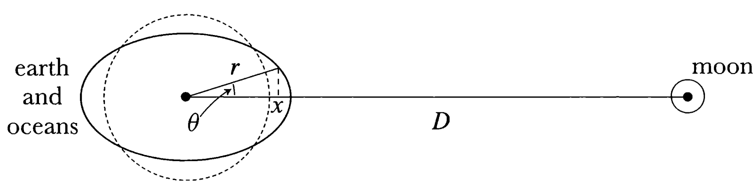
with its distance from the center of the earth, and describe its motion.

19. In the figure below, with a sphere S of radius r surrounded by the surface ∂B, a small cone from the center of S intersects S in a region a that is close to



a rectangle, and it intersects ∂B in another such region A at distance R from the origin, on which the normal ν makes an angle of θ with the normal of the region on S. If $|\mathbf{F}|$ varies inversely as the square of the distance from the center of S, show that the flux of $\mathbf{F}$ through the region a equals the flux of $\mathbf{F}$ through region A. Intuitively, if we think of particles being emitted from the origin at a steady rate, then the flux through any surface surrounding the origin is just the total number of particles emitted per unit time.

361
20. The earth's tides, or more generally, "tidal forces", are entirely due to the fact that gravitational forces are not uniform. The moon's gravitational force on the parts of the oceans nearest the moon is greater than the force on the solid part of the earth, so the water bulges towards the moon, while the force on the parts furthest from the moon are less than the force on the solid earth, so they are pulled less and bulge away from the earth. For a computational analysis, in the figure below the plane of the paper contains the centers of the



earth and moon; the dot at the middle of the earth is close to, but not exactly at, the north pole. Let m be the mass of the moon, D the distance between the center of the earth and the center of the moon, and r the distance from the center of the earth to a particle on the equator making an angle θ with the line from the earth to the moon (if our dot were the north pole, θ would be the particle's longitude), with x the distance along this axis from the center of the earth.

(a) The distance from the center of the moon to this particle is

$$(r^2 - 2rD\cos\theta + D^2)^{1/2},$$

and since $r \ll D$, for the potential function V we have

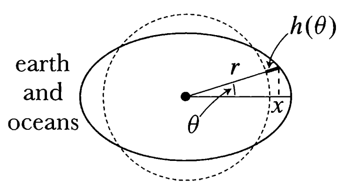
$$V \approx -Gm\left[\frac{1}{D} + \frac{r}{D^2}\cos\theta + \frac{r^2}{D^3}\left(\tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}\right)\right].$$

The constant $-Gm/D$ is irrelevant, and the term $-Gm(r/D^2)\cos\theta$, which can be written $-Gmx/D^2$, corresponds to the constant force Gm/D^2 towards the moon, which we subtract in order to see the effect with respect to the earth's position, leaving us with

$$V_{\text{observed}} \approx \frac{-GMr^2}{D^3}\left(\tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}\right).$$

Notice that this gives the potential function V_{observed} as a function of the distance r *from the earth's center.* For this approximation we can just as well let r simply denote the mean radius of the earth.

(b) To determine the height $h(\theta)$ of the tide above the average radius of the earth, we use the fact the surface of the water should be an equipotential surface



for the difference potential $V_{\text{observed}} + Gm_{\text{e}}/r^2$, where m_{e} is the mass of the earth. Using the fact that $h(\theta)$ is small compared to r, show that

$$h(\theta) \approx \frac{mr^4}{m_{\text{e}}D^3}\left(\tfrac{3}{2}\cos^2\theta - \tfrac{1}{2}\right).$$

(c) If the particle is not on the equator, but has "latitude" ϕ, then $(\frac{3}{2}\cos^2\theta - \frac{1}{2})$ should be replaced in all the formulas by $(3\sin^2\phi\cos^2\theta - 1)$.

The gravitational force of the sun on the earth is about 175 times as great as that of the moon, despite its much greater distance. But that greater distance also diminishes its tidal effect. To compare the effect of the moon and the sun, note that, apart from the factor $\frac{3}{2}\cos^2\theta - \frac{1}{2}$, the ratio $h(\theta)/r$ is $mr^3/m_e D^3$. From the data $m/m_e = 1/81.3$ and $r/D = 1/60.3$, this factor is

$$\frac{mr^3}{m_e D^3} = 5.6 \times 10^{-8}.$$

If M is the mass of the sun, and a is the semimajor axis of the earth's orbit, then $M/m_e = 3.33 \times 10^5$ and $r/a = 4.26 \times 10^{-5}$, and the factor is

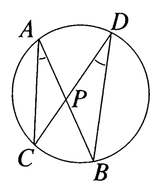
$$\frac{Mr^3}{m_e a^3} = 2.57 \times 10^{-8},$$

so the effect of the sun is somewhat less than half that of the moon.

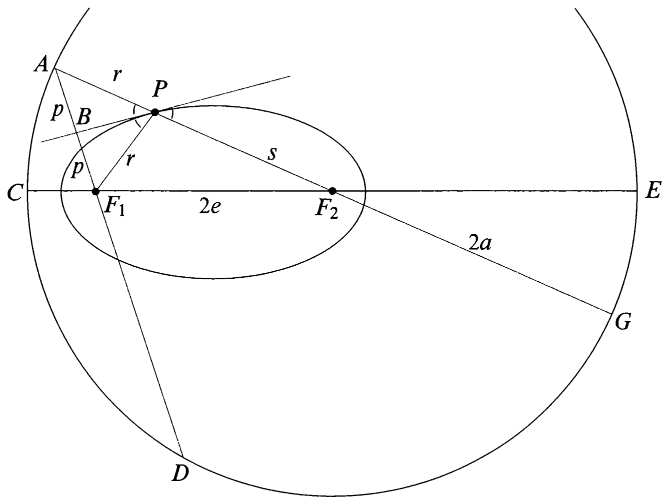
The earth of course exerts tidal forces on the moon, causing it to have a slightly ovaloid shape, with the narrower end tending to be pulled toward the center of the earth, which has resulted in the moon's revolving exactly once for each revolution around the earth, so that the same side is always facing us.

21. For $\beta > 2$, evaluate $\displaystyle\int \frac{dx}{\sqrt{x^\beta - x^2}} = \int \frac{dx}{x\sqrt{x^{\beta-2} - 1}}$ with a substitution of the form $u^2 = \dots$ to get an answer involving arctan, and with a substitution of the form $y = x^\lambda$, for suitable choice of λ, to get an answer involving arcsin.

22. This problem gives an elementary geometric argument, due to Hermann Karcher, for the central force needed to produce an elliptical orbit when the center is one focus of the ellipse, by determining the potential energy function, rather than the magnitude of the force directly; it bears comparison with Newton's alternate proof in Problem 2-4, and uses an elementary geometry theorem very much like the one in Problem 2-3: for two secants intersecting *inside* a circle, as in the figure on page 137, we have $PA \cdot PB = PC \cdot PD$.



(a) Suppose our ellipse has foci F_1 and F_2, with $F_1P + PF_2 = 2a$ for all points P on the ellipse, and let $F_1F_2 = 2e$. We assume that our central force is directed toward the focus F_1, and we consider a circle of radius $2a$ around the other focus F_2. Extend PF_2 to intersect the circle at A and G, and then extend AF_1 to intersect the circle at D. Show that the three angles indicated by small arcs, involving the tangent line to the ellipse at P, are all equal. Hint: Compare Problem 2-4.



(b) Show that AP and F_1P have the same length r, so that the two segments AB and BF_1 have the same length p, which is the distance from F_1 to the tangent line.

(c) Using the elementary geometry theorem mentioned before, show that $2p \cdot F_1D = (2a - 2e) \cdot (2a + 2e)$, which we can write as

$$p \cdot F_1D = \alpha, \tag{1}$$

for a constant α, while conservation of angular momentum tells us that for the velocity v at P we have

$$p \cdot v = \beta \tag{2}$$

for some constant β [so that F_1D must be proportional to v, although only equations (1) and (2) will be needed].

(d) Using similar triangles, show that

$$\frac{p}{r} = \frac{2p + F_1D}{4a} \quad \Longrightarrow \quad \frac{1}{r} = \frac{1}{2a} + \frac{\alpha v^2}{4a\beta^2},$$

and conclude that $1/r$ is proportional to the potential.

23. This problem outlines a proof that the only central forces for which all orbits are conics are, once again, multiples of either $f(r) = r$ or $f(r) = r^{-2}$, a question that Bertrand later posed. This proof, together with remarks on proofs by Darboux and Halphen, and references to analogous results, appears in the classic book Appell [1; Vol. 1, sect. 232]. Not surprisingly, the proof is a lot more algebraic in nature than the proof of Bertrand's Theorem itself.

(a) A conic can generally be written as the graph of the function y defined by

$$y(x) = \alpha x + \beta + \sqrt{ax^2 + 2bx + c}.$$

Compute y'' (carefully!) and conclude that $y''^{-\frac{2}{3}}$ is a second degree polynomial in x, and thus that y satisfies Halphen's 5$^{\text{th}}$ order equation

$$\text{(H)} \qquad \left(y''^{-\frac{2}{3}}\right)''' = 0.$$

(b) Let $c(t) = (x(t), y(t))$ be a solution to

$$\frac{d^2x}{dt^2} = F\frac{x(t)}{r(t)}, \qquad \frac{d^2y}{dt^2} = F\frac{y(t)}{r(t)} \qquad (r(t) = \sqrt{x(t)^2 + y(t)^2}).$$

Then, by conservation of angular momentum (cf. equation (A_c) on page 83)

$$x\frac{dy}{dt} - y\frac{dx}{dt} = \alpha \qquad \text{for a constant } \alpha.$$

For the reparameterization

$$X = \frac{x}{y}, \qquad Y = \frac{1}{y}, \qquad T = -\frac{t}{y^2},$$

show that

$$\frac{dX}{dT} = \alpha, \qquad \frac{dY}{dT} = \frac{dy}{dt}$$

and thus

$$\frac{d^2X}{dT^2} = 0, \qquad \frac{d^2Y}{dT^2} = \frac{d^2y}{dt^2}\frac{dt}{dT} = -F\frac{y^3}{r}.$$

This means that we can replace our problem about central forces with an equivalent one involving forces parallel to the vertical axis.

So we will be considering a function G such that $x(t)$ and $y(t)$ satisfying

$$\frac{d^2x}{dt^2} = 0, \qquad \frac{d^2y}{dt^2} = G(x(t), y(t))$$

always lie on a conic.
(c) We have

(1) $$\frac{dx}{dt} = \alpha, \qquad \frac{d^2y}{dx^2} = \frac{1}{\alpha^2}G.$$

Hence, if we write

(2) $$G = \mu[\phi(x,y)]^{-\frac{3}{2}}, \qquad G^{-\frac{2}{3}} = \mu^{-\frac{2}{3}}\phi(x,y)$$

for some constant μ, then equation (H) gives $[\phi(x,y)]''' = 0$, where primes now denote derivatives d/dx, so that we have, for example,

$$[\phi(x,y)]' = \frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial y}y'.$$

Compute ϕ'' and ϕ''', use (1) and (2) to obtain

$$y'' = \frac{\mu}{\alpha^2}\phi^{-\frac{3}{2}}, \qquad y''' = -\frac{3}{2}\frac{\mu}{\alpha^2}\phi^{-\frac{5}{2}}\left(\frac{\partial\phi}{\partial x} + y'\frac{\partial\phi}{\partial y}\right),$$

and then reduce the equation $\phi''' = 0$ to

$$\frac{\partial^3\phi}{\partial x^3} + 3y'\frac{\partial^3\phi}{\partial x^2\partial y} + 3y'^2\frac{\partial^3\phi}{\partial x\partial y^2} + y'^3\frac{\partial^3\phi}{\partial y^3}$$
$$+\frac{3\mu}{2\alpha^2}\phi^{-\frac{3}{2}}\left(2\phi\frac{\partial^2\phi}{\partial x\partial y} - \frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y}\right) + \frac{3\mu y'}{2\alpha^2}\phi^{-\frac{3}{2}}\left[2\phi\frac{\partial^2\phi}{\partial y^2} - \left(\frac{\partial\phi}{\partial y}\right)^2\right] = 0.$$

(d) Using the fact that this equation holds for any initial values of x, y, y', α, conclude that we have

(i) $$\frac{\partial^3\phi}{\partial x^3} = 0, \qquad \frac{\partial^3\phi}{\partial x^2\partial y} = 0, \qquad \frac{\partial^3\phi}{\partial x\partial y^2} = 0, \qquad \frac{\partial^3\phi}{\partial y^3} = 0,$$

(ii) $$2\phi\frac{\partial^2\phi}{\partial x\partial y} - \frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y} = 0, \qquad 2\phi\frac{\partial^2\phi}{\partial y^2} - \left(\frac{\partial\phi}{\partial y}\right)^2 = 0.$$

Note that (i) shows that ϕ is a second degree polynomial in x and y,

$$\phi(x,y) = Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F.$$

(e) Use (ii) to show that

$$\phi(x,y) = \begin{cases} \dfrac{1}{C}(Bx + Cy + E)^2, & C \neq 0 \\ Ax^2 + 2Dx + F, & C = 0. \end{cases}$$

In terms of equation (2) we have

$$G = \begin{cases} \dfrac{\mu C^{\frac{3}{2}}}{(Bx + Cy + E)^3}, & C \neq 0 \\ \dfrac{\mu}{(Ax^2 + 2Dx + F)^{\frac{3}{2}}}, & C = 0 \end{cases}$$

and the corresponding central forces are

$$F = \begin{cases} -\dfrac{\mu r C^{\frac{3}{2}}}{(Bx + Ey + C)^3}, & C \neq 0 \\ -\dfrac{\mu r}{(Ax^2 + 2Dxy + Fy^2)^{\frac{3}{2}}}, & C = 0. \end{cases}$$

(f) Assuming (as usual) that the central forces are radially symmetric, this gives us the desired result.

24. (a) In the figure on page 147, the nucleus of the gold atom is at $(a\varepsilon, 0)$, and the periapsis of an α-particle on the dashed path is thus $a + a\varepsilon$. Remembering that our $K = q_1q_2/m$, use formula (b_2') to get

$$a = \frac{h^2m}{(\varepsilon^2 - 1)q_1q_2},$$

and thus

$$\varepsilon = \sqrt{1 + \frac{h^2m}{q_1q_2a}}.$$

(b) Using (b_3'), show that

$$a + a\varepsilon = \frac{q_1q_2}{2E}\left(1 + \sqrt{1 + \frac{2mh^2E}{(q_1q_2)^2}}\right).$$

The smallest periapsis clearly occurs for $h = 0$, and has the value

$$\frac{q_1q_2}{E}.$$

(c) Show that this is the same as choosing $s = 0$ (compare Problem 3-27).

CHAPTER 5

RIGID BODIES

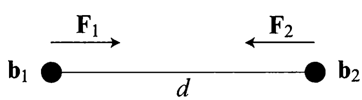
The previous chapter, illustrating the power of Newton's laws in analyzing point masses, or objects that behave in certain respect like point masses, would usually be regarded as material for intermediate or advanced mechanics courses.

But, as we have pointed out numerous times in the previous chapters, many of the "elementary" problems of mechanics do not involve point masses, and instead require the analysis of rigid bodies. Since the study of rigid bodies is also generally regarded as an advanced part of mechanics, elementary mechanics books focus on special cases—often with various unstated assumptions—in order to have some problems to solve.

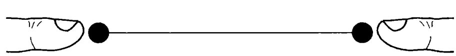
Fortunately, we needn't feel deterred by the use of slightly advanced mathematical notions, so we will be able to examine the basic assumptions that underlie the treatment of rigid bodies without the distraction of various superfluous considerations.

Rigid bodies are obviously idealizations, since in practice nothing is perfectly rigid. (In fact, in special relativity theory rigid bodies are actually impossible even in principle, though we won't be worrying about that here.) Our aim, therefore, is not to produce a "realistic" model of a rigid body, but to define the proper abstract concept that corresponds to it.

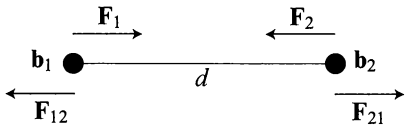
Equilibrium. Before trying to analyze the motion of rigid bodies in general, we first ask when a rigid body should be in equilibrium under certain forces. As the simplest possible example, let's consider a "rigid rod" that consists of just two points $\mathbf{b}_1$ and $\mathbf{b}_2$ (representing two molecules, say) at a distance d apart, and "external" forces $\mathbf{F}_i$ acting on $\mathbf{b}_i$. These forces might be produced, for



example, by some one exerting equal but opposite pressure on both sides of this rod.



If $\mathbf{F}_1 = -\mathbf{F}_2$, then we would expect this rigid rod to be in equilibrium under these forces, and we can justify this expectation by noting that if we consider a force $\mathbf{F}_{21}$ on $\mathbf{b}_2$ equal to $-\mathbf{F}_2$ and a force $\mathbf{F}_{12}$ on $\mathbf{b}_1$ equal to $-\mathbf{F}_1$, then these "internal" forces $\mathbf{F}_{12}$ and $\mathbf{F}_{21}$ do satisfy Newton's third law, and together with the forces $\mathbf{F}_1$ and $\mathbf{F}_2$ they leave our rod, consisting of $\mathbf{b}_1$ and $\mathbf{b}_2$, in equilibrium.



To be sure, as anticipated in Problem 1-4, this picture becomes quite a bit hazier if we try to imagine how these "internal" forces would arise as the forces $\mathbf{F}_i$ are applied. Presumably the internal forces are 0 when the two molecules are at their "natural" distance d apart, but become strongly repulsive if the distance is slightly smaller than d and strongly attractive if the distance is slightly larger than d. So, the forces $\mathbf{F}_i$ initially push the molecules slightly toward each other; as this happens, the molecules produce large repulsive forces, which will not only return the molecules to their original position, but actually cause them to move slightly further apart; this, in turn, will produce large attractive forces, now moving the molecules back toward their "natural" separation, and slightly beyond, causing the repulsive forces to act again. Thus, we would expect the molecules to vibrate around their natural separation, which is more or less what actually happens in our real-world approximation to a rigid rod.

We might hope to describe an ideal rigid rod by considering the limiting situation as the constraining forces of the molecule are made greater and greater. But increasing the constraining forces simply causes the molecules to vibrate more and more rapidly—although they will stay closer and closer to their natural separation, their motions will not approach a limit.

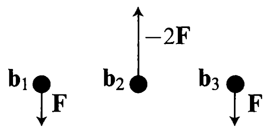
So instead, we will consider our abstract rigid rod to be in equilibrium *simply because such forces* $\mathbf{F}_{ij}$ *can be defined*, without worrying about the details of just how these forces would actually arise in practice, for rods that aren't ideally rigid.

More generally, let us consider a collection of points $\mathbf{b}_1, \dots, \mathbf{b}_K$, which it will sometimes be convenient to regard as a single object, $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_K)$, as well as a collection of forces $\mathbf{F} = (\mathbf{F}_1, \dots \mathbf{F}_K)$, where we regard $\mathbf{F}_i$ as acting on $\mathbf{b}_i$. Then we can make the following definition:

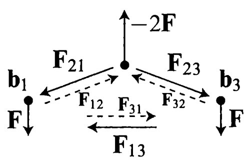
> The collection of points $\mathbf{b}$ is in **rigid equilibrium** under the forces $\mathbf{F}$ if there exist "internal" forces $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ which are multiples of $\mathbf{b}_i - \mathbf{b}_j$ such that
>
> $$\mathbf{F}_i = -\sum_j \mathbf{F}_{ij}.$$
>
> Much more colloquially, of course, we just say that "the rigid body $\mathbf{b}$ is in equilibrium under the forces $\mathbf{F}$".

An important point about this definition is that the masses of the points $\mathbf{b}_i$ do not play a role—though the forces might very well depend on those masses (for example, in a gravitational field); the masses will enter the discussion later on.

Another important point about the definition is that it is inadequate. For example, we presumably ought to have equilibrium for the rod shown below, where there are equal forces $\mathbf{F}$ at the ends of the rod together with a



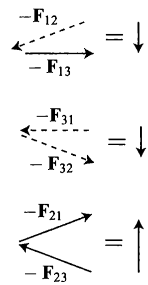
force $-2\mathbf{F}$ in the middle. But these forces obviously can't be balanced by forces that are multiples of the vectors $\mathbf{b}_i - \mathbf{b}_j$. Of course, in practice, the rod will bend a bit, and in this situation the necessary "internal" forces will exist.



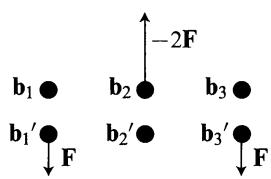
$$\mathbf{F} = -\mathbf{F}_{12} - \mathbf{F}_{13}$$

$$\mathbf{F} = -\mathbf{F}_{31} - \mathbf{F}_{32}$$

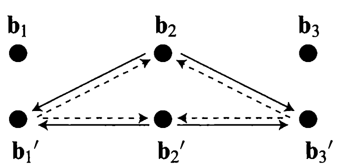
$$-2\mathbf{F} = -\mathbf{F}_{21} - \mathbf{F}_{23}$$



Fortunately, we can stick with our strict theoretical model if we represent the situation by a slightly more realistic figure, with a few extra "molecules", so

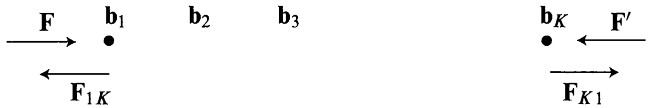


that once again the required internal forces will exist.

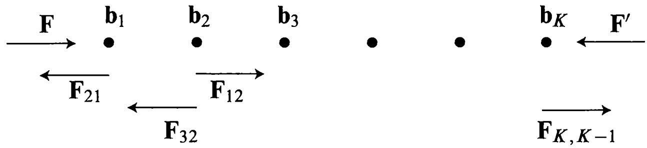


We will normally presume that our particles do not lie on a straight line, or even on a plane, and in realistic situations the number of particles should be much greater, although special cases may be useful for illustration.

It should also be noted that the $\mathbf{F}_{ij}$ of our definition are almost never unique, even for the special case of a "rigid rod" consisting of particles $\mathbf{b}_1, \ldots, \mathbf{b}_K$ lying on a straight line. Given equal and opposite forces $\mathbf{F}$ and $\mathbf{F}'$ on the ends $\mathbf{b}_1$ and $\mathbf{b}_K$, we could choose just two forces $\mathbf{F}_{1K}$ and $\mathbf{F}_{K1}$ between $\mathbf{b}_1$ and $\mathbf{b}_K$,



essentially ignoring all the particles between them, but it would be more natural to balance $\mathbf{F}$ with a force $\mathbf{F}_{21}$ exerted on $\mathbf{b}_1$ by $\mathbf{b}_2$, requiring an equal but



opposite force $\mathbf{F}_{12}$ on $\mathbf{b}_2$, which would in turn be balanced by a force $\mathbf{F}_{32}$ exerted on $\mathbf{b}_2$ by $\mathbf{b}_3$, ... , leading finally to a force $\mathbf{F}_{K,K-1}$ exerted on $\mathbf{b}_K$ by $\mathbf{b}_{K-1}$ that balances $\mathbf{F}'$.

Virtual infinitesimal displacements. Our condition for rigid equilibrium introduces a whole set of unknown forces $\mathbf{F}_{ij}$, but we can obtain a consequence of this condition that does not involve these unknown forces by considering "rigid motions" of $\mathbf{b}$. By this we simply mean a collection of paths $\mathbf{c} = (c_1, \ldots, c_K)$ with $c_i(0) = \mathbf{b}_i$ such that each

$$|c_i(t) - c_j(t)|^2 = \langle c_i(t) - c_j(t),\, c_i(t) - c_j(t)\rangle \quad \text{is constant.}$$

Alternatively, we might think of a rigid motion as a curve $t \mapsto A(t)$ of isometries of $\mathbb{R}^3$, with $c_i(t) = A(t)(c_i(0)) = A(t)(\mathbf{b}_i)$.

Given such a rigid motion, consider the K-tuple of tangent vectors

$$\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_K) = (c_1{}'(0), \ldots, c_K{}'(0)) \in (\mathbb{R}^3)^K.$$

Differentiating the equation

$$\langle c_i(t) - c_j(t),\, c_i(t) - c_j(t)\rangle = \text{constant}$$

and evaluating at 0 gives us

(1) $$\langle \mathbf{v}_i - \mathbf{v}_j,\, \mathbf{b}_i - \mathbf{b}_j\rangle = 0.$$

We have always drawn our forces as if they satisfy the "strong form" of the third law, stated on page 25, and mentioned frequently in Chapter 3, and we will now specifically assume this, leaving further consideration of this assumption to Addendum A. Then, since the force $\mathbf{F}_{ij}$ is a multiple of $\mathbf{b}_i - \mathbf{b}_j$, equation (1) implies that

$$\langle \mathbf{v}_i - \mathbf{v}_j, \mathbf{F}_{ij} \rangle = 0.$$

Consequently,

$$\begin{aligned} \sum_{i,j} \langle \mathbf{v}_i, \mathbf{F}_{ij} \rangle = \sum_{i,j} \langle \mathbf{v}_j, \mathbf{F}_{ij} \rangle &= -\sum_{i,j} \langle \mathbf{v}_j, \mathbf{F}_{ji} \rangle \\ &= -\sum_{i,j} \langle \mathbf{v}_i, \mathbf{F}_{ij} \rangle \quad \text{(interchanging } i \text{ and } j) \end{aligned}$$

and thus

$$\sum_{i,j} \langle \mathbf{v}_i, \mathbf{F}_{ij} \rangle = 0.$$

This in turn means that the external forces $\mathbf{F}_k$ in the condition for rigid equilibrium satisfy

$$\sum\nolimits_k \langle \mathbf{v}_k, \mathbf{F}_k \rangle = -\sum\nolimits_{k,j} \langle \mathbf{v}_k, \mathbf{F}_{kj} \rangle = 0,$$

or simply

$$(*) \qquad \sum_{k} \langle \mathbf{v}_k, \mathbf{F}_k \rangle = 0.$$

Physicists refer to these K-tuples $\mathbf{v} = (c_1{}'(0), \dots, c_K{}'(0))$ for rigid motions $\mathbf{c}$ of $\mathbf{b}$ as "virtual infinitesimal displacements" of $\mathbf{b}$. The word "infinitesimal" in this phrase shouldn't surprise us—it's just the standard physicists' way of referring to tangent vectors. As for the word "virtual" here, it has about as much meaning as it does in the phrase "virtual reality". Basically it refers to the fact that although we have obtained equation $(*)$ under the assumption that our rigid body is in equilibrium, we have done so by considering tangent vectors to "virtual" rigid motions, i.e., motions that our rigid body might have had if it *weren't* in equilibrium.

Configuration space. This can all be expressed in a more familiar, geometric, way by considering the "configuration space" of $\mathbf{b}$, which is the subset $\mathcal{M} \subset (\mathbb{R}^3)^K$ of all points that can be reached from $\mathbf{b}$ at the end of a rigid motion. In other words,

$$\mathcal{M} = \{ (A(\mathbf{b}_1), \dots, A(\mathbf{b}_K)) : A \text{ an orientation preserving isometry of } \mathbb{R}^3 \}.$$

When $\mathbf{b}$ is non-planar, $\mathcal{M}$ is a 6-dimensional manifold diffeomorphic to the set of all orientation preserving isometries A of $\mathbb{R}^3$, and thus to $\mathbb{R}^3 \times \mathrm{SO}(3)$. With this picture, a rigid motion of $\mathbf{b}$ is simply a curve in $\mathcal{M}$, so a virtual infinitesimal displacement $\mathbf{v}$ of $\mathbf{b}$ is simply a tangent vector to $\mathcal{M}$ at $\mathbf{b}$.

We've already found that any such $\mathbf{v}$ satisfies the equation

$$\langle \mathbf{v}_i - \mathbf{v}_j, \mathbf{b}_i - \mathbf{b}_j \rangle = 0. \tag{1}$$

If we define linear functions ϕ_{ij} on $(\mathbb{R}^3)^K$ by

$$\phi_{ij}(\mathbf{v}_1, \dots, \mathbf{v}_K) = \langle \mathbf{v}_i - \mathbf{v}_j, \mathbf{b}_i - \mathbf{b}_j \rangle,$$

this says that

$$\mathcal{M}_{\mathbf{b}} \subset \bigcap_{i,j} \ker \phi_{ij}.$$

1. LEMMA. If $\mathbf{b}$ is non-planar, then

$$\mathcal{M}_{\mathbf{b}} = \bigcap_{i,j} \ker \phi_{ij}.$$

PROOF. By renumbering, we can assume that $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ are points of $\mathbf{b}$ that do not lie in a plane. There is clearly no loss of generality in assuming that $\mathbf{b}_1 = 0$ [as reflected by the fact that we can replace all $\mathbf{b}_i$ by $\mathbf{b}_i - \mathbf{b}_1$ without changing (1)]. Thus our assumption on $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ amounts to $\mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ being linearly independent.

Since we can also replace all $\mathbf{v}_i$ by $\mathbf{v}_i - \mathbf{v}_1$ without changing (1), it follows that

$$\dim\Big(\bigcap_{i,j} \ker \phi_{ij}\Big) = 3 + \dim\Big(\Big\{(0, \mathbf{v}_2, \dots, \mathbf{v}_K) \in \bigcap_{i,j} \ker \phi_{ij}\Big\}\Big).$$

Now for $\mathbf{v}$ with $\mathbf{v}_1 = 0$, a first application of (1) gives

$$\langle \mathbf{v}_i, \mathbf{b}_i \rangle = \langle \mathbf{v}_i - \mathbf{v}_1, \mathbf{b}_i - \mathbf{b}_1 \rangle = 0 \qquad i = 2, 3, 4,$$

and then a second application gives

$$-\langle \mathbf{v}_i, \mathbf{b}_j \rangle = \langle \mathbf{v}_j, \mathbf{b}_i \rangle \qquad i, j = 2, 3, 4.$$

So if $A\colon \mathbb{R}^3 \to \mathbb{R}^3$ is the linear transformation with

$$\mathbf{v}_i = A\mathbf{b}_i \qquad i = 2, 3, 4,$$

then A is skew-adjoint. But the dimension of skew-symmetric 3×3 matrices is 3, so the dimension of $\bigcap_{i,j} \ker \phi_{ij}$ is at most 6, which is the dimension of $\mathcal{M}_{\mathbf{b}}$. ❖

By the way, it shouldn't be too surprising that the mechanics of this proof involved skew-adjoint transformations, since they are the derivatives of orthogonal ones (compare page 186); given the transformation A of the proof, the isometries e^{tA} would produce the given infinitesimal virtual displacement $\mathbf{v}$.

The principle of virtual work. If we use $\langle\ ,\ \rangle$ for the usual inner product on $(\mathbb{R}^3)^K$, then equation $(*)$ on page 179 can be written in the simple form

$$\langle \mathbf{v}, \mathbf{F}\rangle = 0;$$

in other words, $\mathbf{F}$ is perpendicular to the tangent space $\mathcal{M}_{\mathbf{b}}$. As we noted in Chapter 3, the inner product of force and distance is generally called work, so this sum is also called the "(virtual) infinitesimal work" done by the forces $\mathbf{F}$ during the (virtual) infinitesimal displacement $\mathbf{v}$. Our little calculation that $\langle \mathbf{v}, \mathbf{F}\rangle = 0$ if $\mathbf{b}$ is in rigid equilibrium under $\mathbf{F}$ is often referred to by physicists as a proof of the "principle of virtual work". In reality, however, when physicists use the principle of virtual work they almost always assume implicitly that it includes the *converse*:

2. PROPOSITION (THE PRINCIPLE OF VIRTUAL WORK). The non-planar collection of points $\mathbf{b}$ is in rigid equilibrium under the forces $\mathbf{F}$ if and only if the virtual infinitesimal work $\langle \mathbf{v}, \mathbf{F}\rangle$ equals 0 for all virtual infinitesimal displacements $\mathbf{v}$ of $\mathbf{b}$.

PROOF. We have to prove the converse part, that if $\langle \mathbf{v}, \mathbf{F}\rangle = 0$ for all $\mathbf{v}$, then $\mathbf{b}$ is in rigid equilibrium under $\mathbf{F}$. If we consider the linear function Φ on $(\mathbb{R}^3)^K$ defined by

$$\Phi(\mathbf{v}_1, \dots, \mathbf{v}_K) = \sum_k \langle \mathbf{v}_k, \mathbf{F}_k\rangle,$$

then our hypothesis says that Φ vanishes on $\mathcal{M}_{\mathbf{b}}$, and thus by our lemma,

$$\Phi \text{ vanishes on } \bigcap_{i,j} \ker \phi_{ij}, \qquad \phi_{ij}(\mathbf{v}_1, \dots, \mathbf{v}_K) = \langle \mathbf{v}_i - \mathbf{v}_j,\ \mathbf{b}_i - \mathbf{b}_j\rangle.$$

A simple result about vector spaces (see Problem 1 for a refresher) then states that there exist constants λ_{ij} with

$$\Phi = \sum_{i,j} \lambda_{ij} \cdot \phi_{ij}.$$

In other words,

$$\sum_k \langle \mathbf{v}_k, \mathbf{F}_k\rangle = \sum_{i,j} \lambda_{ij} \langle \mathbf{v}_i - \mathbf{v}_j, \mathbf{b}_i - \mathbf{b}_j\rangle \qquad \text{all } \mathbf{v}_1, \dots, \mathbf{v}_K \in (\mathbb{R}^3)^K.$$

Choosing all $\mathbf{v}_i$ to be 0 except for the one vector $\mathbf{v}_l$, we thus obtain

$$\begin{aligned}\langle \mathbf{v}_l, \mathbf{F}_l\rangle &= \sum_j \lambda_{lj}\langle \mathbf{v}_l, \mathbf{b}_l - \mathbf{b}_j\rangle + \sum_i \lambda_{il}\langle -\mathbf{v}_l, \mathbf{b}_i - \mathbf{b}_l\rangle \\ &= \sum_j \lambda_{lj}\langle \mathbf{v}_l, \mathbf{b}_l - \mathbf{b}_j\rangle + \sum_j \lambda_{jl}\langle \mathbf{v}_l, \mathbf{b}_l - \mathbf{b}_j\rangle \\ &= \sum_j (\lambda_{lj} + \lambda_{jl})\langle \mathbf{v}_l, \mathbf{b}_l - \mathbf{b}_j\rangle \\ &= \Big\langle \mathbf{v}_l,\ \sum_j (\lambda_{lj} + \lambda_{jl})(\mathbf{b}_l - \mathbf{b}_j)\Big\rangle,\end{aligned}$$

and since this is true for arbitrary vectors $\mathbf{v}_l$ in $\mathbb{R}^3$, we conclude that

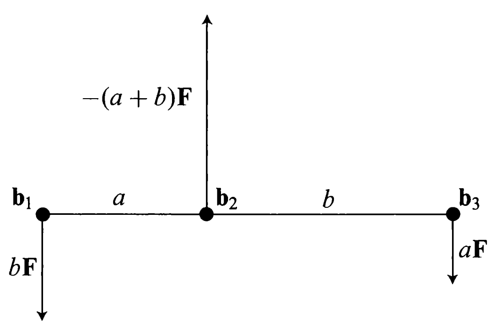
$$\mathbf{F}_l = \sum_j (\lambda_{lj} + \lambda_{jl})(\mathbf{b}_l - \mathbf{b}_j).$$

So we can define

$$\mathbf{F}_{jl} = -(\lambda_{lj} + \lambda_{jl})(\mathbf{b}_l - \mathbf{b}_j)$$

to obtain the required forces. ❖

As an extremely simple example, consider the situation shown below, where the upward force on $\mathbf{b}_2$ balances the two downward forces at points $\mathbf{b}_1$ and $\mathbf{b}_3$, which are at different distances a and b from $\mathbf{b}_2$, with the magnitudes of these forces inversely proportional to those distances. Of course, this is merely a



schematic figure, since it is linear, and we really have to assume that there are other points around, as in our previous examples.

It is easy to see that this collection of points is in rigid equilibrium under these forces:

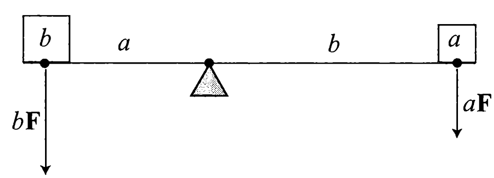
(1) For an infinitesimal displacement given by a vector $\mathbf{z}$ pointing in the vertical direction, the virtual infinitesimal work is 0 because the upward force is the negative of the sum of the two downward forces.

(2) For an infinitesimal displacement given by a vector $\mathbf{z}$ pointing in the horizontal direction, the virtual infinitesimal work is 0 because each individual component is 0.

(3) For an infinitesimal displacement generated by a rotation around $\mathbf{b}_2$ (i.e., around an axis through $\mathbf{b}_2$ perpendicular to the plane of the diagram), the vectors $\mathbf{v}_1$ and $\mathbf{v}_3$ will be in opposite vertical directions, with lengths proportional to the distances a and b. Consequently, the virtual infinitesimal work, involving vectors with length *inversely* proportional to these distances, will be 0. (One can check directly that this is just as true for an infinitesimal displacement generated by a rotation around either $\mathbf{b}_1$ or $\mathbf{b}_3$, but that isn't necessary, since the set of virtual infinitesimal displacements that stay in the plane of the diagram has dimension 3.)

(4) For infinitesimal displacements given by a vector perpendicular to the plane of the figure, or by a rotation through axes perpendicular to our first rotation, the virtual infinitesimal work also works out to be 0; or we can just simplify matters by restricting our attention to the 2-dimensional situation to begin with.

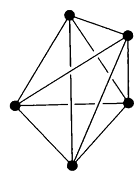
Notice that this provides a fairly good schematic representation of a lever, which of course requires not only a rigid body, but also a *fulcrum*, an immovable point. In practice, this "immobility" is provided in a complicated way by the



connections between the fulcrum and the earth, but it seems reasonable simply to regard this connection as a mechanism that automatically supplies the proper *upward* force to the fulcrum when the downward forces are applied at the ends of the lever.

Naturally, a more realistic picture would use a much large number of points, forming a 3-dimensional object. But in any case, our analysis shows, especially when we think of the lever as bending slightly, as on page 177, that it is the internal forces of the lever that make the weights balance; in short, all the "extra force" that one obtains by pushing at a large distance from the fulcrum is supplied by the lever itself, in its effort to preserve rigidity (together with the force that the earth supplies on the fulcrum, to keep it from moving downward).

Of course, for a truly realistic picture, we would need further information to determine how the forces within the lever actually arise: the internal forces guaranteed by the principle of virtual work certainly won't be unique, any more than they were in the unrealistic case mentioned on page 178. In fact, once we have 4 points $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ not on a plane, all ϕ_{ij} can be expressed as linear



combinations of the ϕ_{ij} for $i, j = 1, \dots 4$, which means that the λ_{ij} in our proof are by no means unique, and thus the $\mathbf{F}_{ij}$ aren't either.

d'Alembert's principle. Although we have so far investigated only a rigid body in equilibrium, our analysis easily extends to the more general situation.

Instead of looking for a condition for equilibrium, we now seek a criterion for a rigid motion $\mathbf{c} = (c_1, \dots, c_K)$ of $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_K)$ to be consistent with the forces $\mathbf{F} = (\mathbf{F}_1, \dots \mathbf{F}_K)$. The $\mathbf{F}_i$ should be considered as functions on $\mathcal{M} \times \mathbb{R}$, to encompass the general situation mentioned at the beginning of Chapter 3, where the forces may depend not only on time, but also on the particular rigid motion that the body has undergone at any particular time.

We now need "internal" forces $\mathbf{F}_{ij}(t) = -\mathbf{F}_{ji}(t)$ with $\mathbf{F}_{ij}(t)$ a multiple of $c_i(t) - c_j(t)$ so that

$$m_i c_i''(t) = \mathbf{F}_i(\mathbf{c}(t), t) + \textstyle\sum_j \mathbf{F}_{ij}(t)$$

or

$$m_i c_i''(t) - \mathbf{F}_i(\mathbf{c}(t), t) = \textstyle\sum_j \mathbf{F}_{ij}(t).$$

The latter equation, which may be regarded as stating that the body is in rigid equilibrium under the forces $\mathbf{F}_i - m_i c_i''$, is often called "d'Alembert's principle" and regarded as the fundamental law—so that, as the physicists like to say, "dynamics reduces to statics". But this really becomes useful only when we apply the principle of virtual work: The conditions on the $\mathbf{F}_{ij}$ imply that

$$\textstyle\sum_i \langle \mathbf{v}_i, m_i c_i''(t) - \mathbf{F}_i(\mathbf{c}(t), t)\rangle = 0$$

for all tangent vectors $\mathbf{v}$, and, conversely, the principle of virtual work implies that if this condition holds, then the requisite $\mathbf{F}_{ij}(t)$ exist. This leads us to the following definition:

> **"d'Alembert's Principle":** The rigid motion $\mathbf{c}$ is a **rigid solution** for the forces $\mathbf{F}$, or, more colloquially, "$\mathbf{c}$ is a possible motion of the rigid body $\mathbf{b}$ under the forces $\mathbf{F}$", if for each t,
>
> $$\sum_i \langle \mathbf{F}_i(\mathbf{c}(t),t) - m_i c_i''(t),\ \mathbf{v}_i\rangle = 0$$
>
> for all tangent vectors $\mathbf{v}$ at $\mathcal{M}_{\mathbf{c}(t)}$.

If we agree to let $m\mathbf{c}$ denote $(m_1c_1,\dots,m_Kc_K)$ and similarly for $m\mathbf{c}''$, and also let $\langle\ ,\ \rangle$ denote the usual inner product on $(\mathbb{R}^3)^K$, then we can write

$$\langle \mathbf{F}(\mathbf{c}(t),t) - m\mathbf{c}''(t),\ \mathbf{v}\rangle = 0 \qquad \text{for all } \mathbf{v} \text{ tangent to } \mathcal{M}_{\mathbf{c}(t)},$$

or, if we are willing to tolerate a little ambiguity in our notation, simply

$$(**) \qquad \langle \mathbf{F} - m\mathbf{c}'',\ \mathbf{v}\rangle = 0 \qquad \text{for all } \mathbf{v} \text{ tangent to } \mathcal{M}_{\mathbf{c}}.$$

Our condition amounts to a system of second order differential equations for vector-valued functions: If we choose a local coordinate system $x^1,\dots,x^6$ on the 6-dimensional manifold $\mathcal{M}$, then we only have to verify $(**)$ for $\mathbf{v} = \partial/\partial x^i$, giving us 6 equations for the vector-valued functions $\mathbf{c}^i = x^i \circ \mathbf{c}$.

Since $\mathcal{M}$ is basically $\mathbb{R}^3 \times \mathrm{SO}(3)$, we can restate this much more concretely. One 3-dimensional collection of vector fields tangent to $\mathcal{M}$ are those of the form $\mathbf{v}_i = \mathbf{z}$ for a constant vector $\mathbf{z}$. Condition $(**)$ becomes

$$\begin{aligned} 0 &= \sum_i \langle \mathbf{F}_i - m_i c_i'', \mathbf{z}\rangle \\ &= \Big\langle \textstyle\sum_i \mathbf{F}_i, \mathbf{z}\Big\rangle - \Big\langle \textstyle\sum_i m_i c_i'', \mathbf{z}\Big\rangle. \end{aligned}$$

Since this must hold for all $\mathbf{z}$, we must have

$$(\mathbf{F}_{\text{rigid}}) \qquad \mathbf{F}_{\text{total}} = \sum_i \mathbf{F}_i = \sum_i m_i c_i'' = MC'',$$

where C is the center of mass, and $M = \sum_i m_i$ is the total mass [recall that $\mathbf{F}_i$ really stands for $t \mapsto \mathbf{F}_i(\mathbf{c}(t),t)$].

We also have to consider the vector fields generated by rotations. As we saw in our discussion of the cross-product, these are of the form $\mathbf{v}_i = c_i \times \boldsymbol{\eta}$. Thus, condition $(**)$ becomes

$$
\begin{aligned}
0 &= \sum_i \langle \mathbf{F}_i - m_i c_i{}'', c_i \times \boldsymbol{\eta} \rangle \\
&= \sum_i \langle \mathbf{F}_i, c_i \times \boldsymbol{\eta} \rangle - \sum_i m_i \langle c_i{}'', c_i \times \boldsymbol{\eta} \rangle \\
&= \sum_i \langle c_i \times \mathbf{F}_i, \boldsymbol{\eta} \rangle - \sum_i m_i \langle c_i \times c_i{}'', \boldsymbol{\eta} \rangle .
\end{aligned}
$$

Since this must hold for all $\boldsymbol{\eta}$, we must have, recalling (L′) on page 83,

$$
(\boldsymbol{\tau}_{\text{rigid}}) \qquad \boldsymbol{\tau} = \sum_i c_i \times \mathbf{F}_i = \sum_i m_i c_i \times c_i{}'' = \mathbf{L}'.
$$

Condition $(\mathbf{F}_{\text{rigid}})$ simply says that the rigid body must move in such a way that the momentum law is satisfied, while condition $(\boldsymbol{\tau}_{\text{rigid}})$ simply says that the rigid body must move in such a way that the angular momentum law is satisfied.

The inertia tensor. To solve these equations we might begin by writing our rigid motion $\mathbf{c} = (c_1, \dots, c_K)$ of $\mathbf{b}$ in the form

$$
c_i(t) = B(t)(\mathbf{b}_i) + \mathbf{w}(t)
$$

for orthogonal $B(t)$. Since $BB^{\mathbf{t}} = I$, we have

$$
\begin{aligned}
B'B^{\mathbf{t}} &= -BB^{\mathbf{t}\prime} \\
&= -(B'B^{\mathbf{t}})^{\mathbf{t}},
\end{aligned}
$$

so $B'B^{-1}(t) = B'B^{\mathbf{t}}(t)$ is skew-adjoint, and its matrix can be written as

$$
\begin{pmatrix}
0 & -\omega_3(t) & \omega_2(t) \\
\omega_3(t) & 0 & -\omega_1(t) \\
-\omega_2(t) & \omega_1(t) & 0
\end{pmatrix}.
$$

Setting $\boldsymbol{\omega}(t) = \big(\omega_1(t), \omega_2(t), \omega_3(t)\big)$, we then have

$$
\begin{aligned}
\mathbf{v}_i(t) = c_i{}'(t) &= B'(t)(\mathbf{b}_i) + \mathbf{w}'(t) \\
&= \big(B'(t)B^{-1}(t)\big)\big(B(t)(\mathbf{b}_i)\big) + \mathbf{w}'(t) \\
&= (B'(t)B^{-1}(t))(c_i(t)) + \mathbf{w}'(t) \\
&= [\boldsymbol{\omega}(t) \times c_i(t)] + \mathbf{w}'(t).
\end{aligned}
$$

There is a natural choice for $B(t)$ and $\mathbf{w}(t)$ in our description of rigid body motion: choose $\mathbf{w}(t) = C(t)$, where $C(t)$ is the center of mass at time t, so that $B(t)$ represents the rotation about the center of mass from its initial position to its position at time t. We then have

$$\mathbf{v}_i(t) = [\boldsymbol{\omega}(t) \times c_i(t)] + C'.$$

We can now write the angular momentum $\mathbf{L}$ of $\mathbf{c}$ as

$$\begin{aligned}
\sum_i m_i c_i \times \mathbf{v}_i &= \sum_i m_i c_i \big[(\boldsymbol{\omega} \times c_i) + C'\big] \\
&= \sum_i \big[m_i c_i \times (\boldsymbol{\omega} \times c_i)\big] + \Big(\sum_i m_i c_i\Big) \times C' \\
&= \sum_i \big[m_i c_i \times (\boldsymbol{\omega} \times c_i)\big] + \big(MC \times C'\big), \qquad M = \textstyle\sum_i m_i.
\end{aligned}$$

Comparing this with the formula on page 84, we see that the quantity

$$\sum_i m_i c_i \times (\boldsymbol{\omega} \times c_i)$$

is the same as the "rotational angular momentum", that is, the angular momentum of $\mathbf{c}$ around its center of mass.

In Chapter 9 we look at specific examples of the equations ($\mathbf{F}_{\text{rigid}}$) and ($\boldsymbol{\tau}_{\text{rigid}}$), but for now we only want to consider some basic aspects of the general problem. To simplify matters, we can ignore the motion of the center of mass, and just look for a solution of the form $c_i(t) = B(t)(\mathbf{b}_i)$, essentially describing how the body rotates about the center of mass; the more general case just involves writing longer equations, without changing the main point we want to make. In fact, we will simply assume that some point in our body, not necessarily the center of mass, is fixed, and consider it to be the origin of our coordinate system.

Thus, we are looking for $\boldsymbol{\omega}$ so that ($\boldsymbol{\tau}_{\text{rigid}}$) holds when we have

$$c_i{}' = \mathbf{v}_i = \boldsymbol{\omega} \times c_i.$$

We have

$$\begin{aligned}
c_i{}'' &= (\boldsymbol{\omega}' \times c_i) + (\boldsymbol{\omega} \times \mathbf{c}_i{}') \\
&= (\boldsymbol{\omega}' \times c_i) + (\boldsymbol{\omega} \times (\boldsymbol{\omega} \times c_i)),
\end{aligned}$$

so equation ($\boldsymbol{\tau}_{\text{rigid}}$) becomes

$$\boldsymbol{\tau} = \sum_i m_i c_i \times (\boldsymbol{\omega}' \times c_i) + \sum_i m_i c_i \times (\boldsymbol{\omega} \times (\boldsymbol{\omega} \times c_i)).$$

In terms of the linear function $\mathbf{I}_c : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\mathbf{I}_c(\eta) = \sum_i m_i c_i \times (\eta \times c_i), \qquad \text{with } \mathbf{I}_c(\boldsymbol{\omega}) = \mathbf{L},$$

we can write this as

$$(*) \qquad \mathbf{I}_c(\boldsymbol{\omega}') = \boldsymbol{\tau} - \sum_i m_i c_i \times (\boldsymbol{\omega} \times (\boldsymbol{\omega} \times c_i)),$$

where the right side depends on $c \in \mathcal{M}$ and $\boldsymbol{\omega}$. The only thing we need to check is that we can always solve this for some $\boldsymbol{\omega}'$ as a function of c and $\boldsymbol{\omega}$, thereby obtaining a system of first order equations for $\boldsymbol{\omega}'$, and thus a system of second order equations for the elements of B. In other words, we need to know that the linear transformation $\mathbf{I}_c$ is an isomorphism for all $c \in \mathcal{M}$.

Since we only have to consider $c \in \mathcal{M}$ of the form $c_i = P(\mathbf{b}_i)$ for some orthogonal P, we have

$$\begin{aligned}\mathbf{I}_c(\eta) &= \sum_i m_i P(\mathbf{b}_i) \times (\eta \times P(\mathbf{b}_i)) \\ &= P\Big(\textstyle\sum_i m_i \mathbf{b}_i \times (P^{-1}(\eta) \times \mathbf{b}_i)\Big) \\ &= (P\mathbf{I}_{\mathbf{b}}P^{-1})(\eta).\end{aligned}$$

Thus we only have to check that $\mathbf{I} = \mathbf{I}_{\mathbf{b}}$ is an isomorphism, where

$$\mathbf{I}(\boldsymbol{\phi}) = \sum_i m_i \mathbf{b}_i \times (\boldsymbol{\phi} \times \mathbf{b}_i) \qquad \text{for } \boldsymbol{\phi} \in \mathbb{R}^3.$$

Now for any $\boldsymbol{\psi} \in \mathbb{R}^3$ we have

$$\begin{aligned}\langle \mathbf{I}(\boldsymbol{\phi}), \boldsymbol{\psi}\rangle &= \sum_i \langle m_i \mathbf{b}_i \times (\boldsymbol{\phi} \times \mathbf{b}_i), \boldsymbol{\psi}\rangle \\ &= \sum_i m_i \langle \boldsymbol{\phi} \times \mathbf{b}_i, \boldsymbol{\psi} \times \mathbf{b}_i\rangle \\ &= \sum_i m_i \langle \boldsymbol{\psi} \times \mathbf{b}_i, \boldsymbol{\phi} \times \mathbf{b}_i\rangle \\ &= \langle \mathbf{I}(\boldsymbol{\psi}), \boldsymbol{\phi}\rangle.\end{aligned}$$

Thus $\mathbf{I}$ is self-adjoint, and consequently has an orthonormal basis of eigenvectors. Since

$$\langle \mathbf{I}(\boldsymbol{\phi}), \boldsymbol{\phi}\rangle = \sum_i m_i |\boldsymbol{\phi} \times \mathbf{b}_i|^2,$$

the corresponding eigenvalues are all ≥ 0, and in fact they are all > 0 because we are assuming that $\mathbf{b}$ is non-planar, and thus at least one $|\boldsymbol{\phi} \times \mathbf{b}_i| > 0$. Since $\mathbf{I}$ always has positive eigenvalues, it is always an isomorphism, so we can indeed always solve equation $(*)$ for $\boldsymbol{\omega}'$.

The map $\mathbf{I} = \mathbf{I_b}$, satisfying

$$\text{(I)} \qquad \mathbf{I}(\omega) = \mathbf{L},$$

is called the **inertia tensor** of $\mathbf{b}$ (with respect to the fixed point). The directions of its eigenvalues are called the **principal axes of inertia**, the corresponding eigenvalues are called the **principal moments of inertia**, and the **inertia ellipsoid** about the fixed point is the surface consisting of all vectors ψ with $\langle \mathbf{I}(\psi), \psi \rangle = 1$ (so the semimajor axes of the inertia ellipsoid are the reciprocals of the square roots of the principal moments of inertia). As our rigid body moves under the rotations $B(t)$, the inertia tensor for $B(t)(\mathbf{b})$ is just the composition $B(t) \circ \mathbf{I} \circ B^{-1}(t)$; the principal moments of inertia remain the same for all positions of the rigid body under the motion, and the whole inertia ellipsoid, including the principal axes of inertia, are transformed by the $B(t)$.

Aside from the values of $\mathbf{F}_{\text{total}}$ and $\boldsymbol{\tau}$, the principal moments of inertia are the only other data entering into our equations, so, in a sense, the whole motion of the rigid body $\mathbf{b}$ depends only on them. In particular, for motion under no external forces, we obtain exactly the same equations for two rigid bodies of arbitrary shape, provided only that they have the same principle moments of inertia.

Calculating the inertia tensor. To write down the matrix of $\mathbf{I}$ with respect to the standard basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ we temporarily adopt the notation $\mathbf{b}_i = (x_i, y_i, z_i)$. Using the identity[1]

$$\mathbf{w} \times (\mathbf{u} \times \mathbf{v}) = \langle \mathbf{w}, \mathbf{v} \rangle \mathbf{u} - \langle \mathbf{w}, \mathbf{u} \rangle \mathbf{v}$$

we can write

$$\mathbf{I}(\omega) = \sum_i m_i (|\mathbf{b}_i|^2 \omega - \langle \mathbf{b}_i, \omega \rangle \mathbf{b}_i),$$

which gives

$$\begin{aligned}
\mathbf{I}(\mathbf{e}_1) &= \sum_i m_i \big(|\mathbf{b}_i|^2 \mathbf{e}_1 - \langle \mathbf{b}_i, \mathbf{e}_1 \rangle \mathbf{b}_i\big) \\
&= \sum_i m_i \big(|\mathbf{b}_i|^2 \mathbf{e}_1 - x_i (x_i, y_i, z_i)\big) \\
&= \sum_i m_i \big(y_i{}^2 + z_i{}^2, -x_i y_i, -x_i z_i\big),
\end{aligned}$$

with similar results for $\mathbf{e}_2$ and $\mathbf{e}_3$. Thus, the matrix of $\mathbf{I}$ with respect to the

[1] Proof: Since $\mathbf{w} \times (\mathbf{u} \times \mathbf{v})$ is perpendicular to $\mathbf{u} \times \mathbf{v}$, it is a linear combination of $\mathbf{u}$ and $\mathbf{v}$, and the appropriate coefficients are easy to determine using the usual identities for $\times$.

standard basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is

$$\mathcal{I} = \begin{pmatrix} \sum_i m_i(y_i{}^2 + z_i{}^2) & -\sum_i m_i x_i y_i & -\sum_i m_i x_i z_i \\ -\sum_i m_i y_i x_i & \sum_i m_i(x_i{}^2 + z_i{}^2) & -\sum_i m_i y_i z_i \\ -\sum_i m_i z_i x_i & -\sum_i m_i z_i y_i & \sum_i m_i(x_i{}^2 + y_i{}^2) \end{pmatrix}.$$

The same result obviously holds for any orthonormal basis $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ if we set $x_i = \langle \mathbf{b}_i, \mathbf{v}_1 \rangle$, $y_i = \langle \mathbf{b}_i, \mathbf{v}_2 \rangle$, and $z_i = \langle \mathbf{b}_i, \mathbf{v}_3 \rangle$.

The diagonal terms of the matrix of $\mathbf{I}$ with respect to an orthonormal basis are the quantities that were classically called the "moments of inertia" of $\mathbf{b}$ about the axes; the off-diagonal terms are sometimes called the "products of inertia". In other words, the moment of inertia I_A of $\mathbf{b}$ about an axis A is

$$I_A = \sum_i m_i r_i{}^2,$$

where r_i is the distance from $\mathbf{b}_i$ to A. If our orthonormal coordinate system happens to point along the principal axes, then these moments of inertia are the principal moments of inertia.

Note that if the three principal axes all have the same moment of inertia I, (see Problems 6 and 8 for examples), then all axes are principal axes, and we have $\mathbf{L} = \mathbf{I}(\boldsymbol{\omega}) = I\boldsymbol{\omega}$. In this case, the equation $\boldsymbol{\tau} = \mathbf{L}'$ simply becomes

$$(\boldsymbol{\tau}_{\text{symmetric}}) \qquad \boldsymbol{\tau} = I\boldsymbol{\omega}',$$

in exact analogy with $\mathbf{F} = m\mathbf{v}'$.

As we might expect, for a rigid body $\mathbf{b}$ a special role is played by the moments of inertia about the axes that pass through the center of mass $C = (x_C, y_C, z_C)$. More generally, there is a simple relationship between the matrix $\mathcal{I}$ and the matrix $\mathcal{I}'$ of the inertia tensor of $\mathbf{b}$ in a parallel coordinate system whose origin is C. Let

$$x_i = \bar{x}_i + x_C, \qquad y_i = \bar{y}_i + y_C, \qquad z_i = \bar{z}_i + z_C,$$

so that $(\bar{x}_i, \bar{y}_i, \bar{z}_i)$ are the coordinates of the rigid body in this system, and let $M = \sum_i m_i$. When we write $\mathcal{I}$ in terms of the $\bar{x}_i, \bar{y}_i, \bar{z}_i$, any cross term like $\sum_i m_i \bar{y}_i y_C$ vanishes, since

$$\sum_i m_i \bar{y}_i = \sum_i m_i(y_i - y_C) = \sum_i m_i y_i - M y_C,$$

which is 0 by definition of y_C. Thus, we obtain simply

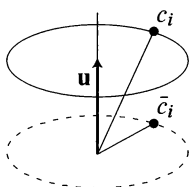
$$
\begin{aligned}
\mathcal{I} &= \mathcal{I}' + M\cdot\begin{pmatrix} {y_C}^2 + {z_C}^2 & -x_C y_C & -x_C z_C \\ -y_C x_C & {x_C}^2 + {z_C}^2 & -y_C z_C \\ -z_C x_C & -z_C y_C & {x_C}^2 + {y_C}^2 \end{pmatrix} \\
&= \mathcal{I}' + \mathcal{I}_C,
\end{aligned}
$$

where $\mathcal{I}_C$ is the matrix of the inertia tensor of the single body C with mass M around the origin of our original coordinate system.

In particular, we have

3. PROPOSITION (THE PARALLEL AXIS THEOREM OR STEINER'S THEOREM). If the point P is at distance d from the center of mass C of $\mathbf{b}$, the moment of inertia of $\mathbf{b}$ about any axis through P is Md^2 plus the moment of inertia about the parallel axis through C.

Rotation about an axis. Moments of inertia play an important role when we consider a rigid body whose motion is a rotation about an axis. For a unit



vector $\mathbf{u}$ pointing along this axis, decompose each c_i as

$$c_i = \bar{c}_i + \langle c_i, \mathbf{u}\rangle \mathbf{u} \tag{1}$$

where $\bar{c}_i$ is in the plane perpendicular to $\mathbf{u}$, so that $|\bar{c}_i|$ is the distance r_i from c_i to the axis. The tangent vector $\bar{c}_i{}'$ is in the same plane as $\bar{c}_i$ and perpendicular to it, and if $\theta(t)$ is the angle through which the body has rotated at time t, then the length of $\bar{c}_i{}'$ is $r_i\,\theta'$. So

$$\bar{c}_i \times \bar{c}_i{}' = {r_i}^2\theta' \cdot \mathbf{u},$$

and it follows easily that

$$\langle c_i \times c_i{}', \mathbf{u}\rangle = \langle \bar{c}_i \times \bar{c}_i{}', \mathbf{u}\rangle = {r_i}^2\theta',$$

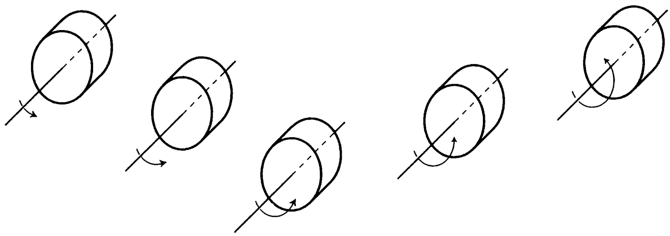
and thus

$$\langle c_i \times c_i{}'', \mathbf{u}\rangle = \langle c_i \times c_i{}', \mathbf{u}\rangle' = {r_i}^2\theta''.$$

So equation ($\boldsymbol{\tau}_{\text{rigid}}$) gives the $\mathbf{u}$ component of $\boldsymbol{\tau}$ as

$$\langle \boldsymbol{\tau}, \mathbf{u}\rangle = I_A \cdot \theta''. \tag{$\boldsymbol{\tau}_{\text{axis}}$}$$

Equation ($\boldsymbol{\tau}_{\text{axis}}$) holds even in the more general case where the motion of a rigid body is the result of combining rotation about an axis A with a motion of



this axis parallel to itself. In fact, this just changes the c_i to

$$c_i + \alpha\mathbf{u} + \mathbf{v},$$

for some functions α and $\mathbf{v}$, with $\mathbf{v}$ always perpendicular to $\mathbf{u}$. It is then easy to check that

$$\langle (c_i + \alpha\mathbf{u} + \mathbf{v}) \times (c_i{}'' + \alpha''\mathbf{u} + \beta''\mathbf{v}), \mathbf{u}\rangle = \langle (c_i \times c_i{}''), \mathbf{u}\rangle,$$

because each of the other terms in the expansion is 0.

If a more "physically intuitive" argument is preferred, we can use the same sort of reasoning that we used in analyzing a rocket on page 33: At any particular time t_0 we work in the inertial system that is moving with the same velocity as the axis at time t_0 to derive our equation, which involves only the second derivative $\theta''(t_0)$ of θ at time t_0, and consequently also holds in our inertial system.

If there are no external forces on our rotating body, then we have $\theta'' = 0$, so that the body rotates with constant angular velocity $\theta' = a$. Since we now have

$$c_i{}' = a\mathbf{u} \times c_i,$$

we get

$$\begin{aligned} c_i{}'' &= a\mathbf{u} \times c_i{}' \\ &= a^2\mathbf{u} \times (\mathbf{u} \times c_i) \\ &= a^2[\langle \mathbf{u}, c_i\rangle\mathbf{u} - c_i], \end{aligned}$$

so that

$$c_i \times c_i{}'' = -a^2\langle \mathbf{u}, c_i\rangle(\mathbf{u} \times c_i).$$

Since we are assuming there are no external forces, equation ($\boldsymbol{\tau}_{\text{rigid}}$) then gives

$$0 = \sum_i m_i \langle c_i, \mathbf{u}\rangle(\mathbf{u} \times c_i).$$

Since

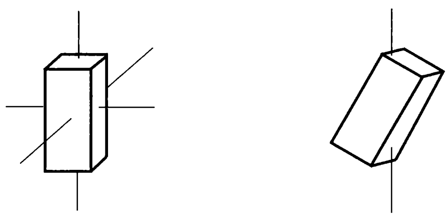
$$\begin{aligned} \mathbf{I}(\mathbf{u}) &= \sum_i m_i c_i \times (\mathbf{u} \times c_i) \\ &= \sum_i m_i |c_i|^2 \mathbf{u} - \sum_i m_i \langle c_i, \mathbf{u}\rangle c_i, \end{aligned}$$

this shows that

$$\mathbf{u} \times \mathbf{I}(\mathbf{u}) = 0,$$

so $\mathbf{I}(\mathbf{u})$ is a multiple of $\mathbf{u}$, and $\mathbf{u}$ must be an eigenvector of $\mathbf{I}$. In the general case, a rigid body has just three axes around which it can rotate without external forces, and then the angular velocity must be constant.

The block shown below can rotate about the three axes of symmetry. You



might naively expect that if it were provided with the right initial push it could also rotate about any other axis that passes through the center of mass of the block, as illustrated in the right hand part of the figure, but a little thought should be able to convince you otherwise (note that the angular momentum vector won't be constant).

Kinetic energy. Although the inertia tensor arose naturally in our investigation of the equations for rigid body motion, physics books usually introduce it in a completely different way, involving an expression for the total kinetic energy T of a rigid body. Using the equations for $\mathbf{v}_i$ on page 186, with $\mathbf{w}'$ the velocity of the center of mass, we have

$$\mathbf{v}_i = \boldsymbol{\omega} \times c_i + \mathbf{w}',$$

so, letting $M = \sum_i m_i$ be the total mass of the rigid body, we obtain

$$T = \tfrac{1}{2}\sum_i m_i |\mathbf{v}_i|^2 = \tfrac{1}{2} M |\mathbf{w}'|^2 + \tfrac{1}{2}\sum_i m_i |\boldsymbol{\omega} \times c_i|^2 + \sum_i m_i \langle \mathbf{w}', \boldsymbol{\omega} \times c_i\rangle.$$

The third term can be written as

$$\left\langle \mathbf{w}', \boldsymbol{\omega} \times \textstyle\sum_i m_i c_i \right\rangle = \langle \mathbf{w}', \boldsymbol{\omega} \times M \cdot C \rangle,$$

where C is the center of mass. So if we choose a (usually non-inertial) coordinate system with the center of mass as the origin, this term vanishes, and we obtain

$$\begin{aligned} T &= \tfrac{1}{2} M |\mathbf{w}'|^2 + \tfrac{1}{2} \sum_i m_i |\boldsymbol{\omega} \times c_i|^2 \\ &= \tfrac{1}{2} M |\mathbf{w}'|^2 + \tfrac{1}{2} \sum_i \langle \mathbf{I}_{c_i}(\boldsymbol{\omega}), \boldsymbol{\omega} \rangle, \end{aligned}$$

breaking up the total kinetic energy into a "translational" part and a "rotational" part:

$$T = T_{\text{transl}} + T_{\text{rot}}.$$

Since $\mathbf{w}'$ is simply the velocity of the center of mass in our original inertial system, T_{transl} is just the usual kinetic energy of the center of mass, while the term T_{rot} is the extra kinetic energy due to rotation about the center of mass, or simply the kinetic energy when the center of mass if fixed. So

$$\text{(T)} \qquad 2T_{\text{rot}} = \langle \mathbf{I}(\boldsymbol{\omega}), \boldsymbol{\omega} \rangle = \big\langle (\omega_1, \omega_2, \omega_3) \cdot \mathfrak{I},\ (\omega_1, \omega_2, \omega_3) \big\rangle.$$

Once we've introduced kinetic energy, it is naturally tempting to consider what happens when our rigid body is moving in a conservative force field. Recall that for a particle c moving in a force field $\mathbf{F}$ with

$$\mathbf{F} = -\left(\frac{\partial V}{\partial x^1}, \frac{\partial V}{\partial x^2}, \frac{\partial V}{\partial x^3} \right)$$

we have

$$\begin{aligned} \tfrac{1}{2} m |\mathbf{v}(t_1)|^2 - \tfrac{1}{2} m |\mathbf{v}(t_0)|^2 &= \int_{t_0}^{t_1} \langle \mathbf{v}(t), \mathbf{F}(c(t)) \rangle \, dt \\ &= -V(c(t_1)) + V(c(t_0)). \end{aligned}$$

So if $\langle\ ,\ \rangle$ denotes, as before, the usual inner product on $(\mathbb{R}^3)^K$, and $\mathbf{c} = (c_1, \ldots, c_K)$, then adding the above equations for the various c_i gives, for the total kinetic energy T, the expression

$$\text{(C)} \qquad \begin{aligned} T(t_1) - T(t_0) &= \int_{t_0}^{t_1} \langle \mathbf{v}(t), \mathbf{F}(\mathbf{c}(t)) \rangle \, dt \\ &= -\big(\textstyle\sum_i V(c_i(t_1))\big) + \big(\textstyle\sum_i V(c_i(t_0))\big). \end{aligned}$$

This equation holds for particles $c_1, \ldots, c_K$ moving independently, but what we are interested in is the motion of $\mathbf{c}$ as a rigid body, which we think of as the motions of the c_i under the forces

$$\bar{\mathbf{F}}_i = \mathbf{F}_i + \sum_j \mathbf{F}_{ij}$$

for suitable "internal" forces $\mathbf{F}_{ij}$. Then we have

$$T(t_1) - T(t_0) = \int_{t_0}^{t_1} \langle \mathbf{v}(t), \bar{\mathbf{F}}(\mathbf{c}(t))\rangle\, dt,$$

where

$$\langle \mathbf{v}(t), \bar{\mathbf{F}}_i(\mathbf{c}(t))\rangle = \langle \mathbf{v}(t), \mathbf{F}_i(\mathbf{c}(t))\rangle + \sum_j \langle \mathbf{v}(t), \mathbf{F}_{ij}(\mathbf{c}(t))\rangle.$$

But the (easy direction of) the principle of virtual work says that

$$\sum_j \langle \mathbf{v}(t), \mathbf{F}_{ij}(\mathbf{c}(t))\rangle = 0;$$

as the physicists like to say, the total work done by the internal forces of a rigid body is always 0. Consequently, equation (C) holds even for the motion of $\mathbf{c}$ as a rigid body:

$$T(t_1) - T(t_0) = -\big(\textstyle\sum_i V(c_i(t_1))\big) + \big(\textstyle\sum_i V(c_i(t_0))\big).$$

For a rigid body moving under the influence of gravity near the earth's surface, each $V(c_i(t))$ is just $m_i h_i(t)$ where $h_i(t)$ is the height of c_i at time t, and it is easy to see that if $h_C(t)$ is the height of the center of mass of the rigid body, and M is its total mass, then

$$\textstyle\sum_i V(c_i(t)) = \big(\textstyle\sum_i m_i\big) h_C(t) = M \cdot h_C(t),$$

where $h_C(t)$ is the height of the center of mass of the rigid body, and M is its total mass. So we obtain, finally,

$$T(t_1) - T(t_0) = M \cdot [h_C(t_0) - h_C(t_1)].$$

Continuous bodies. Mathematically, it is straightforward to generalize the previous considerations to a continuous rigid body B with density ρ: The total mass M is given by

$$M = \int_B \rho = \int_B \rho(x, y, z)\, dx\, dy\, dz,$$

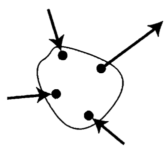
using x, y, z for the standard coordinate functions on $\mathbb{R}^3$, and the center of mass is the vector given by

$$C = \frac{1}{M}\int_B \rho(x,y,z)(x,y,z)\,dx\,dy\,dz,$$

i.e., C is the point of $\mathbb{R}^3$ with coordinates $1/M$ times

$$\int_B x\cdot\rho(x,y,z)\,dx\,dy\,dz, \quad \int_B y\cdot\rho(x,y,z)\,dx\,dy\,dz, \quad \int_B z\cdot\rho(x,y,z)\,dx\,dy\,dz.$$

However, there is one point that might be emphasized to avoid confusion. We often consider forces that are supposed to be acting on a single point of a rigid body. If we think in terms of a continuous body, we then have a finite force



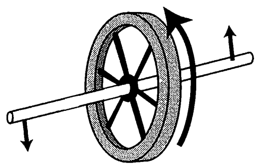
acting on a single point, which ought to have mass 0. It helps to think back to our view of a rigid body as a finite collection of particles, where our analysis shows that the total effect of all the internal forces should be that each force may be considered as acting on the center of mass. For continuous bodies this may be translated into the "principle" that a force exerted anywhere will have the same effect as one exerted at the center of mass.

Continuing with our consideration of continuous bodies, we set the inertia tensor $\mathbf{I}$ of B to be the linear transformation whose matrix $\mathcal{I}$ with respect to the standard basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ is given by [here ρ denotes $\rho(x,y,z)$]

$$\begin{pmatrix} \int_B \rho\cdot(y^2+z^2)\,dx\,dy\,dz & -\int_B \rho\cdot xy\ dx\,dy\,dz & -\int_B \rho\cdot xz\ dx\,dy\,dz \\ -\int_B \rho\cdot yx\ dx\,dy\,dz & \int_B \rho\cdot(x^2+z^2)\,dx\,dy\,dz & -\int_B \rho\cdot yz\ dx\,dy\,dz \\ -\int_B \rho\cdot zx\ dx\,dy\,dz & -\int_B \rho\cdot zy\ dx\,dy\,dz & \int_B \rho\cdot(x^2+y^2)\,dx\,dy\,dz \end{pmatrix}$$

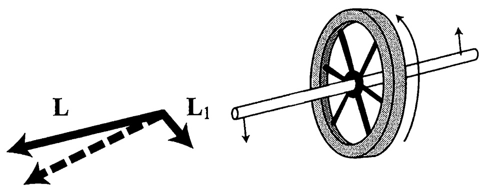
Once we have the inertia tensor, it is irrelevant whether we are considering a continuous body or a discrete set of points: everything becomes a set of equations for the $\boldsymbol{\omega}$. The rotational kinetic energy is given by equation (T), and the diagonal terms of $\mathcal{I}$ are naturally what we call the moments of inertia for continuous bodies; it is left as an easy exercise for the reader to check that the parallel axis theorem still holds.

Elementary examples. Some aspects of rigid body motion can be analyzed without actually solving any equations, simply by using the angular momentum law. In Chapter 3 we mentioned the standard elementary illustration of the special case of conservation of angular momentum, involving some one seated on a rotating stool. An illustration of the more general angular momentum law is provided when the person on the stool holds the ends of an axle with a heavy wheel rapidly rotating on it, and tries to turn the axle, either clockwise or counter clockwise; the rather non-intuitive result is that the axle is fairly hard to



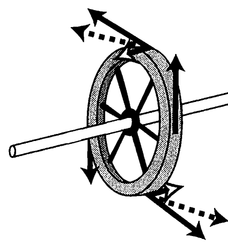
turn, but the effort causes the stool to start spinning. Unlike the first example, in which spin cannot be obtained starting from rest—though a non-zero initial spin can be modified—this second example truly involves a rigid body. To be sure, we have both a rotating wheel and a stationary axle, but that is merely a convenience, and one could imagine a single object consisting of the wheel and the axle rigidly connected and rotating together (though this would be awfully hard on the hands of the person trying to hold it).

Without examining the detailed motion of the spinning wheel, we can see that it has a large angular momentum **L** parallel to the axle, while the twisting



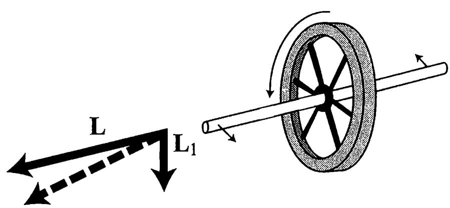
motion of the person on the stool adds a small angular momentum $\mathbf{L}_1$ in the horizontal plane. The resultant, dashed, arrow is still in the same horizontal plane, but rotated from the axis; hence the axis of the rotating wheel needs to rotate in order for the wheel to have this new angular momentum.

It's also nice to know that we can give an "elementary" argument, by considering the velocity vectors at various points of the wheel, together with the additional velocities added by the twisting motion—indicated by solid arrows with white heads—which are oppositely directed at the top and bottom of the



wheel, and 0 at the sides. The dashed arrows show the resultant velocities, and obviously the axis of the wheel must rotate in order for it to have these velocities.

The figure below shows the same situation except that now the forces on the axle are exerted in the horizontal plane, $\mathbf{L}_1$ points downward, and the resultant dashed arrow is in a vertical plane, pointing downward. If this is the view of



a bicycle rider (riding toward the plane of the picture), turning the wheels by means of the handle-bars to which it is attached, the rider soon intuitively learns to lean to the left as the bicycle is thus steered to the left.

Admittedly, there's something a little fishy about all these descriptions. How can the person sitting on the stool suddenly acquire a horizontal rotation from a force directed upwards? How can the completely horizontal force used to turn a bicycle to the left end up producing a motion of the bicycle in the perpendicular plane (a question that many a first-time bicycler may well wonder, though perhaps not in those terms). Obviously, the motion must be somewhat more complicated than what we have described.

We defer this question to Chapter 9, which gives a more detailed investigation of the equations for rigid body motion; though the situation is fairly complicated, it involves the straightforward part of our subject, mathematics, and we have not yet finished dealing with the tricky part, elementary physics.

ADDENDUM 5A

THE STRONG FORM OF THE THIRD LAW

In Chapter 1 we noted that the third law is often accepted rather uncritically, presumably under some mistaken application of the notion of symmetry. In the case of the strong form of the third law, which we've made the basis for our analysis of rigid bodies, the symmetry argument seems more relevant: if we assume that the laws of physics don't distinguish any direction from any other, what other direction could the force between two bodies have except the line between them? Still, we might feel impelled to inquire whether the strong form of the third law is consistent with experiment. And the answer to this question turns out to be No, although it's a fairly complicated No.

In Problem 1-24, we introduced the Lorentz force law

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B})$$

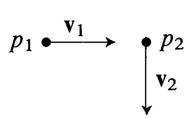
for a moving particle with charge q and velocity $\mathbf{v}$ in a magnetic field $\mathbf{B}$. If our magnetic field is produced by a magnet, rather than a solenoid, we would seem to have a force between the moving charged particle and the magnet that doesn't lie along the line between them. Of course, a magnet isn't a particle (unless some one actually discovers a magnetic monopole!), but this example might still give us pause.

We should begin by noting that, on the face of it, the Lorentz force law $\mathbf{F} = q(\mathbf{v} \times \mathbf{B})$ can't be true, since it would mean that $\mathbf{F}$ would turn out to be different for an observer in another inertial system where the velocity $\mathbf{v}$ of the particle is different! In fact, the proper statement of the law is

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

where E is the "electric field"; the apparent disparity between the results for different inertial systems is accounted for by the fact that a moving magnetic field produces an electric field (as in an electrical generator), and moving charged particles produce a magnetic field (as in an electromagnet).

In particular, two moving charged particles each produce a magnetic field, which affects the other particle according to the Lorentz law, and this can have very strange consequences, as shown for the two charged particles in the figure below. The magnetic field produced by a moving charged particle is 0 along



its line of motion, so particle p_2 is affected only by the charge on p_1. But p_1 not only has the equal but opposite force because of the charge of p_2, but also has an additional force from the magnetic field produced by p_2. Thus the total force on p_1 is not only not in the same direction as the force on p_2, it is actually bigger! We won't try to say much more about this here, because the complete analysis of this situation involves both the principles of electromagnetism and of relativity theory. We should mention, however, that the symmetry argument we were tempted to use on page 25 breaks down precisely because the directions of the velocities of the particles destroys the symmetry of the situation.

Of course, we hope that such effects, involving interactions between the electrons and protons that make up our rigid body, will average out to zero, and when we analyze rigid bodies in terms of point masses we are obviously thinking more in terms of molecules, which presumably ought to be good representatives of point masses without these added weird features. But that is an only a vague approximation, or possibly just a vague hope, not to mention that even at the level of molecules we really should be describing things in terms of quantum mechanics. So our treatment of rigid bodies certainly involves a great deal of idealization and simplification (gracefully finessing, along the way, the whole question of just how one distinguishes a solid from other forms of matter in terms of molecules).

There is a school of thought that dispenses with all these problems, essentially retaining the original mental picture of matter as continuous, totally abjuring the strong form of the third law, and simply regarding equation ($\boldsymbol{\tau}_{\text{rigid}}$) as a basic fact about rigid bodies, verified by experiment (gracefully finessing, along the way, the question of what constitutes a rigid body). The only slight problem with this approach is that the answer to the question raised in our Prologue, just why the lever works as it does, then becomes simply "Because it does".

Take your pick. *De gustibus non est disputandum.*

A lively, not to say cantankerous, discussion of this question may be found in Truesdell [1], in the essay "V. Whence the Law of Moment of Momentum?", which may be regarded as an equal and opposite force to the viewpoint adopted in this chapter; this essay also discusses how the general notion of conservation of angular momentum gradually evolved, as briefly alluded to on page 85, from Newton's corollary to his Proposition about areas swept out under a central force.

PROBLEMS

1. Let $f: V \to \mathbb{R}$ be a linear function on a (possibly infinite-dimensional) vector space V.

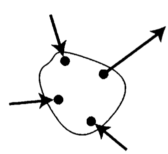
(a) If $v_1, v_2 \in V$, then $f(v_1)v_2 - f(v_2)v_1 \in \ker f$.
(b) We can write $V = \ker f \oplus W$, where W is 1-dimensional.
(c) If $g: V \to \mathbb{R}$ is a linear function with $\ker f \subset \ker g$, then $g = \lambda f$ for some $\lambda \in \mathbb{R}$.
(d) More generally, if $g, f_1, \dots, f_k: V \to \mathbb{R}$ and $\bigcap_i \ker f_i \subset \ker g$, then $g = \sum_i \lambda_i f_i$ for some $\lambda_i \in \mathbb{R}$.

2. Let the Jacobian of $f: \mathbb{R}^n \to \mathbb{R}^k$ have rank k on $f^{-1}(0)$, so that $M = f^{-1}(0)$ is a submanifold of $\mathbb{R}^n$ of dimension $n - k$. Let $g: M \to \mathbb{R}$ be differentiable, and suppose that g has a maximum at $p \in M$.

(a) $M_p = \bigcap_{i=1}^k \ker df^i$, where $df^i: \mathbb{R}^n{}_p \to \mathbb{R}$.
(b) If $X_p \in M_p$ then $dg(X_p) = 0$. *Hint*: $X_p = c'(0)$ for some curve c in M.
(c) Use Problem 1 to conclude that there are $\lambda_1, \dots, \lambda_k$ with

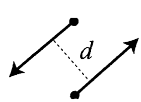
$$D_j g = \sum_{i=1}^k \lambda_i D_j f^i \qquad \text{for } j = 1, \dots, n.$$

3. Consider a rigid body of uniform density, and a finite number of forces at various points of the body. Older books in mechanics consider when two such



sets of forces are "equivalent" in the sense that they have the same effect on the rigid body. This amounts to saying that their sums must be equal, and their total torques about a given point must be equal.

(a) A *couple* is a pair of forces $\mathbf{P}$ and $\mathbf{P}' = -\mathbf{P}$ at two different points. Show that any set of forces is equivalent to a set consisting of a single force $\mathbf{F}$, together with one couple $(\mathbf{P}, -\mathbf{P})$.



(b) Suppose we are considering only a 2-dimensional situation, so that all our forces lie in a plane. Show that the torque of a couple depends only on the distance d between the lines of the forces $\mathbf{P}$ and $\mathbf{P}'$, and that we can consequently choose $\mathbf{P}$ to be collinear with $\mathbf{F}$. Conclude that any set of forces in the 2-dimensional situation is equivalent either to a single force, or to a couple.
(c) In the 3-dimensional situation, consider the system of forces consisting of $\mathbf{F}$ at the point $\mathbf{b}_1$, and the couple consisting of the force $\mathbf{P}$ at the point $\mathbf{b}_2$, and $\mathbf{P} = -\mathbf{P}'$ at the point $\mathbf{b}_3$, and let $\boldsymbol{\tau}$ be its torque around O. Let $\boldsymbol{\tau}'$ be the torque around O of the new system consisting of the same forces, but at the points $\mathbf{b}_i + \mathbf{b}_0$. Show that

$$\boldsymbol{\tau}' = \boldsymbol{\tau} - (\mathbf{b}_0 \times \mathbf{F}).$$

(Recall Corollary 4 of Chapter 3.)

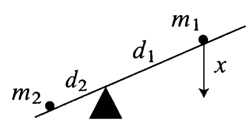
A *wrench* consists of a force $\mathbf{F}$ together with a couple $(\mathbf{P}, -\mathbf{P})$ with $\mathbf{P}$ a multiple of $\mathbf{F}$. If $\mathbf{F} \neq 0$, we can try to reduce our system to a wrench by choosing $\mathbf{b}_0$ so that

$$\boldsymbol{\tau} - \mathbf{b}_0 \times \mathbf{F} = \lambda \mathbf{F}$$

for some λ. Take the inner product with $\mathbf{F}$ to find a formula for λ, and then conclude that any set of forces is equivalent to either a single force, a couple, or a wrench.

4. (a) Consider two particles of masses m_1 and m_2 attached to a lever of negligible weight, with their centers of mass at distances d_1 and d_2 from the fulcrum, and let $x(t)$ be the distance that the first particle falls after time t. Use d'Alembert's Principle to show that

$$x'' = g \cdot \frac{d_1 m_1 - d_2 m_2}{d_1 m_1 + d_2 m_2}.$$



(b) Suppose we start with the lever making an angle of θ with the horizontal, and let T be the time it takes for the lever to reach a horizontal position, so that the first particle has fallen the distance $h = d_1 \sin\theta$. We are going to consider d_2 to be fixed, while d_1 can be varied, by moving the first particle along the lever. Letting $D = d_1/d_2$, show that T satisfies

$$T^2 = \text{constant} \cdot D \cdot \frac{Dm_1 + m_2}{Dm_1 - m_2} = \text{constant} \cdot \frac{D^2 m_1 + Dm_2}{Dm_1 - m_2},$$

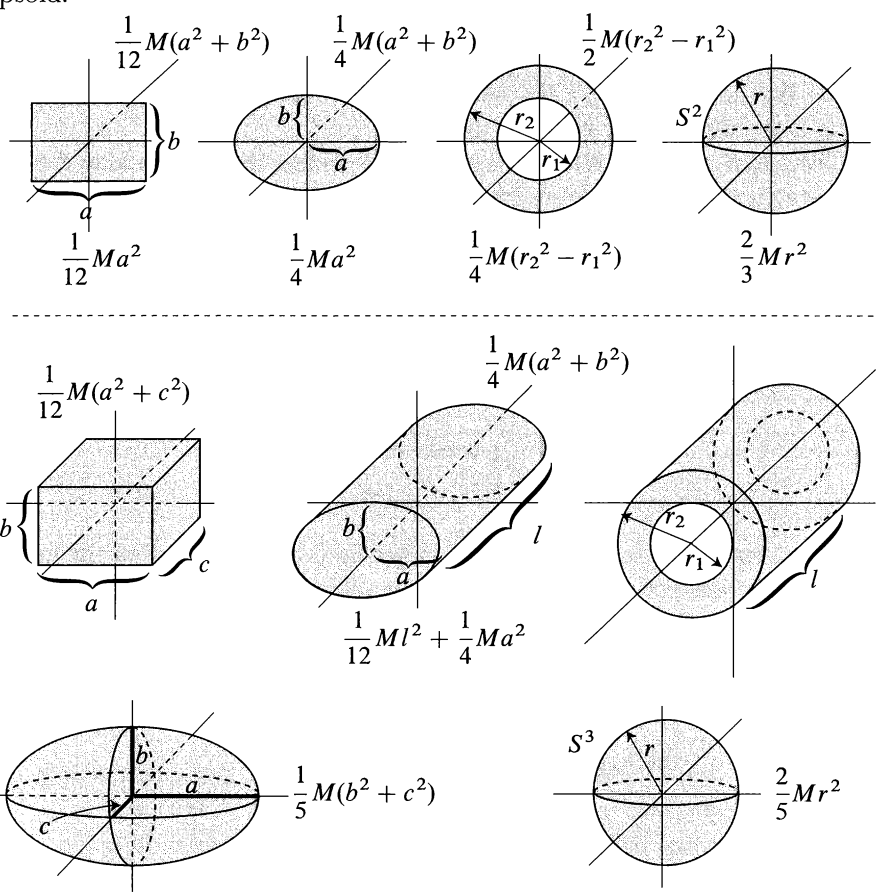
and conclude that the lever will reach a horizontal position in the smallest amount of time when we choose

$$D = \frac{m_2}{m_1} \frac{1 + \sqrt{5}}{2}.$$

5. (a) Prove the "perpendicular axis theorem": For a 2-dimensional figure in the (x, y)-plane, with planar density ρ (a plane "lamina"), the moment of inertia I_z around the z-axis is the sum $I_x + I_y$ of the moments of inertia around the x- and y-axes. (Naturally this applies to any set of orthogonal axes two of which lie in the plane of the lamina.)
(b) For a 3-dimensional object B with density ρ, we have

$$I_x + I_y + I_z = 2 \int_B \rho(x, y, z) r(x, y, z)^2 \, dx \, dy \, dz.$$

225 6. Check or find the moments of inertia about all three axes for the objects shown below, each centered around the origin; all bodies are supposed to be homogeneous, with mass M. The first row contains 2-dimensional objects, which are considered to have planar densities, as in Problem 5, while the rest are 3-dimensional. Note that the inertia ellipsoid of an ellipsoid is a quite different ellipsoid.



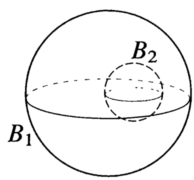
7. Consider a homogeneous disc, regarded as a 2-dimensional object with planar density, as in the previous problem, which at time t has rotated about the axis through its center by the angle $\theta(t)$. Show that its rotational kinetic energy $T_{\text{rot}} = \frac{1}{2}I(\theta')^2$, where I is its moment of inertia about that axis.

8. (a) If a rigid body is symmetric with respect to a plane P, then the direction perpendicular to P is a principal direction, with the other two directions lying in P. (Here the hypothesis means that the reflection R through P takes any point of the rigid body to another point of the rigid body with the same density at the two points.)
(b) Any homogeneous body in the form of a platonic solid or an Archimedean solid has at least three principal directions, with the same principal moment of inertia, and hence all axes through the center of mass are principal directions.
(c) If a rigid body has rotational symmetry around a line L, then L is a principal direction, with the other two directions lying in the plane perpendicular to L.

9. The derivative of the rotational kinetic energy satisfies ${T_{\text{rot}}}' = \langle \boldsymbol{\omega}, \boldsymbol{\tau} \rangle$.

10. (a) For any rigid body, the sum of any two of the principal moments of inertia is greater than or equal to the third.
(b) Given $\alpha, \beta, \gamma > 0$ with the sum of any two greater than or equal to the third, there is a rigid body having α, β, and γ as its principal moments of inertia.

11. Let B_2 be a ball completely contained within the ball B_1, and let B be the difference, $B = B_1 \smallsetminus (\text{interior } B_2)$. Find the principle moments of inertia of B.



12. Consider a rigid body $\mathbf{b}$ in a *uniform* gravitational field, so that the force on particle $\mathbf{b}_i$ of mass m_i is $m_i\mathbf{u}$ for some unit vector $\mathbf{u}$.

(a) As the rigid body $\mathbf{b}$ falls (not necessarily in the direction $\mathbf{u}$, since it may have an initial velocity and rotation), the center of mass C moves the same as if it were a single particle moving under this force. Thus, the center of mass is the "center of gravity".
(b) Assuming that C is actually one of the points in $\mathbf{b}$, show that $\mathbf{b}$ is in equilibrium under the force of gravity together with a force $-M\mathbf{u}$ on C.
(c) If a gravitational field is not uniform, then there are rigid bodies that do not have a "center of gravity" is the sense of part (b).

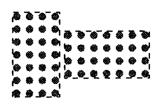
CHAPTER 6

CONSTRAINTS

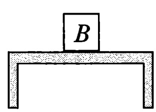
The analysis of rigid bodies represents only the first step in dealing with elementary physics problems, because those problems seldom involve an isolated rigid body in space. We usually have to consider rigid bodies resting on a table, or the floor, or hanging from the ceiling; or we have rigid bodies interacting with each other, colliding or sliding along one another; or our rigid body is restricted by certain "constraints", which, as we will see, have played a role even in the few simple systems that we have already encountered.

Rigid bodies in contact. In some situations we just have to use "common sense", and we might as well get such considerations out of the way in this section, which will consist of various observations, rather than any systematic presentation.

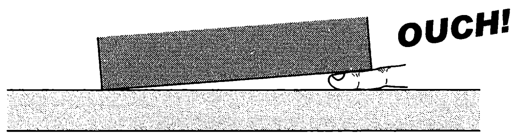
The simple notion of one body in contact with another has no meaning in terms of our theoretical picture: ideal rigid bodies cannot be in contact, but



only very close to each other (just like real ones). If we have a rigid body B of mass M resting on a table, which we regard as yet another rigid body, but with essentially infinite mass, because the rigid table is itself resting on the earth,



then we might as well resort to the usual elementary analysis: The table must exert an upward force on B equal in magnitude to the downward force gM that gravity exerts on B, since B isn't moving. And that of course means that B must be exerting the force gM on the table; it's easy to perceive this force directly by sticking one's finger between B and the table.

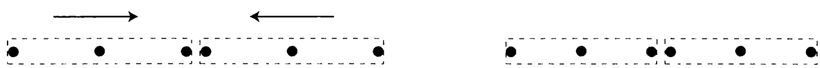


We might ask, instead, what upward forces are exerted on the table top by the table legs to balance the downward force of B. For a 2-dimensional situation, thinking of the table top as a long plank supported by two legs, as in (a), the

(a) (b)

forces are easily determined from the conditions that the total force and total torque on the table top are equal to 0. But if there are three legs, as in (b), there will be more than one possible solution; for example, we could simply ignore the middle leg. Such "statically indeterminate" problems require some considerations of the way that "solid" bodies actually bend. The main ideas involved in such investigations are presented in Addendum B. Although such interesting problems might seem like obvious topics for mechanics courses, nowadays they seem to have been relegated to courses in "applied mechanics" or "mechanical engineering".

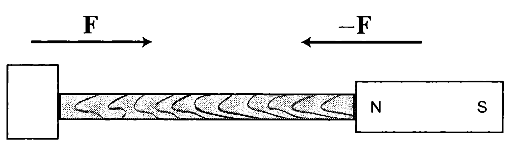
Similarly, the analysis of the previous chapter tells us nothing directly about the collision of two rigid bodies (a); the only reason that realistic, nearly rigid,



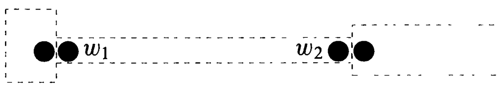
(a) (b)

bodies rebound is because they do compress a bit (b). It would seem reasonable, however, to regard our theoretical rigid body as "perfectly elastic" in the sense discussed in Chapter 3—the forces between the compressed particles simply restore them to their initial positions, increasing the velocity by exactly the same amount that it has been diminished. In that case, we can simply use conservation of kinetic energy to predict the result. Addendum A considers more complex extensions of such reasoning.

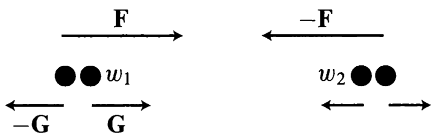
For a final example, involving yet another interesting complication, we consider three bodies in contact: a piece of iron, a long piece of wood, and a magnet. The magnet exerts a force **F** on the iron, and the iron exerts the force $-$**F** on the magnet. We generally presume, probably without even formulating the thought, that the force **F** on the iron is "passed through" to the wood, i.e., that the iron must be exerting a force **F** on the wood, just as in the case of a rigid body resting on a table.



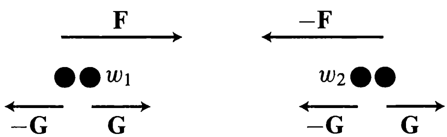
To analyze this situation, the piece of wood can be thought of as rigid body consisting of two particles, w_1 and w_2. The piece of iron can, for the purposes of this problem, simply be considered as a rigid body with only a single particle, and the same is true of the magnet, since we only care about the fact that the magnet exerts a force on the iron (we're not going to start worrying about whether magnetic monopoles exist or not!).



Now let's consider the various forces involved. In addition to the forces **F** and −**F** that the magnet and the iron exert on each other, the iron exerts some force **G** on the particle w_1, and thus w_1 must exert the force −**G** on the iron.



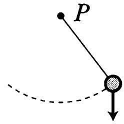
Similarly, w_2 and the magnet must be exerting equal but opposite forces on each other. It seems natural to assume that these forces are −**G** and **G** (after



all, why should the magnet interact with the wood any differently than the iron does?). This means, according to our criterion, that the rigid wooden rod is in equilibrium. Of course, that requires that the iron and magnet also have velocity 0, since they can't pass through the wood! Since the total force on the iron is $\mathbf{F} - \mathbf{G}$, this means that we have $\mathbf{F} = \mathbf{G}$, as we would indeed quite unconsciously presume.

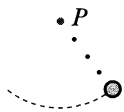
Notice that there seems to be no purely logical way to rule out the possibility that the forces between w_2 and the magnet have values different from **G** and −**G**. If that were the case, however, then our whole apparatus would be continually accelerating—easily solving the energy crisis, among other things—a phenomena that we don't actually encounter in the real world. (These remarks may be compared with those on page 277 of Chapter 7.)

The pendulum. To begin our more systematic investigation of constraints, we consider the pendulum, already mentioned several times in previous chapters, where a string anchored to a pivot point P supposedly constrains a bob—which we will simply regard as point mass—to move along a circular arc. In reality, a pendulum bob can't actually move in a perfectly circular arc (even if we could



have a pivot point P that was really totally immobile). When we initially release the pendulum bob, it starts to fall straight down, and the pendulum string only exerts sufficient force to counteract this force of gravity after it's been stretched somewhat. Then that new force will bring the bob a bit above the circular arc, so that it starts to fall again, etc., etc., etc.

In other words, the seemingly simple example of a pendulum is really an abstraction, presenting the same sorts of problem as a rigid body. In one respect it is simpler, however. We can regard the string as a collection of particles lying linearly along the string. The bob might be affected by a force on one or more

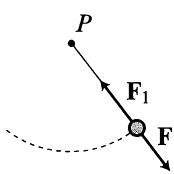


of these particles, as on page 178, but in any case, the net effect is simply a force on the bob in the direction of the string. Thus, besides the force of gravity on the bob, the only other force is a "constraint force" always directed along the line from the bob to the pivot P.

This justifies our previous treatments of the pendulum, but it is also interesting to analyze the pendulum in a manner analogous to our treatment of rigid bodies; to simplify matters we restrict ourselves to a 2-dimensional picture right from the start.

Step 1. Analogous to the condition that a set of points is in rigid equilibrium under various forces, we may say that our pendulum bob, the single particle $\mathbf{a}$, is in equilibrium under the force $\mathbf{F}$, together with the constraint force of the string, keeping it at a fixed distance from the pivot, if

$$\mathbf{F} = -\mathbf{F}_1$$



for some ("internal") force $\mathbf{F}_1$ *in the direction of* $\mathbf{a}$ *to* P.

Step 2. We define a "configuration space" $\mathcal{M}$ for our problem. We now have a single particle $\mathbf{a}$, rather than a set of particles $\mathbf{a}_1, \ldots, \mathbf{a}_K$, and instead of the constraints that the $|\mathbf{a}_i - \mathbf{a}_j|$ remain constant, we have the constraint that $\mathbf{a}$ remain at a fixed distance l from the pivot. So $\mathcal{M} \subset \mathbb{R}^2$ is simply the circle of radius l about the pivot point, and the "virtual infinitesimal displacements", the tangent vectors to possible motions of $\mathbf{a}$ under the constraints, lie along the tangent line of the circle at $\mathbf{a}$.

Step 3. According to Step 1, a force $\mathbf{F}$ causes $\mathbf{a}$ to be in equilibrium under the constraints if and only if it points from $\mathbf{a}$ to P. This is the same as saying that it is perpendicular to the tangent line,

$$\langle \mathbf{F}, \mathbf{v} \rangle = 0 \qquad \mathbf{v} \in \mathcal{M}_{\mathbf{a}}$$

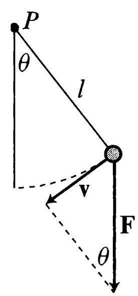
(the principal of virtual work).

Step 4. We conclude that the path c of $\mathbf{a}$ should satisfy

$$(*) \qquad \langle \mathbf{F}(c(t)) - mc''(t), \mathbf{v} \rangle = 0 \qquad \mathbf{v} \in \mathcal{M}_{c(t)}$$

(d'Alembert's principal).

Step 5. To obtain a differential equation, we want to take a coordinate system on $\mathcal{M}$. The natural choice is the angle θ that the line through a point of $\mathcal{M}$ and the pivot P makes with the vertical position. With the standard abuse of



notation, we will let $\theta(t)$ denote $\theta(c(t))$, the θ coordinate of our particle at time t. In $(*)$, we take $\mathbf{v} = \partial/\partial\theta$. If l is the length of the string, then

$$\langle c''(t), \mathbf{v} \rangle = l\theta''(t).$$

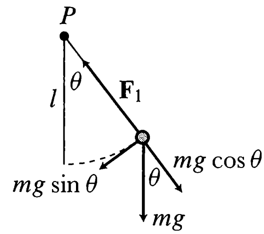
If we choose $\mathbf{F}$ to be the constant downward force of gravity, with magnitude g, then

$$\langle \mathbf{F}(c(t)), \mathbf{v} \rangle = mg \sin\theta(t).$$

So $(*)$ gives us the pendulum equation

$$(\mathrm{P}) \qquad \theta'' + \frac{g}{l}\sin\theta = 0.$$

This equation doesn't involve the internal force $\mathbf{F}_1$ (tension), which is directed along the radius and keeps the bob on the circular arc. Indeed, the whole point of this particular analysis was to eliminate $\mathbf{F}_1$ from consideration—it is simply whatever is necessary to keep the particle on the configuration space $\mathcal{M}$. If we need to know this force, we note from the figure that the component of



the gravitational force along the direction of the string has magnitude $mg \cos\theta$. Avoiding a trap for the unwary, we note that the tension is not simply the negative of this component, because the radial acceleration of the pendulum bob is not zero, but $l\theta'^2$ (Problem 1-5), so $\mathbf{F}_1$ has magnitude $mg\cos\theta + ml\theta'^2$ (we will be able to handle this question differently in Problem 12-3).

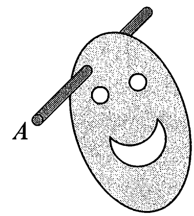
This very special case illustrates a general method for approaching all such "constraint" problems. For rigid bodies, we showed that $\langle \mathbf{F}, \mathbf{v} \rangle = 0$ for all tangent vectors $\mathbf{v}$ of the configuration space by differentiating the constraint equations, and using the fact that the internal forces between particles in Newton's third law lies along the direction between them. In the case of a pendulum, on the other hand, we simply verified the condition $\langle \mathbf{F}, \mathbf{v} \rangle = 0$ explicitly. The same situation will hold for all our constraint problems—we will simply assume that any internal forces are always perpendicular to the configuration space. In fact, a "constraint" more or less means a condition for which this holds (sometimes the term "ideal constraint" is used). Thus, as a general rule we have

d'Alembert's Principle for Constraints: If the constraints on a system confine the system to a configuration space $\mathcal{M}$, and are perpendicular to $\mathcal{M}$, then the motions of the system under the external forces $\mathbf{F}$ satisfy

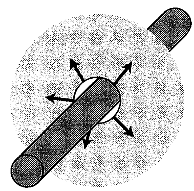
$$\langle \mathbf{F} - m\mathbf{c}'',\ \mathbf{v} \rangle = 0 \qquad \text{for all } \mathbf{v} \text{ tangent to } \mathcal{M}_{\mathbf{c}}.$$

In the case of rigid bodies, our formulation of d'Alembert's principle was more precise: we showed, conversely, that if a motion of the rigid body satisfies this equation, then the appropriate internal forces exist. In the case of constraints, this will usually be more or less automatic, and, as in our pendulum example, the requisite constraint forces can usually be found explicitly, unlike the case of a rigid body, where the internal constraints aren't unique.

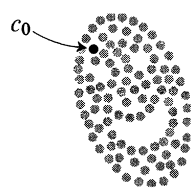
The compound (physical) pendulum. Since the only purpose of the pendulum string was to maintain the distance between the pivot and the bob, we might have simply considered these two particles as constituting a rigid body themselves. More generally, we can consider a "compound pendulum", or "physical pendulum", an arbitrarily shaped pendulum oscillating around an axle A; this is essentially a 2-dimensional problem, although we can think of the pendulum as a thin plate.



This problem illustrates a bit more clearly than the simple pendulum problem the numerous abstractions that our analysis entails. In actuality, the pendulum is constrained to rotate about the axle by various forces all along the circumference



of the circular hole around the axis. But that's obviously a bit more complicated a situation than we want to consider, and for theoretical purposes it will be better to imagine the pendulum not as a continuous body, but as a collection of particles, one of which, c_0, is kept fixed. Thus we are assuming that the



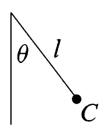
constraint forces keep the point c_0 at distance 0 from some point P. Of course, this theoretical picture is a little weird, since the constraint forces are supposed to act along the line between c_0 and P, which doesn't tell us anything! But our theoretical picture is still capable of encompassing this situation, because of one important fact: Even though we can't specify the direction of this constraint force $\mathbf{C}$, we will have $\langle \mathbf{C}, \mathbf{v}_0 \rangle = 0$ for all virtual infinitesimal displacements $\mathbf{v} = (\mathbf{v}_0, \dots, \mathbf{v}_K)$ of our pendulum under this constraint, for the simple reason that $\mathbf{v}_0 = 0$, since the constraint keeps c_0 fixed.

This means that we can still use d'Alembert's principle for constraints, where we now have the constraint that our collection of particles constitute a rigid body, together with the new constraint that the particle c_0 stays fixed. Our configuration space is just a circle once again, although now we don't think of it as a circle of a particular radius, but simply as the collection of angles θ through which our pendulum can rotate. Restricting ourselves to this configuration space takes care of the constraint that c_0 stays fixed, and all the other constraints have already been analyzed: thus, we just want to apply equation ($\boldsymbol{\tau}_{\text{axis}}$) on page 191.

The torque $\boldsymbol{\tau}$ due to gravity is easily computed: Since the force on particle c_i is $gm_i\mathbf{u}$, where $\mathbf{u}$ is the unit downward vector field, we get

$$\begin{aligned}\boldsymbol{\tau} &= \textstyle\sum_i c_i \times gm_i\mathbf{u} \\ &= g \cdot \left(\textstyle\sum_i m_i c_i\right) \times \mathbf{u} \\ &= gM \cdot C \times \mathbf{u},\end{aligned}$$

where C is the center of mass of the pendulum. This means that $\boldsymbol{\tau}$ points in



the direction of the axis A, and has magnitude

$$gMl \sin\theta,$$

where θ is the angle of C with the vertical, and l its distance from the pivot. So equation ($\boldsymbol{\tau}_{\text{axis}}$) becomes

$$\theta'' + \frac{gMl}{I_A}\sin\theta = 0.$$

Comparing to the pendulum formula (P), we see that our pendulum acts precisely like a single bob pendulum whose distance from the pivot is

$$\frac{I_A}{Ml}.$$

If I_C is the moment of inertia about the center of gravity C, then by the parallel axis theorem we have

$$\frac{I_A}{Ml} = \frac{I_C + Ml^2}{Ml} = l + \frac{I_C/M}{l}.$$

Introducing the *radius of gyration* k by

$$I_C = Mk^2,$$

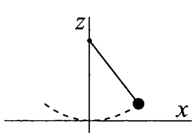
we find that our pendulum acts precisely like a single bob pendulum whose distance from the pivot is

$$l + \frac{k^2}{l}.$$

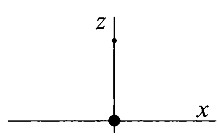
When our rigid body pendulum consists only of the pivot and a single particle at C, we have $k = 0$, but in any other case the pendulum will have a longer period.

Our analysis of the simple pendulum also shows that the force that our rigid body exerts on the pivot—that is, the sum of all the various forces that the particles in our rigid body exert on the pivot—is directed along the line from the center of mass to the pivot, and has magnitude $gM \cos\theta$. The decomposition of this total force is not uniquely determined, just as the necessary internal forces in the pendulum are not uniquely determined, but in practice, of course, most of the force exerted on the pivot will come from particles of the rigid body that are near to it, even though mathematically we might think of it as a single force exerted on the pivot by the center of mass, though this might not even be a particle of the rigid pendulum!

Equilibrium and stability. Before examining other sorts of constraints, we will use the physical pendulum as a simple example to discuss an important concept. For an ordinary pendulum (a), we can ask if there are any *equilibrium* points, a position where the pendulum can simply remain motionless. Of course, the



(a)



(b)

obvious answer is that the equilibrium position is the one where the pendulum is simply hanging straight down (b) with velocity 0. Equilibrium positions are of interest because for a real pendulum, gradually slowing down because of friction at the pivot and air resistance, and for many other realistic mechanisms, the equilibrium positions are those where the mechanisms may eventually come to rest. Before settling down to these positions, they will often exhibit small oscillations, and the equilibrium positions are the ones around which such small oscillations can occur.

If we consider the case of a physical pendulum, so that the "string" is actually a thin rigid rod, whose mass we might consider to be negligible, with a bob being a particle of mass M, there is another equilibrium point, with the pendulum

vertically above the pivot, with velocity 0. Naturally, this situation is clearly a little different—it's easy to obtain the first equilibrium position and virtually impossible to achieve the second, but we'll leave that matter hanging for now.

At the moment, we simply want to recall that the force $\mathbf{F}$ on the pendulum bob, a downward force of magnitude gM, is given by

$$\text{(a)} \qquad \mathbf{F} = -\left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial z}\right)$$

for the potential function

$$V(x,z) = gM \cdot z,$$

and note that the two equilibrium points occur at the maximum and minimum of V when restricted to our configuration space.

In fact, this observation generalizes for any case of d'Alembert's principal for constraints, for a configuration space $\mathcal{M} \subset \mathbb{R}^N$, with a force $\mathbf{F} \in T\mathbb{R}^N$ having a potential function $V\colon \mathbb{R}^N \to \mathbb{R}$. As noted on page 90, we have $\langle \mathbf{F}, \mathbf{v}\rangle = -\mathbf{v}(V)$ for all tangent vectors $\mathbf{v}$. At a critical point p of V on our configuration space, where $\mathbf{v}(V) = 0$ for all tangent vectors $\mathbf{v} \in \mathcal{M}_p$, we thus already have

$$\langle \mathbf{F}(p), \mathbf{v}\rangle = 0 \qquad \text{for all } \mathbf{v} \in \mathcal{M}_{\mathrm{p}},$$

so the constant curve $\mathbf{c}(t) = p$ in $\mathcal{M}$, with $\mathbf{c}'(0) = \mathbf{c}''(0) = 0$, definitely satisfies

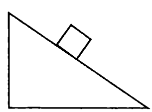
$$\langle \mathbf{F}(p) - m\mathbf{c}''(0), \mathbf{v}\rangle = 0 \qquad \text{for all } \mathbf{v} \in \mathcal{M}_{\mathrm{p}},$$

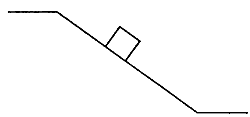
and is thus a possible motion. It is also clear, conversely, that if p is an equilibrium point, then p must be a critical point of V. Several examples are given in the Problems.

In our one-dimensional pendulum example, the minimum critical point is "stable": if we start close enough to this position, with velocity close to 0, then the pendulum will stay close to the minimum equilibrium point. On the other hand, the maximum critical point is not stable; in fact, for any other initial position and velocity, no matter how close to this equilibrium position, the pendulum will eventually move far away.

In general, decisions about stability are much harder in higher dimensions, and we will discuss certain aspects, the basic question of stability of solutions of differential equations near a 0 point of the corresponding vector field, in Addendum 8C.

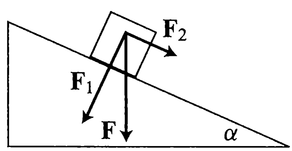
Sliding. The motion of a block sliding down another wedge shaped block is perhaps the simplest problem involving the motion of one rigid body with respect to another. This is often described as a block sliding down an inclined plane,





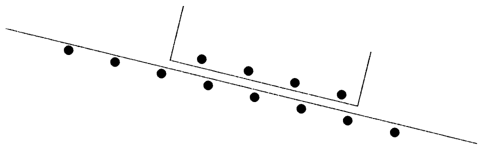
and for the present we are in fact considering the wedge as being immovable, rather than an object that can itself slide horizontally along the floor.

The usual elementary analysis of this problem has already been mentioned in Chapter 1 (page 30): we decompose the force $\mathbf{F}$ of gravity on the block into a force $\mathbf{F}_2$ parallel to the inclined plane, and a force $\mathbf{F}_1$ perpendicular to the inclined plane, and reason that $\mathbf{F}_2$ doesn't act on the inclined plane. In



Chapter 1 we simply said that $\mathbf{F}_1$ is presumed to be the force of the block on the inclined plane, so that the inclined plane must exert the force $-\mathbf{F}_1$ on the block. But we really need to use the hypothesis—now explicitly mentioned—that the inclined plane is stationery, using the same reasoning as for a block resting on a table: $\mathbf{F}_1$ determines the acceleration of the block in the direction perpendicular to the inclined plane, but that must be 0 (since the wedge is not moving and the block slides along it), so the inclined plane must be exerting a force of $-\mathbf{F}_1$ on the block.

On the other hand, as soon as one tries to think about this in terms of the physics of point masses, by imagining the "molecules" in the block and inclined plane, the whole argument appears dubious, since the forces between these point



masses ought to be along the line between them, and thus seldom perpendicular to the inclined plane.

Of course, we've neglected the important (implicit or explicit) qualification that the block is sliding *without friction*. And what does that mean? Why, it must mean precisely that when we look at the forces between the molecules of the block and the inclined plane, the component of such forces in the direction of the plane should be ignored. In other words, the inclined plane exerts a force $-\mathbf{F}_1$ on the block because that's what we're assuming.

Although it is reasonable to add this hypothesis in order to produce a simple theoretical problem—and mechanics is replete with problems in which we ignore friction—it would be nice to have a mental picture that provides some correlation with our notions of friction.

Friction, which we will discuss a bit more thoroughly in Chapter 11, is actually an incredibly complicated phenomenon, because it is an expression of intermolecular forces. As pointed out in the Feynman [1; pp. 12-3 to 12-5]:

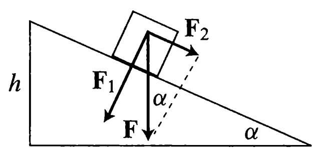
> The tables that list purported values of [friction] for "steel on steel," "copper on copper," and the like, are all false ... The friction is never due to "copper on copper," etc., but to the impurities clinging to the copper. ... If we try to get absolutely pure copper, if we clean and polish the surfaces, outgas the materials in a vacuum, and take every conceivable precaution, ...

then the copper block does not slide more easily along the inclined copper plane, in fact it does not slide at all—the two pieces of copper stick together, even if the inclined plane is completely vertical, because the copper atoms near the surfaces of the two pieces are attracted by the very same forces that keep the atoms within the individual pieces together.

Although that involves considerations far outside those of elementary mechanics, it might suggest that we think of our block as sliding along the inclined plane on protuberances, like furniture "gliders", which reduce the contact between two wooden objects. Indeed, if we imagine the theoretical case of a body

with a curved surface resting tangentially on the inclined plane, it would seem natural to assume in this case that the forces exerted by the two bodies upon each other would be perpendicular to their common tangent plane.

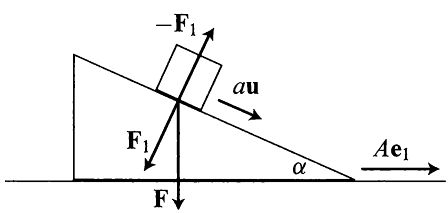
Once we've made the leap of faith to frictionless surfaces, we simply use the analysis on page 30, finding that the force $\mathbf{F}_2$ parallel to the inclined plane has magnitude $mg \sin\alpha$, so if $c(t)$ is the distance that the block has traveled along



the inclined plane after time t, then $c'' = g \sin\alpha$, and our block slides with a uniform acceleration that is $\sin\alpha$ times its free fall acceleration.

When we allow the inclined plane to be a wedge that slides along a horizontal plane, also without friction, our problem becomes more involved. The force $\mathbf{F}_1$ that the block exerts on the wedge can no longer be obtained simply by resolving the downward force of gravity $\mathbf{F}$ into forces perpendicular and parallel to the wedge, because our identification of $\mathbf{F}_1$ with the perpendicular component depended on the wedge being fixed.

Choosing the unit vector $\mathbf{e}_1 = (1, 0)$ parallel to the floor and the unit vector $\mathbf{u}$ parallel to the slope of the wedge, we let $A\mathbf{e}_1$ be the acceleration of the wedge, of mass M, along the horizontal plane, while $a\mathbf{u}$ is the acceleration of the block, of mass m, along the wedge, so that $a\mathbf{u} + A\mathbf{e}$ is the acceleration of the block in



our inertial system. Note that in our picture, we actually have $A < 0$, so that the arrow $A\mathbf{e}_1$ points in the opposite direction, since the force $\mathbf{F}_1$ causes the block to slide to the left.

Breaking up the equation

$$\text{(1)} \qquad -\mathbf{F}_1 + \mathbf{F} = m(a\mathbf{u} + A\mathbf{e}_1)$$

for the motion of the block into the components that are parallel and perpendicular to the slope of the wedge gives

$$\text{(1a)} \qquad mg\sin\alpha = ma + mA\cos\alpha$$

$$\text{(1b)} \qquad |\mathbf{F}_1| - mg\cos\alpha = mA\sin\alpha.$$

The force on the wedge, of mass M, is $\mathbf{F}_1$ plus the gravitational force downward, plus whatever upward force the horizontal plane must exert to keep the wedge from moving downwards. So A is determined by the horizontal component of $\mathbf{F}_1$:

$$\text{(2)} \qquad -|\mathbf{F}_1|\sin\alpha = MA.$$

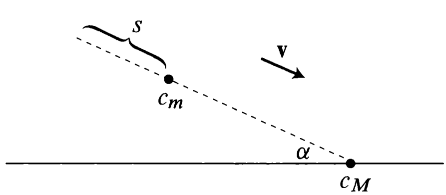
From (1b) and (2) we get

$$A = -g\left(\frac{\sin\alpha\cos\alpha}{\sin^2\alpha + \frac{M}{m}}\right),$$

and then (1a) gives

$$a = g\sin\alpha - A\cos\alpha.$$

Just for fun, we will analyze this problem using configuration spaces. To do this, we regard our system as consisting of two "particles", the block c_m and the wedge c_M. Since the wedge c_M always stays on the horizontal axis, we'll simply



consider our problem as occurring in $(\mathbb{R}^2)\times\mathbb{R}$, with $((a,b),x)$ representing the particle c_m at the point (a,b), and the particle c_M at x.

Our configuration space $\mathcal{M}$ consists of all $((a,b),x)$ which represent the block c_m resting on the wedge c_M [i.e., for which we have $b = (x-a)\cos\alpha$].

A convenient coordinate system on $\mathcal{M}$ is provided by the coordinate $x \in \mathbb{R}$ giving the position of c_M, together with the distance s of c_m from the top of the wedge.

To determine $\partial/\partial s$, we keep x fixed and vary s, obtaining a curve in $\mathcal{M}$ whose $\mathbb{R}^2$ component moves down the slope of the wedge, while its $\mathbb{R}$ component is fixed, so for the unit vector $\mathbf{u}$ in $\mathbb{R}^2$ parallel to the slope of the wedge, we have

$$\frac{\partial}{\partial s} = (\mathbf{u}, 0).$$

On the other hand, if we keep s fixed and vary x, then we obtain a curve in $\mathcal{M}$ whose $\mathbb{R}^2$ component moves parallel to the first axis along with its $\mathbb{R}$ component, so for $\mathbf{e}_1 = (1, 0)$ we have

$$\frac{\partial}{\partial x} = (\mathbf{e}_1, 1).$$

Now if $s(t), x(t)$ are the functions describing the motion of c_m, c_M, we have

$$\begin{aligned} c_m(t) &= s(t)\mathbf{u} + x(t)\mathbf{e}_1 \in \mathbb{R}^2 \\ c_M(t) &= x(t) \in \mathbb{R}, \end{aligned}$$

so

$$\begin{aligned} c_m{}'' &= s''\mathbf{u} + x''\mathbf{e}_1 \in \mathbb{R}^2 \\ c_M{}'' &= x'' \in \mathbb{R}. \end{aligned}$$

The external forces $\mathbf{F}_m$ on c_m and $\mathbf{F}_M$ on c_M are given by

$$\begin{aligned} \mathbf{F}_m &= -mg\mathbf{e}_2 \\ \mathbf{F}_M &= 0, \end{aligned}$$

so our condition for a solution is that

$$\langle -mg\mathbf{e}_2 - mc_m{}'', \mathbf{v}_1 \rangle + (0 - Mc_M{}'') \cdot \mathbf{v}_2 = 0$$

for all $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2)$ tangent to $\mathcal{M}$, where $\langle\ ,\ \rangle$ is the usual inner product in the first factor $\mathbb{R}^2$, while the inner product in the second factor $\mathbb{R}$ is just ordinary multiplication. Choosing

$$\mathbf{v} = \frac{\partial}{\partial s} = (\mathbf{u}, 0) \qquad \text{and then} \qquad \mathbf{v} = \frac{\partial}{\partial x} = (\mathbf{e}_1, 1)$$

gives the two equations

$$\begin{aligned} 0 &= \langle -mg\mathbf{e}_2 - ms''\mathbf{u} - mx''\mathbf{e}_1, \mathbf{u} \rangle - Mx'' \cdot 0 \\ 0 &= \langle -mg\mathbf{e}_2 - ms''\mathbf{u} - mx''\mathbf{e}_1, \mathbf{e}_1 \rangle - Mx'' \cdot 1, \end{aligned}$$

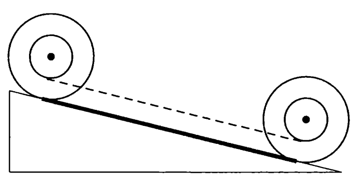
which amount to the equations

$$0 = mg\sin\alpha - ms'' - mx''\cos\alpha$$
$$0 = -ms''\cos\alpha - mx'' - Mx''.$$

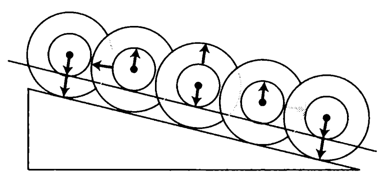
Solving for x'' and s'' gives us the same results that we obtained previously, when they were called A and a, respectively. The first of the above equations is precisely (1a), while the second is a combination of (1a), (1b), and (2).

Although this configuration space method doesn't explicitly use the force $\mathbf{F}_1$, it probably seems more complicated than the elementary method. In Part III we will see that Lagrangian mechanics provides a more convenient way of handling the problem, but the basic reasoning behind the use of Lagrangian mechanics for constraint problems is indicated by this straightforward use of configuration spaces, even though virtually all of the problems for this chapter are most easily solved by the elementary method.

Rolling. I do not know whether the wheel first appeared as the invention of some *Homo sapiens* genius whose identity has been obscured by the mists of human history, or, in the manner of "2001, A Space Odyssey", as a gift from an advanced intelligence. However, wheels have continued to bedevil the minds of people ever since, and even today one can encounter the bemusing "paradox" of two wheels rotating on a common axle, supposedly implying that the circumference of the smaller circle is the same as that of the larger one (something akin to this paradox appears near the beginning of Galileo [2].) One can easily



dispel any mystery this might present by using an apparatus that allows the two wheels to roll on separate tracks, rotating independently on the same axle, so

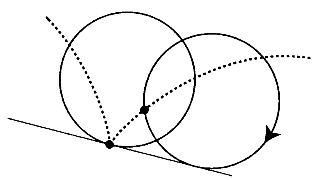


that one sees the smaller wheel rotating more than once in the same time that the larger wheel rotates exactly once. If one then secures the inner wheel to the axle on which the outer wheel is attached, say by a screw, and there is sufficient friction on the two tracks to prevent the wheels from sliding, the wheels will simply not move, or do so in a very jerky manner, because they both can't roll at once.

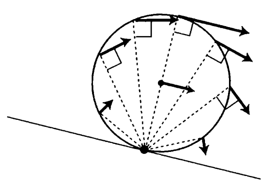
The moral of all this seems to be that "rolling" must mean moving in such a way that each arc of the wheel traces out a line of the same length—if a wheel doesn't have this property, then it just isn't "rolling". But this pronouncement is not very useful from the point of view of physics, because it doesn't explain physically "what is going on" (which, in a way, is what bothered Galileo); more precisely, it doesn't provide a way of analyzing a wheel rolling down an inclined plane in terms of various forces.

Physics books that discuss this problem in any detail point out that rolling depends on a truly paradoxical fact: a wheel only rolls because of frictional forces, ones that "oppose sliding": a wheel not only displays the characteristics of our abstract rigid body, but it also has the strange feature that it is affected by a frictional force that is exerted only at the (always changing) contact point of the wheel on the inclined plane. To make the picture even more confusing, this frictional force essential for rolling won't affect conservation of energy, because, the physics books note, the path followed by any point on the circumference of the wheel has velocity 0 at the moment it hits the plane!

It is indeed well known that a cycloid, the path followed a point on the circumference of a wheel, has velocity 0 at the point of contact, but physics books

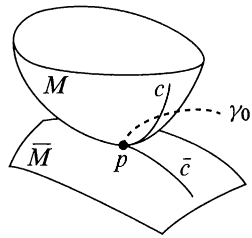


seem to regard this fact as intuitively clear, as well as another claim, that the motion at any time is essentially just a pure rotation about the contact point,



a fact that is sometimes invoked in discussing the phenomenon of two wheels rotating on a common axle. For those of us not endowed with the requisite physical intuition, here is a proof, for the general case of one surface rolling on another.

1. PROPOSITION. Consider two surfaces M and $\overline{M}$ in $\mathbb{R}^3$ that are tangent at a point p. Let c be a curve in M, and $\bar{c}$ a curve in $\overline{M}$ such that $c(0) = p = \bar{c}(0)$,



and such that $c'(0) \neq 0$ is a multiple of $\bar{c}'(0)$. For each t let $A(t)$ be the rigid motion for which

(a) $A(t)(c(t)) = \bar{c}(t)$,

(b) $A(t)(M)$ is tangent to $\overline{M}$ at $\bar{c}(t)$,

(c) $A(t)(c'(t))$ points in the same direction as $\bar{c}(t)$, so that

$$A(t)(c'(t)) = \alpha(t) \cdot \bar{c}'(t) \text{ for some function } \alpha.$$

Also, for each point $c(\tau)$ on the curve c, let γ_τ be the curve that this point follows under these rigid motions,

$$\gamma_\tau(t) = A(t)(c(\tau)).$$

Then the following are equivalent:

(1) $\alpha(t) = 1$ for all t, so that the lengths of c and $\bar{c}$ are the same on any time interval $[t_0, t_1]$.

(2) Each $\gamma_\tau{}'(\tau) = 0$, so that γ_τ has velocity 0 at the time that it hits $\overline{M}$.

(3) For each t we have

$$A'(t)(\mathbf{x}) = C(\mathbf{x} - \bar{c}(t))$$

for some skew-adjoint C (so that $A(t)$ is, "up to first order", a rotation about the point $\bar{c}(t)$).

PROOF. Write $A(t)$ in the form

$$A(t)(\mathbf{x}) = B(t)(\mathbf{x}) + \mathbf{w}(t) \qquad \mathbf{x} \in \mathbb{R}^3,$$

for orthogonal $B(t)$. Setting $\mathbf{x} = c(t)$ and using $A(t)(c(t)) = \bar{c}(t)$, we see that $\mathbf{w}(t) = \bar{c}(t) - B(t)(c(t))$, so we can write

$$A(t)(\mathbf{x}) = B(t)(\mathbf{x}) + [\bar{c}(t) - B(t)(c(t))].$$

The definition $\gamma_0(t) = A(t)(c(0)) = A(t)(p)$ gives

$$\text{(a)} \qquad \gamma_0(t) = B(t)(p) + \bar{c}(t) - B(t)(c(t)),$$

so that we can also write

$$\text{(b)} \qquad \begin{aligned} A(t)(\mathbf{x}) &= B(t)(\mathbf{x}) + \gamma_0(t) - B(t)(p) \\ &= B(t)(\mathbf{x} - p) + \gamma_0(t). \end{aligned}$$

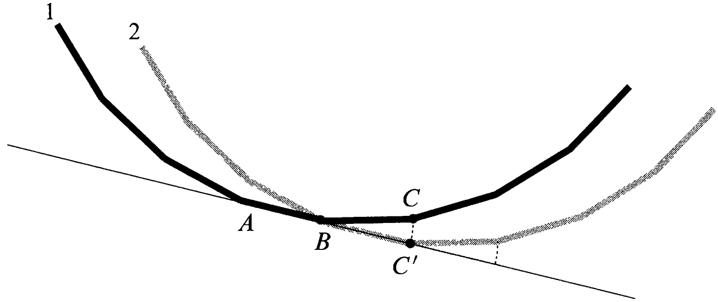
Since $B(0)$ is the identity, differentiating (a) gives

$$\begin{aligned} \gamma_0{}'(0) &= B'(0)(p) + \bar{c}'(0) - B'(0)(p) - c'(0) \\ &= \bar{c}'(0) - c'(0) \\ &= \bar{c}'(0) - \alpha(0) \cdot \bar{c}'(0). \end{aligned}$$

So $\alpha(0) = 1$ if and only if $\gamma_0{}'(0) = 0$, and by (b) this is true if and only if $A'(0)(\mathbf{x}) = B'(0)(\mathbf{x} - p)$, where $B'(0) = B'(0)B(0)^{\mathbf{t}}$ is skew-adjoint.

Our hypotheses on $A(t)$ then allow us to use this same argument at any point $c(\tau)$ by considering the reparameterization $t \mapsto t + \tau$. ❖

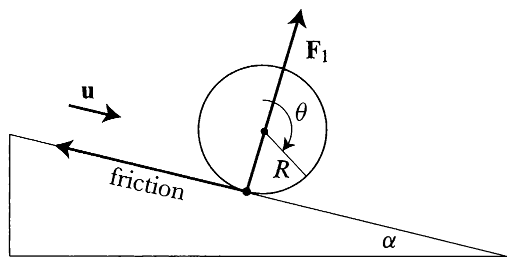
For the case of a wheel rolling down an inclined plane, we can provide a picture that both reinforces this geometric information and also allows us to see "what is going on" physically. We regard our circular wheel as a polygon with a very large number of sides, and suppose that initially, in position 1, it is lying on the inclined plane along the segment AB. Now, instead of sliding down the plane, it *rotates* about the point B, reaching position 2 when vertex C hits the



inclined plane, at C'. Then it rotates around C', and so forth. (Galileo uses a picture of this sort for the case of two wheels rotating on a common axle, and his analysis makes for very interesting reading—it is the sort of inspired nonsense that only a genius would come up with, or at any rate, get away with.)

Let's return to the special case of a wheel rolling down an inclined plane, once again assumed to be immovable. Our "wheel" is really supposed to be a cylinder, so that it is forced to roll along a straight line, but the 2-dimensional cross-section picture provides all the interesting information.

We consider a wheel of radius R and mass M and uniform density, and let $\theta(t)$ be the angle through which it has rotated after time t. For a unit vector $\mathbf{u}$



pointing down along the inclined plane, we will let $a \cdot \mathbf{u}$ be the acceleration of the center of mass, and let $-f \cdot \mathbf{u}$ be the frictional force along the inclined plane at the contact point of the wheel and the plane. The total force on the wheel is the sum of the downward force of gravity, a constraining force $\mathbf{F}_1$ perpendicular to the plane, which keeps the wheel from moving perpendicularly to the plane, and the frictional force $-f \cdot \mathbf{u}$.

For the acceleration $a \cdot \mathbf{u}$ of the center of mass we have

$$(1) \qquad Ma = Mg\sin\alpha - f,$$

since $Mg\sin\alpha$ is the magnitude of the component of the gravitational force along the inclined plane, while the constraining force $\mathbf{F}_1$ is perpendicular to $\mathbf{u}$. Note that since our wheel is a rigid body, the force f acts on the center of mass (cf. remarks on page 196).

For the rotational motion about the center of mass we can apply equation ($\boldsymbol{\tau}_{\text{axis}}$) on page 191 to the axis through the center of mass that is perpendicular to the plane of the drawing to get

$$(2) \qquad I\theta'' = Rf,$$

where I is the moment of inertia of the wheel about its center of mass.

Finally, the fact that our wheel is rolling tells us that the distance traveled by the center of mass at time t is equal to $R \cdot \theta(t)$, which means that

$$(3) \qquad a = R \cdot \theta''.$$

Solving (1)–(3) gives

$$a = g \sin \alpha \frac{1}{1 + \dfrac{I}{R^2 M}},$$

and substituting into (1) then gives

$$f = Mg \sin \alpha \frac{I}{R^2 M + I}.$$

In terms of the radius of gyration k defined (page 213) by

$$I = Mk^2,$$

we have

$$a = g \sin \alpha \frac{1}{1 + (k^2/R^2)}, \qquad f = Mg \sin \alpha \frac{k^2}{k^2 + R^2}.$$

Notice, by the way, that this is always non-negative. If the wheel is rolling up the incline plane, because of some initial impulsive force at the bottom, the frictional force opposing sliding is still directed upwards.

The coefficient $(Mg \sin \alpha k^2)/(k^2 + R^2)$ gives the amount of frictional force that the inclined plane must be able to produce in order to prevent sliding. As a very general rule, we can say that the frictional force that a body produces on an inclined plane is proportional to the normal component of the gravitational force, i.e., it equals $\mu \cdot Mg \cos \alpha$ for a constant μ, the "coefficient of friction". So to prevent sliding at the angle α we need to have

$$\mu \cdot Mg \cos \alpha = Mg \sin \alpha \frac{k^2}{k^2 + R^2}$$

or

$$\mu = \tan \alpha \frac{k^2}{k^2 + R^2}.$$

Thus we will need a "perfectly rough" surface, with "$\mu = \infty$" if we want to prevent sliding at any angle.

If our wheel is simply a homogeneous disc of radius R and mass M, then (Problem 5-6) the moment of inertia I is $\frac{1}{2}MR^2$, with $k^2 = \frac{1}{2}$, so we have

$$a = \tfrac{2}{3} g \sin \alpha.$$

Thus our wheel rolls down the incline plane at only $2/3$ of the speed that a block slides down a frictionless inclined plane. But the difference is not due to the "infinite friction" of the inclined plane on the wheel (the frictional force does no work), but to the fact that the kinetic energy of the wheel has both a translational part T_{transl} and a rotational part T_{rot} (page 194) and we find (Problem 8) that the total kinetic energy at the bottom of the plane is precisely Mgh where h is the initial height of the wheel.

One aspect of our solution deserves particular notice. When $\alpha = 0$, so that the wheel is simply rolling under no force at all, we have $f = 0$. This shouldn't be surprising—with no forces acting on it, the wheel presumably moves with constant velocity, and a frictional force would change that. On the other hand, friction is supposed to be what causes rolling. So in this case it's the frictional force of 0 that is responsible for rolling! One apparently has to accept this as an inevitable consequence of our idealizations.

Even though a rolling wheel on an inclined plane does involve friction, it is still a natural candidate for treatment by d'Alembert's principle for constraints: If we consider our wheel as a rigid body made up of a large collection of particles, and let $\mathbf{v}$ be any virtual infinitesimal displacement of the wheel, then the inner product $\langle \mathbf{f}, \mathbf{v}_p \rangle$ of the frictional force $\mathbf{f}$ and the velocity $\mathbf{v}_p$ at the point of contact p is always 0, since $\mathbf{v}_p = 0$ according to Proposition 1.

The main difficulty is that we have a sort of hybrid between our initial pendulum bob problem, where we considered a single particle acted upon by a constraint force, and the problem of a pendulum as a rigid body, where almost all our constraints had already been considered in the analysis of rigid body motion. We really need to think of our wheel as representing two different "particles" s and θ in $\mathbb{R}$,

$$\begin{aligned} s &= \text{position of the center of mass} \\ \theta &= \text{angle through which wheel has turned,} \end{aligned}$$

having the respective masses

$$\begin{aligned} M &= \text{the total mass of the wheel} \\ I &= \text{the moment of inertia of the wheel.} \end{aligned}$$

[Lagrangian mechanics will enable us to handle this more directly].

We thus have a problem in $\mathbb{R}^2$ that is reduced to a problem on a 1-dimensional submanifold $\mathcal{M} \subset \mathbb{R}^2$ by the rolling condition

$$\theta(t) = s(t)/R.$$

On $\mathcal{M}$ we have the obvious single coordinate s, the distance along the inclined plane, and the corresponding tangent vector to $\mathcal{M}$ represents the pair

$$(1, 1/R).$$

We are looking for a function $s(t)$ such that

$$(s(t), \theta(t)) = (s(t), s(t)/R)$$

satisfies

$$\begin{aligned} 0 &= \langle\, (-\mathbf{F}_s - Ms'', 0 - I\theta''),\ (1, 1/R)\,\rangle \\ &= \langle\, (-\mathbf{F}_s - Ms'', 0 - Is''/R),\ (1, 1/R)\,\rangle, \end{aligned}$$

where $\mathbf{F}_s$ is the component of the force on the center of mass that is parallel to the inclined plane. Thus we get

$$\begin{aligned} 0 &= Mg\sin\alpha - Ms'' - (Is''/R)\cdot 1/R \\ &= Mg\sin\alpha - Ms'' - Is''/R^2, \end{aligned}$$

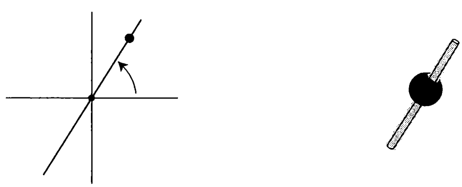
which gives the same result

$$s'' = g\sin\alpha\,\frac{1}{1 + \dfrac{I}{R^2M}}$$

that we obtained previously by combining three equations.

Some subsidiary topics. Before continuing with the major considerations of this chapter, we tie up some loose ends by considering a couple of subsidiary topics.

1. *Time-dependent constraints.* Consider the 2-dimensional problem of a bead that can slide without friction along a rigid wire that is rotating about the origin in some plane, with no other external forces on the bead. Naturally, we won't worry about the particular details of the forces involved around the hole in the



bead, but simply consider the bead to be a point mass constrained to lie on the wire at all times.

We can't choose our configuration space $\mathcal{M}$ to be the set of all possible positions of the bead, since that is all of $\mathbb{R}^2$, rather than a 1-dimensional manifold. But we can apply the configuration space method by means of the standard trick of introducing time as another variable. In other words, instead of a particle in $\mathbb{R}^2$, we consider a particle $c = (x, y, \tau)$ in $\mathbb{R}^3$ with the constraint that $(x(t), y(t))$ lies on the wire at time t and the additional constraint that $\tau(t) = t$.

If $\theta(t)$ is the angle that the wire makes at time t, and $r(t)$ is the distance of the bead from the origin, then we have

$$c(t) = \big(r(t)\cos\theta(t),\ r(t)\sin\theta(t),\ t\big)$$

and we find that

$$c'' = [r'' - r\theta'^2]\cdot(\cos\theta, \sin\theta, 0) + [r'\theta' + (r\theta')']\cdot(-\sin\theta, \cos\theta, 0).$$

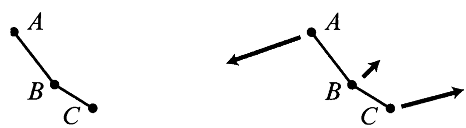
Our constraints restrict c to lie on a 1-dimensional manifold for which r is a coordinate system, and $\partial/\partial r = (\cos\theta, \sin\theta, 0)$. Our one equation

$$0 = \langle -c'',\ \partial/\partial r\rangle$$

then reduces to $r'' = \theta'^2 r$.

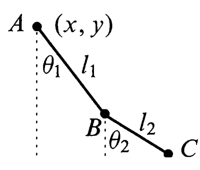
2. *Hinges.* So far, we have been considering almost exclusively only rigid bodies, or systems like a pendulum that act essentially like one, but we often need to consider systems involving rigid pieces that are "hinged" together, if, for example, we want to be able to analyze the motions of living objects, as discussed on page 86.

Naturally, the actual details of such hinges are mind-boggling complex, but for theoretical purposes we can simply imagine a "linkage" of three particles



where the distances between the particles A and B and between B and C must remain constant, while the angle $\angle ABC$ can vary freely, with forces applied at any of these points.

In terms of the lengths l_1 of AB and l_2 of BC, our system can thus be described by specifying the position (x, y) of A, and the angles θ_1 and θ_2 that AB and BC make with the perpendicular, so our configuration space is just $\mathbb{R}^2 \times S^1 \times S^1$.



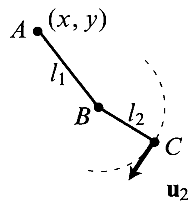
If we consider the case where A is fixed, reducing our configuration space to $S^1 \times S^1$, the position of C is given in terms of the coordinates θ_1, θ_2 by

$$(l_1\sin\theta_1 + l_2\sin\theta_2, l_1\cos\theta_1 + l_2\cos\theta_2),$$

and

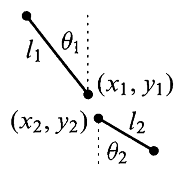
$$\begin{aligned}\partial/\partial\theta_1 &= (l_1\cos\theta_1, -l_1\sin\theta_1)\\ \partial/\partial\theta_2 &= (l_2\cos\theta_2, -l_2\sin\theta_2).\end{aligned}$$

Geometrically, $\partial/\partial\theta_2$ is l_2 times the unit vector $\mathbf{u}_2$ at C that is perpendicular to the circle of radius l_2 with center B, while $\partial/\partial\theta_1$ has length l_1, but points along a vector in the direction of the vector at B that is perpendicular to the circle of radius l_1 with center A.



When A and B are particles of masses m_1 and m_2 we obtain the "double pendulum", with rather complicated equations of motion. We're not going to derive these equations here, although we will mention a special case in Chapter 8. For now, we merely want to point out that such "hinges" can be treated by the methods at our disposal, and we can theoretically consider the motion of even more complicated linkages.

It should also be pointed out that we might consider our problem in a rather different way: We begin with two rigid rods, one described by the coordinates (x_1, y_1) and θ_1, the other by the coordinates (x_2, y_2) and θ_2, and then add the



additional constraint that $(x_1, y_1) = (x_2, y_2)$. Here we would be imagining a pair of additional constraint forces, $\mathbf{F}_{12} = -\mathbf{F}_{21}$, directed along the line from (x_1, y_1) to (x_2, y_2). Since the configuration space $\mathcal{M}$ will involve only points with $(x_1, y_1) = (x_2, y_2)$, tangent vectors to $\mathcal{M}$ will involve two equal vectors $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{w}$ at (x_1, y_1) and (x_2, y_2), and consequently

$$\langle \mathbf{v}_1, \mathbf{F}_{21}\rangle + \langle \mathbf{v}_2, \mathbf{F}_{12}\rangle = \langle \mathbf{w}, \mathbf{F}_{21} + \mathbf{F}_{12}\rangle = 0,$$

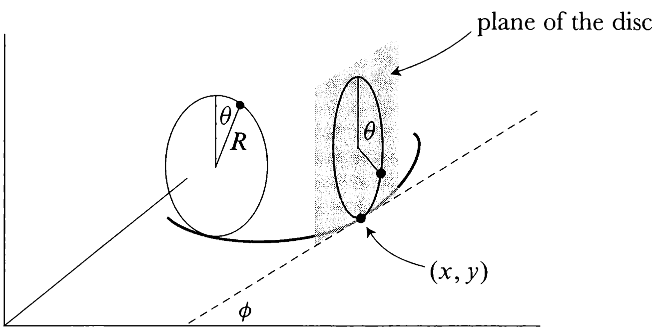
so that d'Alembert's principle for constraints will still apply.

Holonomic and differential constraints. All the constraint problems examined thus far were amenable to a treatment paralleling our treatment of rigid bodies, with the constraints in each case restricting our solutions to lie in a "configuration space" $\mathcal{M}$ that was a submanifold of the larger space for which the problem was originally posed. We have used the obvious principle that if you are looking for the solutions of a differential equation on a manifold $\mathcal{N}$, and you know that the solution lies on a submanifold $\mathcal{M} \subset \mathcal{N}$, then you might as well just consider what the equation says on $\mathcal{M}$, thereby obtaining an equation in fewer variables.

Physicists call such constraints "holonomic", and physics texts usually present holonomic constraints only in the context of Lagrangian mechanics. As we will see in Part III, Lagrangian mechanics relies on the same basic principle, but it allows us to circumvent the main difficulty encountered in all our examples— the task of expressing the tangent vectors $\partial/\partial x_i$ for a coordinate system on $\mathcal{M}$ in terms of the standard coordinates on the enveloping space, and of expressing the acceleration $c''(t)$ of each particle in terms of the coordinate functions $x_i(c(t))$ on $\mathcal{M}$. Lagrangian mechanics instead provides a systematic way of writing down the final equations, without going through such intermediate steps.

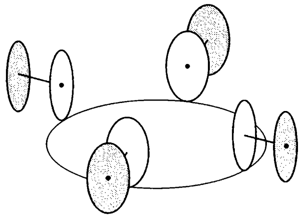
Physics texts also mention all sorts of other constraints, like the constraint that particles remain within a given box, or outside of a given sphere. Such constraints are expressed by inequalities or more complicated conditions, and obviously require special considerations in each individual case. But there is one other very important sort of non-holonomic constraint that allows a systematic treatment.

The standard example of this kind of constraint is provided by an upright disc rolling on a plane. The possible positions of the disc are determined by the



coordinates (x, y) of the point at which the disc rests on the plane, the angle θ that a fixed point on the disc makes with the vertical, and the angle ϕ that the plane of the disc makes with the x-axis.

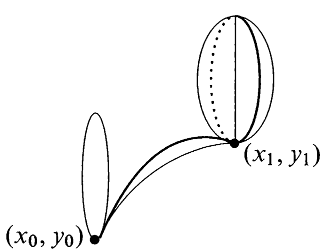
This example is rather idealized. To begin with, in order for the disc to remain upright, we might imagine that it has a companion disc attached to it by an axle (for the non-upright disc, see Addendum 9A and Addendum 12A). We will want to assume that the axle and the companion disc both have negligible weight, and it is also important that the two discs be able to rotate independently



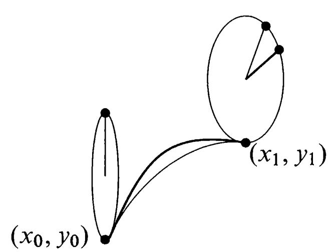
about this axle, so that, for example, the disc can revolve around a circle, with its companion "shadow" disc revolving around a circle of a different radius. In addition, although our disc has to have *some* thickness, we want to imagine it to be so small that it actually can roll along a circle—or, indeed, along any path—rather than being constrained to roll along a straight line.

In the simplest case, where there are no external forces, it is easy to guess from the symmetry of the situation that the disc of mass m and angular momentum $\mathbf{L}$ will roll with constant speed along a circle of radius $m/|\mathbf{L}|$, or along a straight line when $\mathbf{L} = 0$. But that doesn't suggest a general method for solving the problem where there are external forces, for example if we tilt the plane, so that now the force of gravity is only partially offset by the constraining perpendicular force of the plane.

Unfortunately, we are stymied when we try to use our method of configuration spaces to reduce the problem to one in fewer variables. Starting with our disc at a point (x_0, y_0), we can roll it, as in (a), to a nearby point (x_1, y_1) along paths that all start in the same direction at (x_0, y_0) but reach (x_1, y_1) at different angles, so that we obtain a whole interval of possible ϕ values. Moreover, we can also roll it, as in (b), along paths that all have the same direction at both (x_0, y_0) and (x_1, y_1), thereby obtaining a whole interval of possible θ values. Thus, the proper configuration space for this problem is a whole neighborhood of $\mathbb{R}^2 \times S^1 \times S^1$, rather than a lower-dimensional submanifold.



(a)



(b)

This phenomenon is a reflection of a simple fact about the relations between the coordinates of our disc moving in the space with coordinate functions (x, y, θ, ϕ). Letting $x(t), y(t), \theta(t), \phi(t)$ denote the components of the coordinates of the disc, the velocity of the center of mass is $R\theta'$, where R is the radius of the disc, and consequently (refer to the figure on page 230) we have

$$
\begin{aligned}
x' &= R\theta' \cos\phi \\
y' &= R\theta' \sin\phi.
\end{aligned} \tag{1}
$$

This means that the tangent vectors of the curve satisfy

$$
\begin{aligned}
0 &= dx - R\cos\phi\, d\theta \\
0 &= dy - R\sin\phi\, d\theta,
\end{aligned} \tag{1'}
$$

In particular, they therefore satisfy the condition

$$dy - \tan\phi\, dx = 0.$$

This determines a 3-dimensional subspace of all tangent vectors at each point, but this 3-dimensional distribution isn't integrable, as we can easily see from the standard integrability conditions. For example, to apply the differential form version of the Frobenius integrability theorem, we simply note that the 2-form

$$d(dy - \tan\phi\, dx) = \sec^2\phi\, d\phi \wedge dx$$

isn't in the ideal generated by $dy - \tan\phi\, dx$. Equivalently, we can note that the distribution is spanned by the vectors

$$X_1 = \frac{\partial}{\partial x} - \tan\phi \frac{\partial}{\partial y}, \qquad X_2 = \frac{\partial}{\partial \phi}, \qquad X_3 = \frac{\partial}{\partial \theta},$$

but the bracket

$$\left[\frac{\partial}{\partial x} - \tan\phi \frac{\partial}{\partial y}, \frac{\partial}{\partial \phi}\right] = -\sec^2\phi \frac{\partial}{\partial y}$$

obviously cannot be written as a linear combination of the three vectors X_1, X_2, and X_3.

Thus, although we have a condition that must be satisfied by tangent vectors to a solution curve, we can't select a 3-dimensional configuration space on which the solution curves must lie. We can only say the following:

$$\text{We must have } \langle \mathbf{F} - m c'',\ \mathbf{v}\rangle = 0$$
$$\text{for all } \mathbf{v} \in \ker(dx - R\cos\phi\, d\theta) \cap \ker(dy - R\sin\phi\, d\theta).$$

Here $\mathbf{F}$ is evaluated at $(\mathbf{c}(t), t)$ and c'' is evaluated at t, while $dx - R\cos\phi\, d\theta$ and $dy - R\sin\phi\, d\theta$ are evaluated at $c(t)$.

More generally, if the conditions in (1′) are replaced by the vanishing of certain 1-forms $\omega_1, \ldots, \omega_L$, then

$$\text{We must have } \langle \mathbf{F} - m\mathbf{c}'', \mathbf{v}\rangle = 0$$
$$\text{for all } \mathbf{v} \in \ker\omega_1 \cap \cdots \cap \ker\omega_L.$$

In terms of the linear functional

$$\Phi(\mathbf{v}) = \langle \mathbf{F} - m\mathbf{c}'', \mathbf{v}\rangle$$

this condition says that

$$\ker\Phi \supset \ker\omega_1 \cap \cdots \cap \ker\omega_L.$$

We can now appeal to the very same vector space fact (Problem 5-1) that was used in the proof of Lemma 1 of the previous chapter, and conclude, applying the argument at each point, that

$$\Phi = \lambda_1\omega_1 + \cdots + \lambda_L\omega_L$$

for some functions $\lambda_1, \ldots, \lambda_L$, known as *Lagrange multipliers*. This leads us to the following criterion for solutions:

> **d'Alembert's Principle for Differential Constraints:** If the constraints on a system require the tangent vector of the motion to lie in the subspace $\ker(\omega_1) \cap \cdots \cap \ker(\omega_L)$, then there are Lagrange multipliers $\lambda_1, \ldots, \lambda_L$ such that the motions of the system under the external forces $\mathbf{F}$ satisfy
>
> $$\langle \mathbf{F} - m\mathbf{c}'', \mathbf{v}\rangle = \lambda_1\omega_1(\mathbf{v}) + \cdots + \lambda_L\omega_L(\mathbf{v})$$
> $$\text{for } \textit{all} \text{ tangent vectors } \mathbf{v} \text{ at } \mathbf{c}.$$

We want to apply this to the problem of an upright disc rolling on a plane, where we again have the relations

$$\begin{aligned} x' &= R\theta' \cos\phi \\ y' &= R\theta' \sin\phi \end{aligned} \tag{1}$$

and

$$\begin{aligned} 0 &= dx - R\cos\phi\, d\theta \\ 0 &= dy - R\sin\phi\, d\theta. \end{aligned} \tag{1$'$}$$

As in the case of the rolling wheel, we are not dealing with a single particle $c(t)$; and in the present situation we have to think of the disc as *three* different "particles", the particle (x, y) with mass M, the particle θ with mass I, the moment of inertia of the disc about the axle, and the particle ϕ with mass I_ϕ, the moment of inertia of the disc about a diameter. We thus have

$$(*)\quad \langle(-Mx'', -My'', -I_\phi\phi'', -I\theta''), \mathbf{v}\rangle$$
$$= \lambda_1(dx - R\cos\phi\, d\theta)(\mathbf{v}) + \lambda_2(dy - R\sin\phi\, d\theta)(\mathbf{v}) \qquad \text{for all } \mathbf{v}.$$

Taking $\mathbf{v} = \partial/\partial x$, $\partial/\partial y$, $\partial/\partial\phi$, and $\partial/\partial\theta$, this gives us the equations

$$-Mx'' = \lambda_1 \tag{2x}$$
$$-My'' = \lambda_2 \tag{2y}$$
$$I_\phi\phi'' = 0 \tag{2ϕ}$$
$$I\theta'' = \lambda_1 R\cos\phi + \lambda_2 R\sin\phi. \tag{2θ}$$

Differentiating our original constraint equations (1) gives

$$\begin{aligned} x'' &= R\theta''\cos\phi - R\theta'\phi'\sin\phi \\ y'' &= R\theta''\sin\phi + R\theta'\phi'\cos\phi, \end{aligned}$$

so substituting $(2x)$ and $(2y)$ into (2θ) gives

$$\begin{aligned} I\theta'' &= -MR\big[\big(R\theta''\cos\phi - R\theta'\phi'\sin\phi\big)\cos\phi \\ &\qquad\qquad +\big(R\theta''\sin\phi + R\theta'\phi'\cos\phi\big)\sin\phi\big] \\ &= -MR^2\theta''. \end{aligned} \tag{3}$$

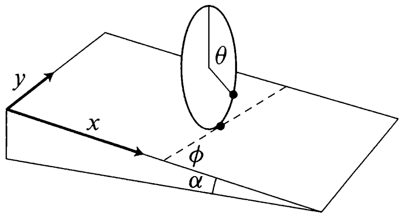
Thus $(I + MR^2)\theta'' = 0$, and θ' is constant.

Equation (2ϕ) shows that ϕ' is also constant, and if we substitute the two expressions

$$\begin{aligned} \theta(t) &= at + b \\ \phi(t) &= ct + d \end{aligned}$$

into (1) and solve, we find that (x, y) moves along a circle of radius Ra/c for $c \neq 0$, or a straight line if $c = 0$, in which case ϕ is constant.

Some discussion of the much more complicated case where the disc need not be vertical is postponed until Addendum 9A, but for now we can apply the same analysis to the more interesting case where our vertical disc is rolling down an inclined plane with slope α. The only change to $(*)$ is that the term $-Mx''$



must be replaced with $Mg\sin\alpha - Mx''$. In the set of equations (2), the only change is that equation (2x) is replaced by

$$gM\sin\alpha - Mx'' = \lambda_1.$$

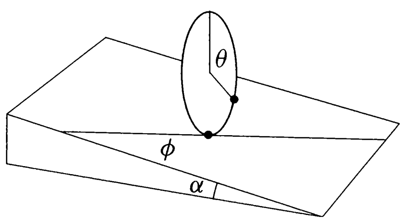
Proceeding as before, (3) then becomes

$$\begin{aligned}(I + MR^2)\theta''(t) &= (MgR\sin\alpha)\cos\phi(t)\\ &= (MgR\sin\alpha)\cos(ct + d),\end{aligned}$$

which, introducing an appropriate constant A, we write simply as

$$\theta''(t) = A\cos(ct + d).$$

The solutions of this equation—whether derived by Lagrangian mechanics or by our current method—yield interestingly complex possibilities for the motion. In the special case $c = 0$, the angle ϕ will be constant. This case is essentially just the same as the case of a wheel rolling down an inclined plane: the disc rolls down the inclined plane along the straight line that makes a constant angle with the x-axis.



For $c \neq 0$ we might as well take $d = 0$, so that $\phi(t) = ct$, since this just amounts to changing the point from which θ is measured, so there are constants B and C with

$$\theta'(t) = B\sin(ct) + C,$$

and thus

$$\begin{aligned} x' &= \tfrac{1}{2}RB\sin(2ct) + RC\cos(ct) \\ y' &= \tfrac{1}{2}RB(1-\cos(2ct)) + RC\sin(ct). \end{aligned}$$

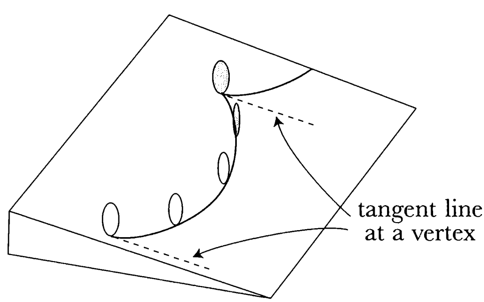
To see the general shape of the curve along which the wheel rolls, we can take $c = 1/2$, $R = 1$, and $B = 2$, so that

$$(*) \qquad \begin{aligned} x'(t) &= \sin t + C\cos t/2 \qquad & \phi(t) &= t/2 \\ y'(t) &= (1-\cos t) + C\sin t/2 \qquad & \theta'(t) &= 2\sin t/2 + C. \end{aligned}$$

First we consider the special case $C = 0$. We then have, up to additive constants,

$$\begin{aligned} x(t) &= 1 - \cos t \\ y(t) &= t - \sin t, \end{aligned}$$

which is the standard parameterization of a cycloid, with the x and y axes reversed. As in the following picture, the disc—white on one side and gray on the other—moves along the cycloid, which appears upside-down because x increases in the downward direction. To get this behavior we would need to



start the disc rolling straight down, but with a bit of spin. After the disc has traveled along the cycloid to the next vertex, we have $x'(t) = y'(t) = 0$, as well as $\theta'(t) = 0$. The disc then continues along the next arc of the cycloid; note that it doesn't just fall back along the arc already traversed, because it still has a non-zero spin $\phi'(t) = 1/2$.

The figure below shows the solutions to (∗), over the fundamental interval of length 4π, for increasing values of C. The curves become progressively taller and less symmetrical, and the tangent lines at the vertex become less vertical,

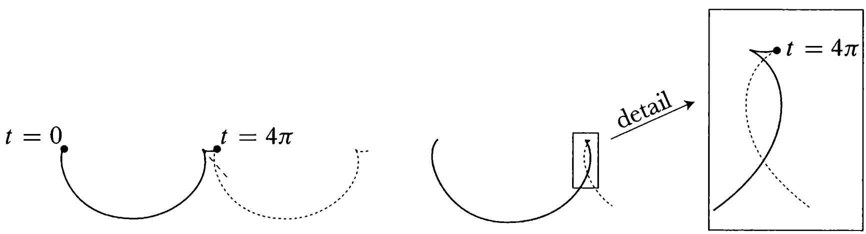
$C = 0$ $\quad C = .2 \quad$ $C = .6$ $\quad C = .8$

so the axis of the disc goes further beyond the vertical. The vertices occur at t with $x'(t) = y'(t) = 0$, both equations giving the same condition

$$\text{(V)} \qquad \sin\frac{t}{2} = -\frac{C}{2},$$

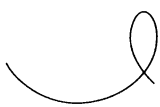
and we then have $\theta'(t) = 0$ at these vertices.

By $C = 1.4$ the curve is quite unsymmetrical, and by $C = 1.7$ it crosses over itself. On the other hand, since the equation (V) has no solutions at all for

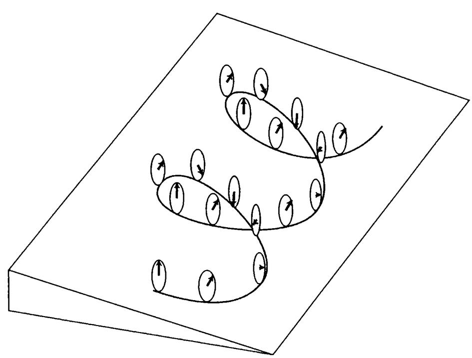


$C = 1.4$ $\quad C = 1.7$

$|C| > 2$, those curves are completely smooth, as in the case $C = 2.2$ illustrated below. The case $|C| = 2$, the transition between crossing curves and smooth



$C = 2.2$



ones, is also smooth, with $x'(\pi) = y'(\pi) = 0$; it looks essentially the same as the curve for $C = 2.2$, though it looks much closer to having a corner at the top loop, where it has $y''(\pi) = 0$.

Finding the constraint forces. Although the Lagrange multipliers λ_i that occur in d'Alembert's principle for differential constraints may seem to have appeared out of the blue, they may be interpreted in terms of the constraint forces $\mathbf{C}$ on our system S. In fact, consider two systems:

(a) the system S with constraints $\mathbf{C}$ and external forces $\mathbf{F}$,

(b) the system S with no constraints and external forces $\mathbf{F} + \mathbf{C}$.

The systems (a) and (b) obviously have the same solutions. But the solutions to (a) satisfy

$$\langle \mathbf{F} - mc'', \mathbf{v}\rangle = \lambda_1\omega_1(\mathbf{v}) + \cdots + \lambda_L\omega_L(\mathbf{v}) \qquad \text{for all } \mathbf{v},$$

while the solutions to (b) satisfy

$$\langle \mathbf{F} + \mathbf{C} - mc'', \mathbf{v}\rangle = 0 \qquad \text{for all } \mathbf{v};$$

subtracting the first equation from the second, we find that we must have

$$\langle \mathbf{C}, \mathbf{v}\rangle = -(\lambda_1\omega_1 + \cdots + \lambda_L\omega_L)(\mathbf{v}) \qquad \text{for all } \mathbf{v}.$$

By writing out $\lambda_1\omega_1 + \cdots + \lambda_L\omega_L$ in terms of the coordinates, we can then find all the components of $\mathbf{C}$.

For example, in our original problem of the disc rolling on a horizontal plane, where we have

$$\begin{aligned}\lambda_1\omega_1 + \lambda_2\omega_2 &= \lambda_1(dx - R\cos\phi\, d\theta) + \lambda_2(dy - R\sin\phi\, d\theta)\\ &= \lambda_1\, dx + \lambda_2\, dy - (\lambda_1 R\cos\phi + \lambda_2 R\sin\phi)\, d\theta\end{aligned}$$

we find that the components $\mathbf{C}_x$ and $\mathbf{C}_y$ are given by

$$\begin{aligned}\mathbf{C}_x &= \langle \mathbf{C}, \partial/\partial x\rangle = -\lambda_1\\ \mathbf{C}_y &= \langle \mathbf{C}, \partial/\partial y\rangle = -\lambda_2.\end{aligned}$$

Thus, the constraint forces can be found in terms of λ_1 and λ_2, which we can determine from $(2x)$ and $(2y)$ once we've solved explicitly for $x(t)$ and $y(t)$. The x and y components together, the vector $(x''(t), y''(t))$, represents the constraint force on our "particle" $(x(t), y(t))$, and thus the frictional force exerted by the plane to keep the center of mass in its circular orbit. Since the center of mass moves in a circle with constant angular velocity, $(x''(t), y''(t))$ is always perpendicular to the velocity vector $\mathbf{v} = (x'(t), y'(t))$ of the center of mass, as we would expect.

In the case of holonomic constraints, we didn't need to use the Lagrange multipliers λ_i, but we *can* use them, if we want to obtain the constraint forces. For example, consider the simple rolling wheel problem on page 226. Now we will simply use the coordinates s and θ and the relation

$$s'(t) = R\theta'(t)$$

between their derivatives. We then have the following condition for all $\mathbf{v}$:

$$\left\langle (Mg\sin\alpha - Ms'')\frac{\partial}{\partial s} - I\theta''\frac{\partial}{\partial\theta}, \mathbf{v} \right\rangle = \lambda(ds - R\,d\theta)(\mathbf{v}).$$

Taking $\mathbf{v} = \partial/\partial s$ and then $\mathbf{v} = \partial/\partial\theta$ we get

$$\text{(a)} \qquad Mg\sin\alpha - Ms'' = \lambda$$

$$\text{(b)} \qquad -I\theta'' = -R\lambda,$$

so

$$Mg\sin\alpha - Ms'' = \frac{I}{R}\theta''$$

and differentiating the constraint $s' = R\theta'$ gives $s'' = R\theta''$, so this becomes

$$Mg\sin\alpha - Ms'' = \frac{I}{R^2}s'',$$

with the same solution

$$s'' = g\sin\alpha\,\frac{1}{1 + \dfrac{I}{R^2M}}$$

as before. Substituting back into (a) then gives

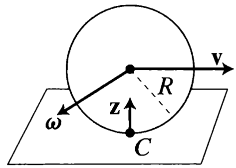
$$\lambda = Mg\sin\alpha\,\frac{I}{R^2M + I},$$

which agrees with the formula for the frictional force f on page 225.

Finally, we should point out that we can easily formulate and use a "mixed" version of d'Alembert's principle, where some of the constraints restrict our system to lie in a configuration space $\mathcal{M}$, while other constraints restrict the tangent vector of the system to be in the kernels of various 1-forms.

The rolling sphere. After all these exertions, the question of a rolling sphere (mathematically, a ball), even on a level plane, might seem quite intimidating. If $\mathbf{v}$ is the velocity of the center of the sphere of radius R, while $\mathbf{v}_C$ is the velocity of the contact point C, it is easily seen that

$$\mathbf{v}_C = \mathbf{v} + R\mathbf{z} \times \boldsymbol{\omega},$$



where $\mathbf{z}$ is the unit vector pointing upwards. So condition (2) of Proposition 1 on page 222 for rolling, $\mathbf{v}_C = 0$, gives

$$\mathbf{v} = -R\mathbf{z} \times \boldsymbol{\omega}$$

(which also follows immediately from the fact the sphere is instantaneously rotating about C). We also have (page 190) the equation

$$\boldsymbol{\tau} = I\boldsymbol{\omega}'.$$

The upwards force at the contact point C that balances the weight of the ball makes no contribution to the torsion $\boldsymbol{\tau}$ around the center, so $\boldsymbol{\tau}$ depends only on the frictional force $\mathbf{F}$ at the point of contact in the plane. Now from the considerations on page 226, we would suspect that $\mathbf{F} = 0$, so that $\boldsymbol{\tau} = 0$, and hence $\boldsymbol{\omega}$ is constant. But if $\boldsymbol{\omega}$ is constant, then the first equation shows that $\mathbf{v}$ is constant. In other words, it appears that the sphere should simply roll along a straight line with uniform speed.

To prove this, we substitute the above two equations into

$$\mathbf{F} = m\mathbf{v}'$$
$$\boldsymbol{\tau} = -R\mathbf{z} \times \mathbf{F}$$

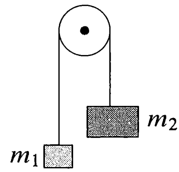
to deduce that

$$\begin{aligned} I\omega' &= -mR\mathbf{z} \times (-R\mathbf{z} \times \omega') \quad (\Longrightarrow \ \omega' \text{ is horizontal}) \\ &= -mR^2\omega', \end{aligned}$$

so that $\boldsymbol{\omega}' = 0$. Problem 14 explains that this result doesn't mean quite what you might expect.

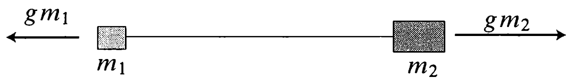
The paradoxical contrast between the disc rolling on a plane and the more restricted motion of a sphere rolling on a plane arises because the condition that a sphere is rolling is stronger than the condition that a disc is rolling, since we have to account for the motion of the contact point in two directions rather than just one: when the upright disc moves in a circle, there must be a centripetal force directed toward the center, but this comes from a frictional force perpendicular to the plane of the disc, and thus irrelevant to the question of whether the disc rolls.

Give a physics student enough rope problems ... Elementary physics texts seem to delight in presenting problems involving ropes (or strings, or some other type of "filament"), like the classic example of two weights on opposite sides of a



pulley, which is often referred to in mechanics texts by the rather mysterious name of "The Atwood machine".

The simplest way of dealing with the Atwood machine is to regard the pulley merely as a device that allows the oppositely directed forces to be produced by the force of gravity acting in the same direction: in other words, we treat



the problem as if it involved oppositely directed forces on two weights that are attached to a long, weightless, rigid rod. The implicitly assumed unstretchability of the rope, or the rigidity of the corresponding rod, merely insures that the two weights stay a constant distance apart.

Since we will soon be analyzing the forces that ropes entail in some detail, for now let's simply solve this presumably equivalent problem like any other constraint problem. Our system is determined by the positions x_1, x_2 of the weights along the line, and since $x_2 - x_1$ is constant, we have a 1-dimensional configuration space $\mathcal{M}$, and the single equation

$$(gm_1 - m_1x_1'') + (-gm_2 - m_2x_2'') = 0;$$

together with $x_2'' = x_1''$, this gives

$$x_1'' = g\frac{m_1 - m_2}{m_1 + m_2}$$

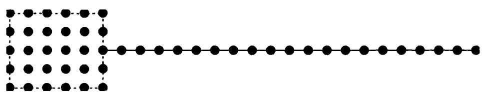
In particular, if we start with equal masses M, and then add a very small mass m to one side, the acceleration

$$x_1'' = g\frac{m}{2M + m}$$

will have a very small value, allowing us to measure it much more accurately than we could directly measure the acceleration g of a body falling freely. The first mechanism of this sort was constructed by the Rev. George Atwood (1746–1807), a tutor at Trinity College, Cambridge. Atwood, not having ball-bearings, employed a rather ingenious mechanism to reduce friction; for some very nice pictures go to the web site `physics.kenyon.edu/EarlyApparatus` and click on `Mechanics`, and then on `Atwood's Machine`.

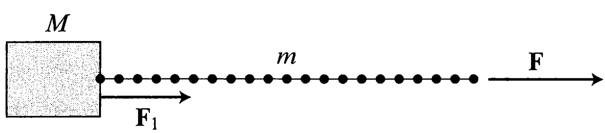
For a more detailed analysis of ropes and such, we certainly don't want to worry about the details of how ropes are attached to other objects, either by

the friction of knots, or some sort of glue. Instead, we'll simply attach a rope to



another object by making an endpoint of the rope a particle on the surface of that object.

Our rope, or other filament, may be regarded as the limiting case of a very large collection of very small rigid rods that are linked together by hinges. When a force $\mathbf{F}$ is applied to the free end of a filament, the filament will become "taut", arranging itself along a line in direction of $\mathbf{F}$, which is how we will always picture it. If a force $\mathbf{F}$ is applied to one end of a filament of mass m attached to an



object of mass M, then the filament and attached object have an acceleration a satisfying

$$\mathbf{F} = (M + m)a.$$

If $\mathbf{F}_1$ is the force that the last particle of the filament exerts on the object of mass M, then we also have (compare Problem 1-4)

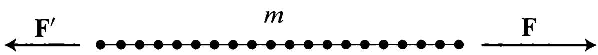
$$\mathbf{F}_1 = Ma,$$

and consequently

$$\mathbf{F}_1 = \frac{M}{M+m}\mathbf{F}.$$

Thus, the force transmitted to the object of mass M is less than our original force $\mathbf{F}$. Physicists like to consider the idealized case of a "massless" filament, with $m = 0$, which can be regarded as the limiting case when m is very small compared to M. In this special case of an idealized massless filament we will have $\mathbf{F}_1 = \mathbf{F}$.

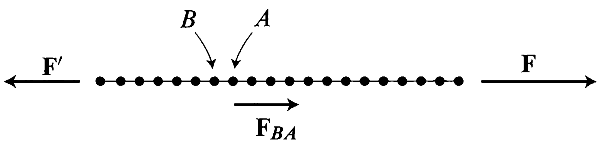
Consider a filament of mass m with a force $\mathbf{F}$ at one end and a force $\mathbf{F}'$ at



the other, and thus an acceleration a given by

(a) $$\mathbf{F} + \mathbf{F}' = ma.$$

Although the internal forces of a real rope extend in all sorts of directions, for our idealized filament it's convenient to assume that each particle exerts a force only on the particles right next to it. If A and B are two adjacent particles,



and m_1 is the mass of the part of the filament to the left of A, then the force $\mathbf{F}_{BA}$ that A exerts on B satisfies

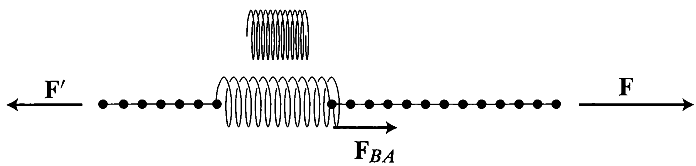
(b) $$\mathbf{F}_{BA} + \mathbf{F}' = m_1 a.$$

Thus,

(c) $$\begin{aligned} \mathbf{F}_{BA} &= m_1 a - \mathbf{F}' \\ &= \frac{m_1}{m}(\mathbf{F} + \mathbf{F}') - \mathbf{F}' \qquad \text{by (a)} \\ &= \frac{m_1}{m}\mathbf{F} + \left(\frac{m_1}{m} - 1\right)\mathbf{F}'. \end{aligned}$$

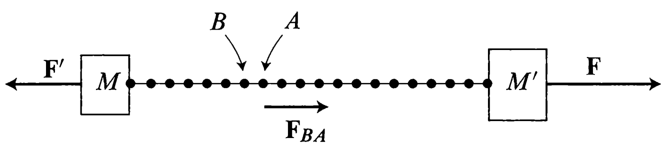
So the force varies in magnitude from $|\mathbf{F}|$, at the far end, where $m_1 = m$, to $|\mathbf{F}'|$ at the other end, where $m_1 = 0$. The magnitude of the force at any point of the filament is called its *tension* T at that point.

Notice that T is a number, not a vector: at any point A of the filament, it is the magnitude of the force $\mathbf{F}_{BA}$ that A exerts on the adjacent particle B, and thus the magnitude of the equal but opposite force $\mathbf{F}_{AB}$ that B exerts on A. This tension could be determined by inserting a spring into our filament at A, and measuring how much the spring is stretched.



In a situation like a tightened violin string we have $\mathbf{F}' = -\mathbf{F}$, with a non-accelerating string, and the tension T has the constant value $|\mathbf{F}|$. Of course, the tension can be increased only by stretching the violin string, so our idealized filament represents the usual sort of strange hybrid, where we pretend to be working with rigid bodies, or unstretchable filaments, even though the required internal forces can only arise from minute amounts of stretchability.

When the forces at the ends do not balance, so that our string is accelerating, the tension will not be constant. However, the idealized case of a massless filament will make sense when at least one end of our filament is attached to an object which has a non-negligible mass. If we enlarge our previous picture, of



a filament with forces at either end, to include objects of masses M and M' at the ends, then equation (c) is replaced by a rather more complicated relation,

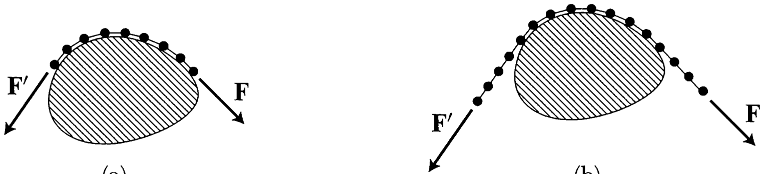
$$\mathbf{F}_{BA} = \frac{m_1 + M'}{m + M + M'}\mathbf{F} + \left(\frac{m_1 + M'}{m + M + M'} - 1\right)\mathbf{F}',$$

but when we take m, and thus m_1, to be very small, all $\mathbf{F}_{BA}$ have the constant value

$$\frac{M'}{M + M'}\mathbf{F} + \left(\frac{M'}{M + M'} - 1\right)\mathbf{F}' = \frac{M'}{M + M'}\mathbf{F} - \frac{M}{M + M'}\mathbf{F}',$$

so that the tension will again be constant along the filament.

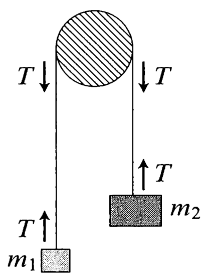
Now let us consider a filament that is wrapped partly around a fixed object, with forces **F** and **F**$'$ acting tangentially at the ends (a). Of course, in practice we



(a) (b)

would normally have some additional filament on each side (b), but we already know how to handle those situations. We can show that for the case of a massless filament with at least one end attached to an object of non-negligible mass, **F** and **F**$'$ will again have the same magnitudes, even though their directions may vanish; the analysis is just a more complicated version of the previous one, noting that any normal forces that arise in the analysis are balanced out by the normal forces that the object exerts on the various particles.

That is the basis for the usual analysis of the Atwood machine, if instead of thinking of a cord passing over a wheel, we imagine that the cord is sliding frictionlessly over a fixed rod (in practice we would have to grease the rod and/or the cord pretty heavily to obtain anything close to the theoretical ideal of frictionless motion). The object of mass m_2 has an acceleration a downward given



by

$$m_2 g - T = m_2 a,$$

while the object of mass m_1 has an acceleration $-a$ given by

$$m_1 g - T = m_1(-a),$$

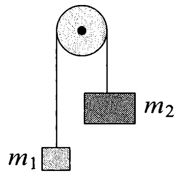
so

$$a = \frac{g(m_2 - m_1)}{m_1 + m_2},$$

as before, and we also find the tension $T = 2gm_1m_2/(m_1 + m_2)$.

Of course, in practice, we use a pulley in which the cords pass over a wheel, with nice low-friction ball-bearings, rather than a heavily greased cord sliding over a fixed rod. But the fixed rod picture actually gives the best representation of the usual elementary analysis of the Atwood machine, because this analysis conveniently assumes implicitly that the wheel over which the cord passes has negligible mass.

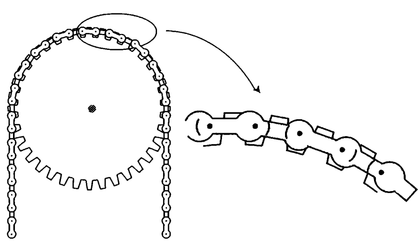
The case of an Atwood machine with a "massive" wheel (i.e., a wheel with non-negligible mass) is seldom discussed in physics textbooks, and those that



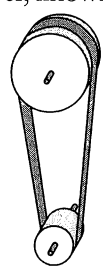
do discuss it engage in the usual maddeningly nonchalant assumption that any fool would know how to analyze it. In reality, however, the Atwood machine with a massive wheel represents something quite unlike anything else we have considered.

We are still going to ignore losses due to friction, but this just means that we are going to assume that the wheel rotates on its axis without friction. Unlike the case of a fixed rod, on which the cord slides without friction, we now want the motion of the cord to cause the wheel to turn. But that would seem to require the friction of the cord on the wheel, which we would also like to ignore! We appear to have something like the paradox of a rolling wheel, but with even greater complications: our wheel touches the cord on which it "rolls" along a whole stretch of cord, and the velocity of the points of the wheel are not 0 at these points.

A good representation of this apparently contradictory theoretical picture is given by an actual mechanism, the chain and sprocket wheel, known to everyone



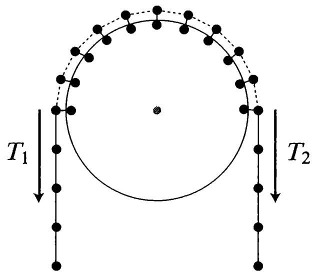
chain and sprocket wheel



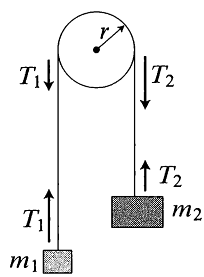
belt drive

who rides a bicycle. Industrially, a chain and sprocket wheel is used when a belt drive would be inadequate—precisely because the belt would have too much slippage.

As an idealized version, we can imagine that as the individual particles in our filament make contact with the wheel, they attach themselves to particles



in the wheel, dutifully detaching themselves when it is time to leave the path of the wheel. These particles just go along for the ride while they are in contact with the wheel, and the usual internal forces between them temporarily vanish. Consequently, at each moment we basically have two different cords attached to the wheel at opposite points, each of which can have its own tension.



For a wheel of radius r, the torque on the wheel is $rT_2 - rT_1$, so if α is the angular acceleration of the wheel, we have $rT_2 - rT_2 = \alpha I$, where I is the moment of inertia of the wheel around its center, and thus

$$T_2 - T_1 = (I/r)\alpha.$$

If a is the acceleration of the cord, we also have

$$\begin{aligned} gm_2 - T_2 &= m_2 a \\ T_1 - gm_1 &= m_1 a. \end{aligned}$$

Adding these three equations, we obtain

$$gm_2 - gm_1 = m_1 a + m_2 a + \frac{I}{r}\alpha.$$

But $r\alpha$ is the magnitude of the acceleration of a point on the circumference of the wheel, so we must have $a = r\alpha$, and thus

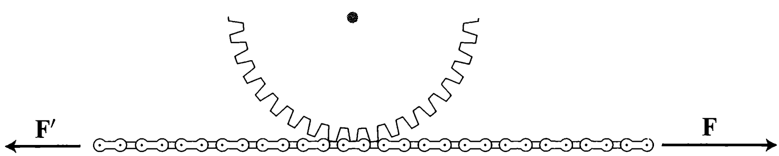
$$gm_2 - gm_1 = m_1 a + m_2 a + \frac{I}{r^2}a,$$

giving us

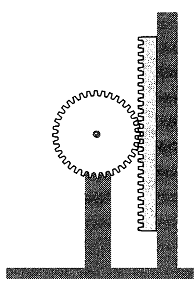
$$(*)\qquad a = \frac{g(m_2 - m_1)}{m_1 + m_2 + (I/r^2)}.$$

Since the total force $\mathbf{F}$ on our system is $g(m_2 - m_1)$, equation $(*)$ may be regarded as saying that the "effective total mass" of our system is $m_1 + m_2 + (I/r^2)$, so that the wheel has an "effective mass" of I/r^2, rather than I. This might have been expected, since motion through distance s for the end masses corresponds to a rotation of s/r radians for the wheel, and the torque equation $r\mathbf{F} = I \cdot \theta'' = I \cdot s''/r$ can be written as $\mathbf{F} = (I/r^2)s''$. It also suggests an easy way to find the solution (Problem 16), if we don't need to find the tensions.

Note, by the way, that even after we "straighten out" the problem in the same way that we originally treated the Atwood machine, we need to think in terms



of this chain and sprocket wheel picture if our rope causes a massive wheel to rotate as it is being pulled. A similar analysis is necessary if we have an object constrained to move along a track so that it causes a massive wheel to rotate as it falls.

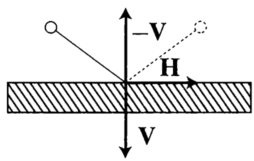


ADDENDUM 6A

THE BOUNCING SUPERBALL

In Chapter 3 we considered a ball bouncing off a flat surface to justify the term "perfectly elastic" for a collision preserving momentum and kinetic energy. The arguments at the beginning of this chapter would lead us to say that if a rigid ball is thrown straight at a flat surface (with effectively infinite mass), then the reasonable presumption is that it will simply bounce back with the same velocity.

For the case where the ball is thrown at the floor at an angle, it seems reasonable to argue that when the ball hits the floor, the horizontal component **H**



of its velocity is unchanged, while the vertical component **V** causes the wall to produce a force giving an opposite horizontal component of $-\mathbf{V}$, so that the ball bounces back according to old rule "angle of incidence equals angle of reflection".

But what happens when the ball is spinning? The common pink ball used in many children's games seems to follow the same rule, at least approximately, while a more sturdy tennis ball definitely shows some deviance from that rule. The SuperBall (see Problem 3-26) represents a fair approximation to both perfect elasticity and rigidity, and it bounces in a quite remarkable way.

Let h_0 and v_0 be the signed magnitudes of the initial components of the velocity of the ball (positive h_0 represents motion to the right, positive v_0 motion downwards). We will take the simple case where the ball is spinning about the axis perpendicular to the plane of the figure, and we will let ω_0 be its angular velocity about that axis. Similarly, h_1, v_1, and ω_1 will be the values after the ball bounces back.

We will assume, as before, that $v_1 = -v_0$. If I is the moment of inertia of the ball about its center, then conservation of kinetic energy gives

$$\tfrac{1}{2}I\omega_0{}^2 + \tfrac{1}{2}mh_0{}^2 = \tfrac{1}{2}I\omega_1{}^2 + \tfrac{1}{2}mh_1{}^2,$$

which can be written as

$$\text{(a)} \qquad m(h_1 - h_0)(h_1 + h_0) = -I(\omega_1 - \omega_0)(\omega_1 + \omega_0).$$

It also seems reasonable to assume that the floor is "perfectly rough", with conservation of angular momentum at the point of contact of the collision. This gives the equation

(b) $$ma(h_1 - h_0) = -I(\omega_1 - \omega_0).$$

We will ignore the solution $h_1 = h_0$, $\omega_1 = \omega_0$, since this would imply that there were no frictional forces from the floor to change the spin. So we can divide (b) into (a), to obtain $h_1 + h_0 = a(\omega_1 + \omega_0)$, and thus

(c) $$h_1 - a\omega_1 = -(h_0 - a\omega_0).$$

[Since $h - a\omega$ is the horizontal velocity at the point of contact, this says that it is exactly reversed at the bounce.]

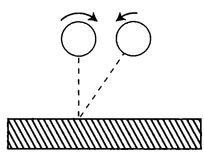
Solving (b) and (c) for h_1 and ω_1 in terms of h_0 and ω_0, and using $I = \frac{2}{5}ma^2$, we end up with

(*) $$\begin{aligned} h_1 &= \tfrac{3}{7}h_0 + \tfrac{4}{7}\omega_0 a \\ \omega_1 &= -\tfrac{3}{7}\omega_0 + \tfrac{10}{7}\frac{h_0}{a}. \end{aligned}$$

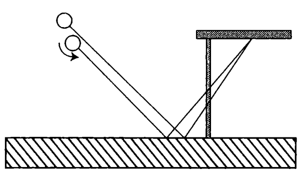
For $h_0 = 0$ we obtain

$$h_1 = \tfrac{4}{7}\omega_0 a, \qquad \omega_1 = -\tfrac{3}{7}\omega_0,$$

so if the ball is thrown vertically downwards with spin it bounces up at an angle, with a (smaller) reversed spin!



Similarly, if we the ball throw at an angle, but without spin, it will end up with spin. As a matter of fact, there is a SuperBall trick that involves bouncing the



ball under a table, without spin, and having it return, with a spin. If we start with $\omega_0 = 0$, then for the first bounce (*) gives

(B1) $$\begin{aligned} h_1 &= \tfrac{3}{7}h_0 \\ \omega_1 &= \tfrac{10}{7}\frac{h_0}{a}. \end{aligned}$$

For the second bounce, downwards, the angular velocities need to be calculated in the opposite direction, so we have to change the final + signs in each equation of (∗) to minus signs, so that we get

$$\text{(B2)}\qquad \begin{aligned} h_2 &= \tfrac{3}{7}h_1 - \tfrac{4}{7}\omega_1 a \\ \omega_2 &= -\tfrac{3}{7}\omega_1 - \tfrac{10}{7}\frac{h_1}{a} \end{aligned} \qquad \text{or} \qquad \begin{aligned} h_2 &= -\tfrac{31}{49}h_0 \\ \omega_2 &= -\tfrac{60}{49}\frac{h_0}{a}. \end{aligned}$$

Similarly, for the third bounce we obtain, finally,

$$\text{(B3)}\qquad \begin{aligned} h_3 &= -\tfrac{333}{343}h_0 \\ \omega_3 &= -\tfrac{130}{343}h_0, \end{aligned}$$

and the ball returns in practically the same direction, with slightly slower speed, but with the same total kinetic energy.

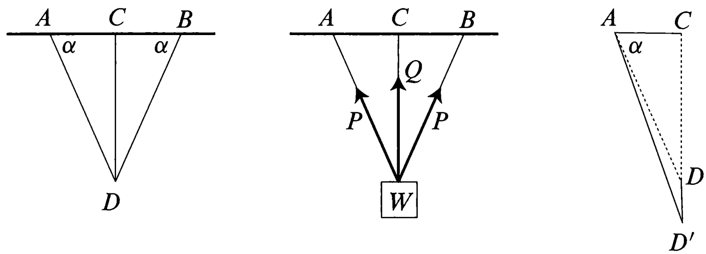
An analysis of this sort first appeared in Garwin [1], where some arguments, which I don't understand, are given to justify the various assumptions that we have made. Other justifications, which I also don't understand, may be found in Barger and Olsson [1].

ADDENDUM 6B

STATICALLY INDETERMINATE PROBLEMS

All statically indeterminate problems require some information about the elastic properties of materials, but in some cases the information is elementary and easy to apply.

For example, consider three filaments, of the same material, arranged as in (a), so that ABD is an isosceles triangle, and CD is its altitude. A weight W is hung



(a) (b) (c)

from the end, as in (b), so that the side filaments exert a force of magnitude P along their directions, while the middle filament exerts a force of magnitude Q. These forces actually come about because the filaments stretch slightly as in (c).

We assume that the filaments obey *Hooke's law*: If the filament is fixed at one end and the force $\mathbf{F}$ pulls on the other end, then

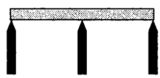
$$|\mathbf{F}| = \lambda \cdot \frac{\Delta l}{l} \qquad \text{for some constant } \lambda,$$

where l is the length of the unstretched filament, and Δl is the increase in length (this holds only for a certain interval of Δl values).

We could use geometry to find the length $AD' - AD$ of the left filament in terms of the change in length $\delta = DD'$, express both P and Q in terms of δ and the constant λ, and then use $2P + Q = W$ to find δ, leading to rather messy formulas for P and Q, in terms of λ. But it is much easier (Problem 22) to see what the limiting values are for $\varepsilon \to 0$, and these limiting values, which don't involve λ at all, give a reasonable answer for most filaments, like a piece of wire, where λ is very large, and ε very small.

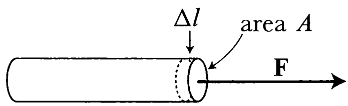
A more involved problem is encountered when we consider a plank resting on three identical supports. If the plank were perfectly rigid, the supports would

all have to be compressed by the same amount, and thus all have to provide an upward force exactly 1/3 of the weight of the plank. But we are interested instead in analyzing the effect of the plank's bending, which results in different compressions of the supports, and thus different upward forces. In fact, we aren't interested in the extremely tiny compression of the supports at all, only in the upward forces that they will have to provide to balance the bent plank. For simplicity we consider "knife edge" supports, which touch the plank along



a line, appearing as a single point in our 2-dimensional section.

Hooke's law also holds for a plank or rod, except that it is stated a bit differently, because we want to take into account the cross-section A, which was



essentially assumed to be 0 for the case of a filament. So we consider the ratios

$$\text{stress } \sigma = \frac{|\mathbf{F}|}{A}, \qquad\qquad \text{strain } \varepsilon = \frac{\Delta l}{l},$$

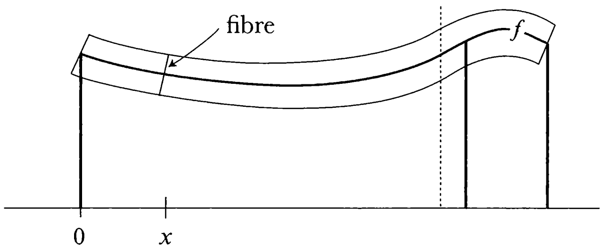
and write Hook's law in the form

$$\sigma = E\varepsilon$$

for a constant E, the *modulus of elasticity*. Of course, A actually changes a bit when the force $\mathbf{F}$ is applied, but the change is so minute that it is disregarded. It should also be noted that the modulus of elasticity for stretching might not be the same as that for compression (concrete is supposed to be an example), but we will be only consider the case where they are the same.

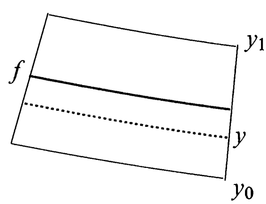
For steel, $E \approx 29 \times 10^6$psi (pounds per square inch), while for molybdenum, $E \approx 49 \times 10^6$psi. For wood E varies between $\approx .6 \times 10^6$psi and $\approx 1.7 \times 10^6$psi, depending on the type and grade of wood, the direction of the load, etc.

Now consider a long plank supported by various knife edge supports. The plank is actually going to sag a small amount, so that viewing it head on we see something like the picture shown below. To the left of the dotted line, there is a compressing stress along the top and a stretching stress along the bottom, while to the right of the dotted line just the opposite is true. The stress is thus 0



at some intermediate surface, the *neutral plane*, whose profile, shown as a heavy line in the figure, is the graph of some function f. The figure also shows a cross-section of the "fibre" through $(x, f(x))$, that is, the surface into which a vertical section of the plank is deformed.

The figure below is a greatly enlarged view of a small portion of the figure near the point $(x, f(x))$, bounded by two fibres. While the heavy line is the

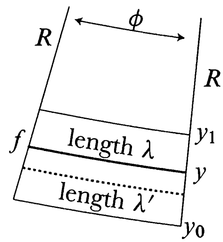


graph of f, the cross-section of the neutral plane, the dotted line indicates the cross-section of a surface where the stress has the constant value y. The whole region is filled up by such surfaces having values in some interval (y_0, y_1).

Let the curvature of the graph of f at the point $(x, f(x))$ be $\kappa(x)$, where (see page 62) we have

$$\kappa(x) = \frac{f''(x)}{1 + (f'(x))^2}.$$

Near the point $(x, f(x))$, the graph is very close to a segment of a circle subtending an angle ϕ with radius $R = 1/(-\kappa)$, where the minus sign is necessary because $f'' \leq 0$ at this point. So the solid and dotted lines have lengths λ and λ'



given, to first order, by

$$\lambda = R\phi$$
$$\lambda' = (R + y)\phi.$$

Consequently, the strain ε along the surface indicated by the dotted line is given by

$$\varepsilon = \frac{\lambda' - \lambda}{\lambda} = \frac{y\phi}{R\phi} = \frac{y}{R} = -y\kappa,$$

and the stress along this surface is

$$\sigma = \varepsilon E = -y\kappa E.$$

It is easy to check that for points where $f'' \geq 0$ we get exactly the same formula.

For the fibre A through the point $(x, f(x))$, let $\boldsymbol{\tau}(x)$ be the total torque on A, with respect to the point $(x, f(x))$, from all the external forces to the left of the point (gravity acting down on the portion of the board to the left, together with the upward force of any supports to the left). Since A isn't rotating, we must have the following, where b is the width of the plank:

$$\begin{aligned}\boldsymbol{\tau}(x) &= \int_A \sigma \\ &= b \int_{y_0}^{y_1} \sigma(x, y) \cdot y \, dy \\ &= -E\kappa(x) \int_{y_0}^{y_1} by^2 \, dy \\ &= -EI\kappa(x),\end{aligned}$$

where

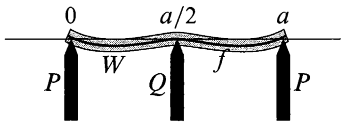
$$I = \int_{y_0}^{y_1} by^2 \, dy$$

is, by definition, the moment of inertia of A, which we will assume can be taken to be a constant.

This gives us the equation $\tau(x) = -EI\kappa(x)$. Finally, since f' is usually going to be extremely small, we simply throw away the $f'(x)$ term in the expression for $\kappa(x)$, leading to the *Euler-Bernoulli equation* for thin plank bending,

$$(*) \qquad -EIf''(x) = \tau(x).$$

As a simple application of the Euler-Bernoulli equation, consider a plank of length a resting on three knife edge supports, two at the ends, and one in the



middle. The plank, of weight W, is assumed to have uniform density $w = W/a$; the outside supports each exert an upward force of P and the middle support exerts an upward force of Q, with $2P + Q = W$. For convenience, we choose the position of the x-axis so that our function f is 0 at 0, $a/2$, and a.

For $0 \le x < a/2$ we have

$$\tau(x) = -Px + \tfrac{1}{2}wx^2,$$

where the first term is the moment of the upward force P at distance x from our point, and the second term comes from the uniformly distributed force of w along the plank of length x to the left of our point. Thus

$$EIf''(x) = Px - \tfrac{1}{2}wx^2$$

and

$$EIf'(x) = \frac{Px^2}{2} - \frac{wx^3}{6} + C_1.$$

There is another equation for $a/2 < x \le a$, involving another constant C_2, but in this case we can use symmetry to dispense with the second expression. Since we clearly have $f'(a/2) = 0$, we can immediately solve for C_1, to get

$$EIf'(x) = \frac{Px^2}{2} - \frac{wx^3}{6} + \frac{wa^3}{48} - \frac{Pa^2}{8},$$

and since $f(0) = 0$ this gives

$$EIf(x) = \frac{Px^3}{6} - \frac{wx^4}{24} + \left(\frac{wa^3}{48} - \frac{Pa^2}{8}\right)x.$$

Finally, using $f(a/2) = 0$, and remembering that $aw = W$, this gives $P = \frac{3}{16}W$. So each end provides an upward force of $\frac{3}{16}W$, while the middle support bears most of the weight, providing an upward force of $\frac{10}{16}W$.

By the way, from symmetry, on the intervals $0 \le x < a/2$ and $a/2 < x \le a$ we obviously have

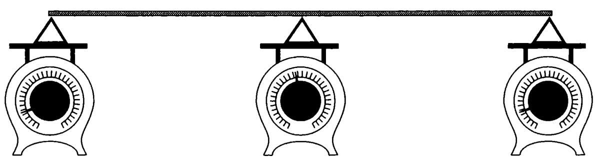
$$EIf''(x) = \begin{cases} Px - \frac{1}{2}wx^2 \\ P(a-x) - \frac{1}{2}w(a-x)^2 \end{cases} \implies EIf'''(x) = \begin{cases} P - wx \\ -P + w(a-x). \end{cases}$$

Thus at $a/2$ there is a jump discontinuity in $EIf''' = -\tau'$ of

$$2\left(P - \frac{wa}{2}\right) = 2P - W = -Q,$$

due to the upward force Q concentrated at $a/2$.

This simple example, as well as several others, comes from the delightfully old-fashioned book Synge and Griffith [1], which does not eschew "engineering" type problems. There are, of course, many subtleties that have been overlooked in this brief description, and modern books often don't even mention the equation explicitly, because it is basically a simplification (Hooke's law itself is basically just a simplification). Novices can easily be misled, as I discovered when I tried a a simple "home-lab" experiment, using three scales and a brass strip. Though it gives a result vaguely close to the theoretical one when the bass



strip is about 3 feet long, for shorter lengths the results are dramatically wrong, with the reading on the middle scale *lower* than the readings at the ends. It took me a long time to realize that this is because the results are reliable only for *thin* beams, ones whose thickness is small compared to their length.

Although you might not even find the name Bernoulli in a modern mechanical engineering book, you probably will find Euler's name, though usually in connection with his theory of column buckling, which is a sort of vertical version of plank bending. For an interesting "home-lab" experiment that one can do concerning Euler buckling, see Casey [1].

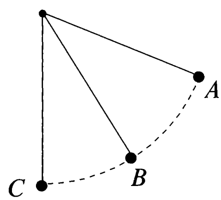
More sophisticated analyses of all these matters are to be found in studies of elasticity. For mathematicians interested in exploring this subject, I suspect that a good place to start would be Marsden and Hughes [1].

PROBLEMS

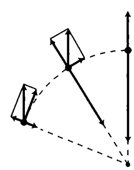
1. Consider the iron, wood, magnet combination on page 206, except that the iron and wood are in contact, but separated from the magnet, towards which they are accelerating. If **F** is the force that the magnet exerts on the iron, of

mass M, what is the force that the iron is exerting on the wood, of mass m? If instead the wood is in contact with the magnet, of mass M', what force is the magnet exerting on the wood? What happens at the moment that the wood comes in contact with both the iron and the magnet?

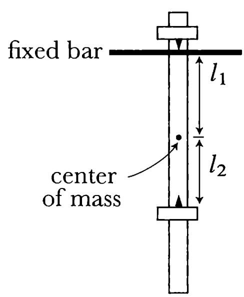
2.[1] For the three positions A, B, C of a pendulum bob released from position A, draw the four vectors shown in the figure on page 210. The vectors needn't be drawn precisely to scale, but the relative sizes should be clearly indicated.



Answer:



3. Measuring g with a pendulum bob on a string is very inexact because the string length is difficult to measure very precisely. A physical pendulum avoids this problem, but introduces a new problem, the difficulty of finding the center of mass, or equivalently the radius of gyration k. The Kater pendulum cleverly circumvents this problem. In essence, the pendulum consists of a rod on which



there are two knife edges pointing in opposite directions, either of which can be placed on a fixed bar and used as the pivot point of a pendulum.

[1] Cf. agm.cat/recerca-divulgacio/pendulum-TPT.pdf.

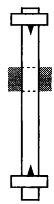
(a) If l_1 and l_2 are the distances from the center of mass to the two knife edges, then for small oscillations the periods T_1 and T_2 are given by

$$T_i{}^2 = 4\pi^2\left(\frac{k^2 + l_i{}^2}{gl_i}\right).$$

(b) If the positions of the knife edges are movable, and adjusted so that $T_1 = T_2 = T$, then $k^2 = l_1 l_2$ (*provided that* $l_1 \neq l_2$). So

$$g = 4\pi^2\left(\frac{l_1 + l_2}{T^2}\right) = 4\pi^2(L/T^2),$$

where $L = l_1 + l_2$ can be measured very accurately without having to measure l_1 and l_2. In practice, it is much better to keep the delicate knife edges stationary, and add another adjustable weight.



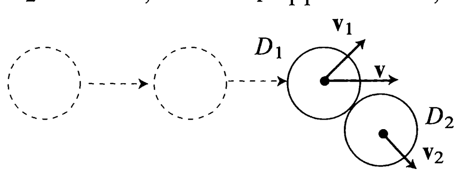
(c) We cannot expect to get $T_1 = T_2$ exactly, no matter how obsessively we measure, adjust weights, remeasure, ... , so we need to know how the inaccuracies affect the measurement for g. Show that

$$\frac{4\pi^2}{g} = \frac{l_1 T_1{}^2 - l_2 T_2{}^2}{l_1{}^2 - l_2{}^2} = \frac{T_1{}^2 + T_2{}^2}{2(l_1 + l_2)} + \frac{T_1{}^2 - T_2{}^2}{2(l_1 - l_2)},$$

so when $T_1 \approx T_2$ we get a good value when we know $l_1 + l_2$ very accurately, even though $l_1 - l_2$ might be known with much less accuracy.

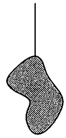
Bessel was one of the first to use these ideas to obtain very accurate measurements of g and to demonstrate the proportionality of mass and weight mentioned on page 38.

423 4. Let D_1 and D_2 be homogeneous 3-balls of the same mass, sliding on a frictionless surface. D_2 is at rest, while D_1 approaches it, not directly head on,

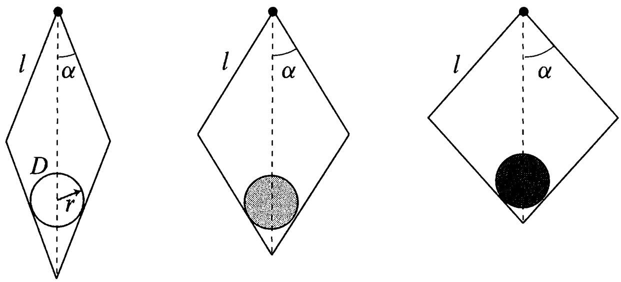


with velocity $\mathbf{v}$. We assume that the collision is perfectly elastic, and also that the surfaces of the balls are perfectly smooth, so that the collision will not cause any rotation. Using Problem 3-12 and the assumption that the restraint forces at the point of collision are perpendicular to the surfaces of the balls, find the resultant velocities $\mathbf{v}_1$ and $\mathbf{v}_2$.

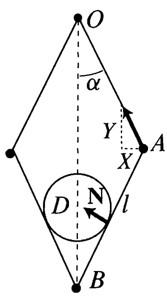
5. An object hangs in equilibrium from a thread. Show that the line of the thread passes through the center of gravity of the object.



6. Consider a rhomboid made of four rigid rods hinged at their ends, each of length l and weight w. The rhomboid is suspended from a fixed point, and a disc D of radius r and weight W is placed within this rhomboid. The weight of the rhomboid tends to make α smaller, while the weight of the disc tends to make α larger. The problem is to find the angle α at equilibrium. As usual, we are assuming that there is no friction between the disc and the rods.



(a) Let X and Y be the components of the force that AB exerts on AO, and



let $N = |\mathbf{N}|$ be the length of the reaction force $\mathbf{N}$ that AB exerts on the disc. Establish the equations

(O) $$l(X \sin\alpha - Y \cos\alpha) + \tfrac{1}{2} l w \sin\alpha = 0$$

(B) $$l(X \sin\alpha + Y \cos\alpha) - rN \cot\alpha - \tfrac{1}{2} l w \sin\alpha = 0$$

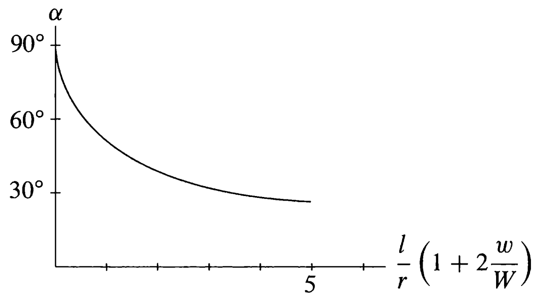
(AB) $$-X + N \sin\alpha + w = 0$$

(D) $$W - 2N \sin\alpha = 0$$

(hint: the equation labels indicate the objects on which to compute forces, or the points about which to compute torques), and use them to deduce that

$$(*) \qquad \cot^3\alpha + \cot\alpha - 2\frac{l}{r}\left(1 + 2\frac{w}{W}\right) = 0,$$

which has only one real root for $\cot\alpha$. The diagram below shows how α varies with $(l/r)[1 + 2(w/W)]$.



(b) Noting that the vertical diameter of the rhomboid has length $2l\cos\alpha$, so that the distances from the fixed point to the center of gravity of the rhomboid and to the center of the disc are

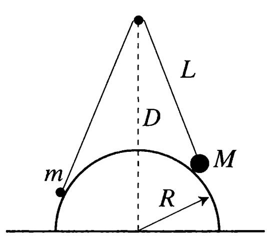
$$l\cos\alpha \quad \text{and} \quad 2l\cos\alpha - \frac{r}{\sin\alpha},$$

show that $V(\alpha)$, the potential energy of the system for α, is

$$V(\alpha) = C - 4wl\cos\alpha - W\left(2l\cos\alpha - \frac{r}{\sin\alpha}\right)$$

for a constant C. Then use $V'(\alpha) = 0$ to obtain $(*)$.

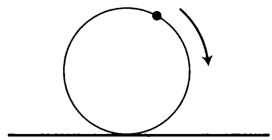
7. A fixed peg at distance D above the floor is centered above a fixed semicircular track of radius R. A string of total length L passing over the peg is attached



to objects of mass m and M; for convenience, we assume that the dimensions of the peg and these objects is negligible. Find the equilibrium position.

8. (a) For the wheel analysed on pages 224–226, starting at rest at height h from the floor, compute the speed and angular velocity at the bottom of the inclined plane, and then the rotational energy T_{rot}, and conclude that the total energy is again Mgh.
(b) Conversely, use conservation of energy to determine the motion of the wheel.

9. A classical paradox involves a rolling rigid weightless hoop with a point mass on it (one has to pretend that this is not already a paradox). When the hoop

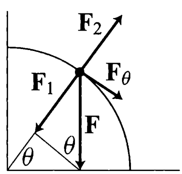


is in the position shown, the potential energy, as well as the kinetic energy, is positive. When the point mass reaches the bottom, its potential energy is 0, and so is its kinetic energy, since the velocity of the contact point is 0, contradicting conservation of energy. Explain.

Mathematically, what is the tacit assumption in Proposition 1? Physically, what happens to the forces between the particles of the hoop as the point mass approaches the bottom?

10. Analyze the motion of a cylinder rolling down a wedge, which can slide without friction on the floor.

11. Sliding particle. Consider a particle of mass m falling from height h along a frictionless circular path of radius l. Let $\mathbf{F}$ be the downward force due to



gravity, of magnitude mg, and $\mathbf{F}_1$ the component perpendicular to the path. Then the path exerts a force $\mathbf{F}_2 = -\mathbf{F}_1$ on the particle, so that the total force $\mathbf{F}_\theta$ is $\mathbf{F} - \mathbf{F}_1$.

(a) Conclude that the acceleration of the particle, tangent to the circle, has magnitude $a_\theta = g\cos\theta$, so that

$$\theta'' + \frac{g}{l}\cos\theta = 0. \tag{a}$$

Notice that this problem is essentially the same as the pendulum problem, except

that the angle ϕ for the pendulum is measured from the lowest point, so that $\phi = \pi/2 + \theta$, and the equation $\phi'' + (g/l)\sin\phi = 0$ is equivalent to (a).
(b) One of the classical elementary mechanics problems is to determine when the particle loses contact with the path, a fine example of a physics problem where the main difficulty is figuring out what the problem is actually saying, and, for good measure, where the answer given is often incomplete.

We need to determine when the total normal force vanishes. The inward acceleration is v^2/l (Problem 1-5), while the outward force $\mathbf{F}_2$ gives an acceleration of magnitude $g\sin\theta$. So we need

$$g\sin\theta = \frac{v^2}{l}$$

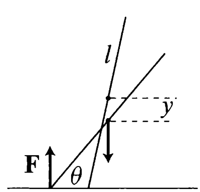
(compare page 210). Using conservation of energy, show that if the particle starts at rest ($v = 0$) at height h, this happens when the height $l\sin\theta$ is $2h/3$.
(c) Also solve the problem by multiplying (a) by θ' to show that

$$\frac{\theta'^2}{2} + \frac{g}{l}\sin\theta = \text{constant},$$

and conclude that if the particle starts at rest at θ_0, we have $\sin\theta = 2\sin\theta_0/3$. (Of course, this trick is basically the one we used to obtain conservation of energy in the first place.)

We have actually only shown that the total normal force vanishes at $2/3$ of the original height. To show that the particle actually leaves the path, note that the formula for the total normal force would give a negative value past this point if the particle stayed on the path (if our particle were a small loop sliding along a circular wire, this would mean that it is pulling on the wire instead of being pushed by the wire).

12. Sliding stick I. Consider a stick of mass m and length $2l$ sliding along a frictionless floor as it falls. By Problem 5-6, the moment of inertia of the stick



is $ml^2/3$. Since the only forces are gravity acting downward and a reaction force $\mathbf{F}$ acting upward at the point where the stick hits the plane, the center of mass must fall straight down.

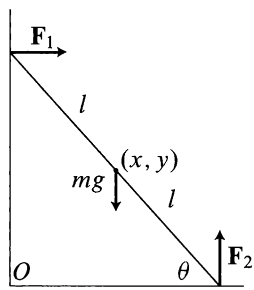
(a) Show that the total energy after the center of mass has fallen distance y is

$$\frac{my'}{2} + \frac{ml^2}{6}\theta'^2 + mg(l - y)$$

and conclude that the velocity of the center of mass after falling distance y is

$$y' = \sqrt{\frac{6gy\sin^2\theta}{3\sin^2\theta + 1}}.$$

13. Sliding stick II. Consider a stick of length $2l$ that is sliding between a frictionless wall and a frictionless floor. In addition to the downward force of



magnitude mg, there are forces $\mathbf{F}_1$ of magnitude F_1 and $\mathbf{F}_2$ of magnitude F_2 acting horizontally and vertically at the ends of the stick. This looks even worse than the previous problem, but we can get much more interesting information.

(a) We have

$$\begin{aligned} mx'' &= F_1 \\ my'' &= F_2 - g \\ \frac{ml^2}{3}\theta'' &= lF_1\sin\theta - lF_2\cos\theta, \end{aligned}$$

and thus

$$\frac{ml^2}{3}\theta'' = ml\sin\theta x'' - ml\cos\theta y'' - mlg\cos\theta.$$

(b) Differentiate the equations $x = l\cos\theta$ and $y = l\sin\theta$ to obtain expressions for x'' and y'', and reduce the above equation to

$$\theta'' + \frac{3g}{4l}\cos\theta = 0.$$

(c) Use the trick in Problem 11 (c) to show that if the stick starts at rest at angle θ_0, then

$$\theta'^2 = \frac{3g}{2l}(\sin\theta_0 - \sin\theta).$$

Conclude that the stick loses contact with the wall ($0 = F_1 = mx''$) when

$$3\sin\theta = 2\sin\theta_0,$$

and thus when its top point reaches $2/3$ of its initial height.

As in Problem 11, the formula for F_1 would give a negative value beyond this point, showing that the stick really leaves the wall, and we could also use conservation of energy.[1]

Comparison between Problems 11 and 13. Note that the line segment from O to (x, y) also has length l, so

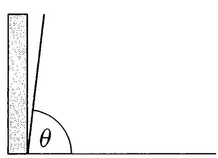
(i) The center of mass (x, y) moves along a circle of radius l,

(ii) The angle between the horizontal axis and the line from O to (x, y) is also θ.

Thus the equation in part (b) and equation (a) of Problem 11 differ only by a factor of $3/4$, and the center of mass of the sliding stick moves $\sqrt{3}/2$ times as fast as the particle sliding down the circle of radius l.

After the fall. Note that when the stick hits the floor we will have $x' > 0$, since whenever $\mathbf{F}_1 \neq 0$ it is always pointing away from the wall. Consequently, the stick will continue to move away from the wall, until stopped by friction.

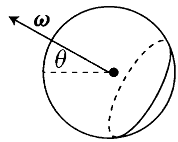
If we consider a stick starting nearly vertical and rotating in the other direction, or for a convenient experiment, a ruler placed next to a book standing upright on a desk, we end up with the same equations, and now $x' > 0$ when



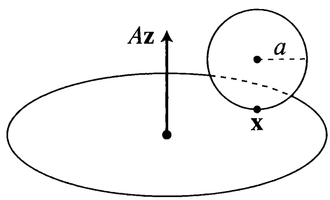
the ruler hits the desk—the ruler cannot simply pivot around the point where the book and desk meet, but must end up to the right of the book.

[1] Parts of this problem come from the venerable book Osgood [1], which, like Synge and Griffith [1], also contains material off the beaten track.

14. For the constant $\boldsymbol{\omega}$ of a rolling sphere we can choose our first axis so that $\boldsymbol{\omega} = (0, \omega\cos\theta, \omega\sin\theta)$ for some θ. Use our equation $\mathbf{v} = -R\mathbf{z} \times \boldsymbol{\omega}$ to show that as the sphere rolls along a straight line in the plane, the path traced out on the sphere is a circle tilted at angle $\pi/2 - \theta$ to the horizontal.



15. Consider a sphere of mass m and radius a rolling on a rotating turntable with constant angular momentum $A\mathbf{z}$.



(a) If $\mathbf{x}(t)$ is the contact point at time t, then the velocity $\mathbf{v}$ of the center of the ball is

$$\mathbf{v} = A\mathbf{z} - a\mathbf{z} \times \boldsymbol{\omega}$$
$$\mathbf{v}' = A\mathbf{z} - a\mathbf{z} \times \boldsymbol{\omega}'.$$

(b) As before, $\boldsymbol{\tau} = -a\mathbf{z} \times \mathbf{F} = -a\mathbf{z} \times m\mathbf{v}'$, so

$$\boldsymbol{\omega}' = -\frac{am}{I} \cdot \mathbf{z} \times \mathbf{v}',$$

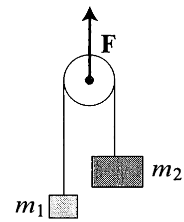
and thus

$$\mathbf{v}' = \frac{A}{1 + ma^2/I} \cdot \mathbf{z} \times \mathbf{v}$$
$$= \frac{2}{7} A \cdot \mathbf{z} \times \mathbf{v}.$$

(c) The ball travels in a circle with a frequency $2/7$ of the rotation frequence A of the turntable.

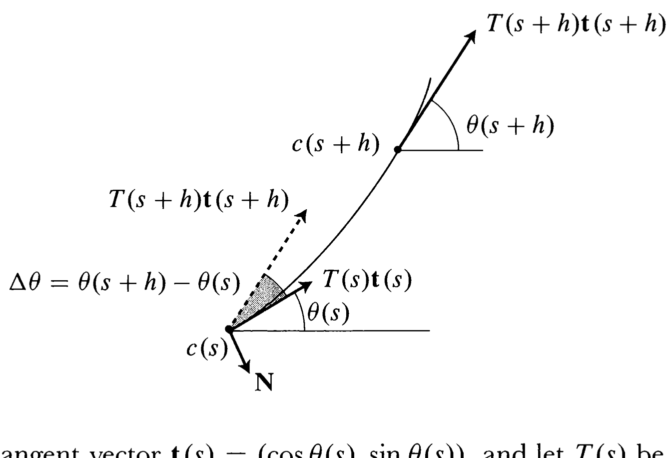
16. Find the acceleration for an Atwood machine with a massive pulley having moment of inertia I by using conservation of energy.

17. Analyze the motion, including the tensions of the rope, for an Atwood machine with a massive pulley having moment of inertia I when the axle of the pulley wheel is free to move and a constant force $\mathbf{F}$ is applied to it. Here $\mathbf{F}$



denotes the total force, i.e., the applied force minus the downward force of gravity on the pulley wheel.

18. (a) Consider a massless filament wrapped around a fixed object, in equilibrium. Let the profile curve of the object be an arclength parameterized curve c,



with unit tangent vector $\mathbf{t}(s) = (\cos\theta(s), \sin\theta(s))$, and let $T(s)$ be the tension at $c(s)$. For small h, the segment of the filament between $c(s)$ and $c(s+h)$ is practically a straight line of length h so that the normal force along this segment has magnitude close to $|\mathbf{N}|h$, where $\mathbf{N}$ is the normal force exerted by the fixed object on the filament at $c(s)$. Since the total normal force at $c(s)$ must be 0, conclude that

$$T(s+h)\sin\Delta\theta = |\mathbf{N}|h + o(h)$$

(things can be made more precise by using the Mean Value Theorem, for those so inclined) and then that

$$|\mathbf{N}| = T(s)\theta'(s) = T(s)\kappa(s),$$

where κ is the curvature.

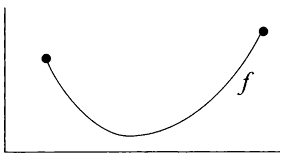
Of course, the filament exerts an equal and opposite force on the object, and this force is greater at points of greater curvature, so a string wrapped around a parcel bites in deeper at the edges of the parcel.

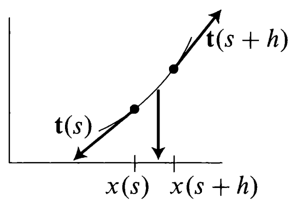
(b) Let $\Delta T = T(s+h) - T(s)$. Since the tangential force at $c(s)$ should also be 0, conclude that

$$(T(s) + \Delta T)\cos\Delta\theta = T(s),$$

and then that $T' = 0$.

19. Consider a cable anchored at two points, and hanging under the influence of its own weight, plus possibly other weights attached to it. We will let the





function w be the density of this load, the weight per unit length. The cable lies along the graph of a function f, and we will let $T(x)$ be the tension of the cable at $(x, f(x))$. Let $s \mapsto (x(s), y(s))$ be the arclength parameterization of the graph, so that $x'(s)$ and $y'(s)$ are the cosine and sine of the unit tangent vector $\mathbf{t}(s)$ at $(x(s), y(x))$.

(a) The piece of the cable between $(x(s), y(s))$ and $(x(s+h), y(s+h))$ is acted upon by tension forces at the ends, and forces due to the weight, the later of which are all directed straight¡ down. Conclude that the *horizontal component* of $T \cdot \mathbf{t}$ is constant, that is,

(1) $$T\frac{dx}{ds} = T_h \qquad T_h \text{ constant.}$$

(b) Conclude also that if Δl is the length of the cable between these two points, then

$$T(x(s+h))y'(s+h) - T(x(s))y'(s) = w(\xi)\cdot\Delta l$$

for some ξ with $x(s) \le \xi \le x(s+h)$, which implies that

(2) $$\frac{d}{ds}\left(T\frac{dy}{ds}\right) = w,$$

and thus finally

(*) $$\frac{d}{ds}\left(\frac{dy}{dx}\right) = \frac{w}{T_h}.$$

(c) Consider a suspension bridge, where the cable is regarded as weightless, while the load, the road the cable supports, has a uniform weight w_0 *per unit horizontal length*, which means that

$$w = w_0\frac{dx}{ds}.$$

Conclude from (*) that

$$\frac{d^2y}{dx^2} = \frac{w_0}{T_h},$$

so that the graph of f has the shape of the parabola

$$x \mapsto \frac{w_0}{2T_h}x^2.$$

(d) Using $x'(s)^2 + y'(s)^2 = 1$, conclude that

(3) $$\frac{ds}{dx} = \sqrt{1+\left(\frac{dy}{dx}\right)^2},$$

and thus by (1)

$$T = T_h\sqrt{1+\left(\frac{dy}{dx}\right)^2}.$$

If the bridge spans the interval $[-S, S]$ and has height H at these ends, we have $T_h = w_0S^2/2H$, and the maximum of dy/dx on the interval is w_0S/T_h, so the maximum of T on the interval turns out to be

$$Sw_0\sqrt{\frac{1}{4H^2}+1}.$$

So we have to make H large to keep the maximum tension small enough to insure that the cable doesn't break.

20. Now consider a cable of uniform density w per unit length, which is hanging freely under its own weight, with no additional weights. Then $(*)$ and (3) of the previous problem give

$$y'' = \frac{w}{T_h}\sqrt{1+y'^2} = \frac{1}{c}\sqrt{1+y'^2} \qquad\qquad c = \frac{T_h}{w}$$

(where the primes now indicate d/dx).

This can be solved in various ways. Familiarity with the hyperbolic functions would suggest writing

$$f' = \sinh \circ v,$$

since this reduces the equation to

$$\cosh \circ v \cdot v' = \frac{1}{c}\sqrt{1+\sinh^2 \circ v} = \frac{1}{c}\cosh \circ v,$$

or simply

$$v(x) = \frac{1}{c}x + \text{constant}.$$

It is convenient to choose the lowest point of the cable at the origin, so that $f'(0) = 0 \implies v(0) = 0$, so that we have

$$f'(x) = \sinh\frac{x}{c} \implies f(x) = c\cosh\frac{x}{c} - c,$$

a catenary. Also find the tension T.

Another way to solve the equation, written in classical notation as $d^2y/dx^2 = (1/c)\sqrt{1+(dy/dx)^2}$, is to use the substitution $p = dy/dx \implies dp/dx = d^2y/dx^2$, leading to

$$\frac{dp}{dx} = \frac{1}{c}\sqrt{1+p^2} \implies dx = \frac{k\,dp}{\sqrt{1+p^2}} \implies x = c\sinh^{-1}(p) + C$$

for a constant C and thus to $dy/dx = p = \sinh([x-C]/c)$, a trick apparently introduced by Jacopo Francesco Riccati (1676–1754) in 1712.

21. In the previous problem we essentially found an equilibrium position for the cable, so we naturally wonder whether we can also find a solution by looking at the minimum point for the "potential function" of y,

$$V(y) = \int_a^b wy\,ds = \int_a^b wy\sqrt{1+y'^2}\,dx.$$

The trick here is that we need to find the minimum under the constraint that the length of the cable is fixed,

$$L(y) = \int_a^b \sqrt{1 + (y')^2}\,dx = \text{constant}.$$

As explained in Addendum 13A, such problems can be solved, analogously to Problem 5-2 (c), by finding the Euler equations for the calculus of variations problem for $V + \lambda L$ for some "Lagrange multiplier" λ. So for

$$F(x, y, y') = wy\sqrt{1 + y'^2} + \lambda\sqrt{1 + y'^2}$$

we have the Euler equations

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 0$$

where $\partial F/\partial y$ and $\partial F/\partial y'$ simply denote the derivatives of F with respect to its second and third arguments, and the terms in the equation are evaluated at $(x, f(x), f'(x))$.

(a) In this case, where F doesn't depend on x, we have "Beltrami's identity"

$$\frac{d}{dx}\left(F - f'\frac{\partial F}{\partial y'}\right) = f'\left[\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'}\right],$$

leading to the simpler equation

$$F - f'\frac{\partial F}{\partial y'} = C$$

for a constant C.

(b) For our F, deduce that

$$y'^2 = \frac{(wy + \lambda)^2 - C^2}{C^2}$$

and thus

$$wy + \lambda = C\sqrt{1 + y'^2}.$$

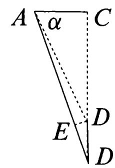
(c) By differentiating the first of these equations deduce finally that

$$y'' = \frac{w}{C}\sqrt{1 + y'^2},$$

basically the same equation as on page 270.

An interesting discussion of these and related problems, with historical side lights, may be found in Nahin [1 Chap. 5, pp. 240–251].

22. In the figure below, the left part of the figure on page 252, the point E is chosen so that $AE = AD$.



(a) Show that

$$\frac{P}{Q} = \frac{ED'}{DD'} \cdot \frac{CD}{AD} = \frac{ED'}{DD'} \sin\alpha.$$

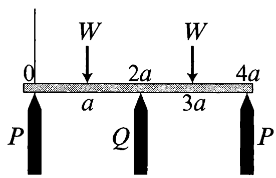
(b) For small DD', the line DE is practically perpendicular to AD. Conclude that, in the limit,

$$\frac{P}{Q} = \sin^2\alpha,$$

and thus

$$P = \frac{\sin^2\alpha}{1 + 2\sin^3\alpha} W \qquad \text{and} \qquad Q = \frac{1}{1 + 2\sin^2\alpha} W.$$

23. Consider a plank of length $4a$ resting on three knife edge supports, two at



the ends, and one in the middle, except that now the mass of the plank itself is negligible, while there are forces of magnitude W exerted at the middle of each section.

(a) We have

$$EIf''(x) = \begin{cases} Px & 0 \le x < a \\ Px - W(x-a) = (P-W)x + Wa & a < x < 2a. \end{cases}$$

(b) Using the fact that $f'(a) = 0$ and $f(0) = 0$, show that

$$f(x) = \frac{Px^3}{6} + \left(\frac{M}{2} - 2P\right) a^2 x \qquad \text{on } (0, a).$$

(c) Using the fact that $f'(2a) = 0$ and $f(2a) = 0$, show that

$$f(x) = \frac{Px^3}{6} - \frac{Wx^3}{6} + \frac{Wax^2}{2} - 2Pa^2x + a^3\left[\frac{8P}{3} - \frac{2M}{3}\right] \qquad \text{on } (a, 2a).$$

(d) Setting these expressions equal at a, conclude that $P = \frac{5}{16} W$, and $Q = \frac{22}{16} W$.

CHAPTER 7

PHILOSOPHICAL AND HISTORICAL QUESTIONS

This chapter contains several remarks concerning philosophical questions, some of which will be relevant to Chapter 10, together with a few tidbits of an historical nature.

Early notions of conservation of momentum. The first statement of the law of conservation of momentum might be attributed to Descartes, who asserted in his *Principles*, 1644, that "God in his omnipotence has created matter together with the motion and the rest of its parts, and with his day-to-day interference, he keeps as much motion and rest in the Universe now as he put there when he created it ... " (Dugas [1; pg. 161]). Arguments of this sort might not get a very favorable reception in modern physics journals, but the real problem with Descartes' formulation was that he confused mv with $m|v|$ (since every one was thinking of collisions along a straight line at the time, the scalar speed, rather than the vector velocity, was the quantity at issue). Partly as a result of this, almost all the rules of impact formulated by Descartes are simply wrong, a circumstance that Descartes seems inclined to dismiss as due to experimental error (see Dugas [1; pg. 163]).

It is therefore not surprising that in 1668 the Royal Society proposed a discussion on the subject of the laws of colliding bodies, the impetus for the investigations of Wren, Wallis, and Huygens mentioned on page 22, which Newton reformulated as the third law. Note that when Newton derives conservation of momentum from this law (page 22), it is carefully stated so that Descartes' error is corrected—quantity of motion "is determined by adding the motions made in one direction and subtracting the motions made in the opposite direction".

Huygens and Galilean Invariance. On page 26ff. we were able to present Huygens' argument in just a few lines. But Huygens himself gave a much more detailed argument, which appeared in his book *De Motu Corporum ex Percussione*, published posthumously in 1700.

Huygens begins by first explicitly stating three hypotheses that he will use in his argument:

Hypothesis I was basically the law of inertia (Newton's first law).

Hypothesis II was basically our assumption, on the basis of symmetry, that identical bodies moving toward each other with equal speeds must rebound with equal speeds. (Huygens' hypothesis was actually stronger, and his investigations were quite a bit more complicated—see Problem 3-9 for references.)

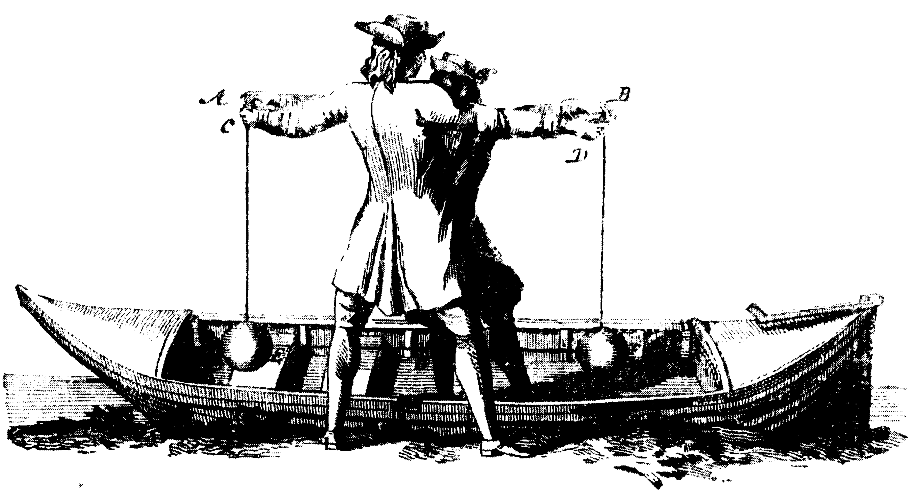
Hypothesis III stated (see Dugas [1; pg. 176]):

> The expressions 'motion of bodies' and 'equal or unequal velocities' should be understood relatively to other bodies that are considered as at rest, although it may be that the second and the first both participate in a common motion. And when two bodies collide, even if both are subject to a uniform motion as well, to an observer who has this common motion they will repel each other just as if this parasitical motion did not exist.
>
> Thus let an experimenter be carried by a ship in uniform motion and let him make two equal spheres, that have equal and opposite velocities with respect to him and the ship, collide. We say that the two bodies will rebound with velocities that are equal with respect to the ship, just as if the impact were produced in a ship at rest or on *terra firma.*

Huygens' presentation of his arguments then begins

> Imagine that a ship is carried alongside the bank by the current of a river and that it is so close to the edge that a passenger on the ship can hold the hands of an assistant on the bank. . . .

and continues for nearly two pages (see Dugas [1; pp. 177–178] for the complete text), together with a delightfully quaint illustration.



Huygens' detailed explanation probably arose from his stance on a contentious question. The statement in Hypothesis III that

> the expressions 'motion of bodies' and 'equal or unequal velocities' should be understood relatively to other bodies that are considered at rest,

might seem like nothing more than our recognition that the notion of position, and thus of velocity, depends on the coordinate system used by the observer, but Huygens, like Leibniz and Descartes and his followers, maintained that only relative motion had a meaning, whereas Newton felt that one had to resort to a notion of "absolute space", objecting that otherwise velocities could be assigned arbitrarily, making it meaningless to say that a body unacted upon by forces has constant velocity.

Nowadays we avoid, or at any rate hope we have managed to avoid, the whole problem by stating the first law as on page 11, in terms of the *existence* of an inertial system. No doubt, however, objections to this approach can be raised also. From Newton's day on, there have been extensive arguments about this matter, all of which might be characterized as being of a philosophical nature. Without meaning to assign too pejorative a meaning to that term, let us simply say here that between the two viewpoints there was no disagreement on actual experimental results. Newton certainly wouldn't have disputed the claim that any laws of mechanics that we discover in the one coordinate system ought to hold just as well in the other, because this is, in fact, an immediate consequence of Newton's laws, the crucial point being that the second law involves only $\mathbf{v}'$, and not $\mathbf{v}$. The only question is whether we want to give prominence to the claim, and note that it implies that Newton's laws should involve only $\mathbf{v}'$, and not $\mathbf{v}$, or instead regard the claim as a consequence of Newton's laws.[1]

In any case, whether we decide to note it as a consequence of Newton's laws, or regard is as a fundamental assumption, the basic notion that the laws of mechanics will appear the same in two coordinate systems moving at uniform velocity with respect to each other is nowadays often called the *Galilean relativity principle*, and Huygens' argument, despite its limitations, is certainly an alluring application of this principle.

Galileo, of course, didn't use the term "relativity principle"—that terminology was introduced only after the appearance of "Einstein's principle of relativity"—but he did enunciate the principle quite explicitly, and argued it in great detail

[1] Actually, there's a whole other aspect of this argument, which we touch upon at the end of Chapter 10.

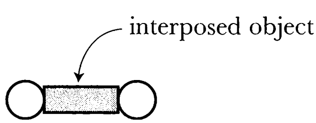
in Galileo [1; pp. 186-88 of the University of California Press edition], a very amusing account, too long to quote here, from which we give a short extract:

> Shut yourself up with some friend in the main cabin below decks on some large ship, and have with you there some flies, butterflies, and other small flying animals. With the ship standing still, observe carefully how the little animals fly with equal speed to all side of the cabin.
>
> ...
>
> in throwing something to your friend, you need throw it no more strongly in one direction than another, the distances being equal; jumping with your feet together, you pass equal spaces in every direction.
>
> ...
>
> [then] have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still.
>
> ...
>
> The cause of all these correspondences of effects is the fact that the ship's motion is common to all the things contained in it, and to the air also. ...

Galileo used this argument to explain why naive objections to the Copernican system—that if the earth rotated from west to east, then cannonballs shot eastward should fly further than those shot westward, or that objects falling from tall buildings should end up to the west of the building when they hit the earth—were mistaken, and that book was the one that got him into trouble with the Inquisition.

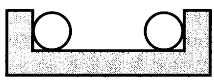
Newton's proof of the third law. The little experiment with a magnet and piece of iron described on page 23ff. was inspired by something in the Principia, though it may well be the silliest thing that Newton ever said, at least among scientific statements.

After his description of his pendulum experiments mentioned on page 22, which involved the *repulsive* forces of collisions, Newton also wanted to say something about *attractive* forces, since he had gravity in mind. So after three pages describing his careful experimentation, he immediately adds the following paragraph, concerning two bodies separated by an interposed obstacle:



> I demonstrate the third law of motion for attractions briefly as follows. Suppose that between any two bodies A and B that attract each other any obstacle is interposed so as to impede their coming together. If one body A is more attracted toward the other body B than that other body B is attracted toward the first body A, then the obstacle will be more strongly pressed by body A than by body B and accordingly will not remain in equilibrium. The stronger pressure will prevail and will make the system of the two bodies and the obstacle move straight forward in the direction from A to B and, in empty space, go on indefinitely with a motion that is always accelerated, which is absurd and contrary to the first law of motion. . . .

Thus, after three pages of careful experiment, Newton provides a one paragraph theoretical argument, and this argument is patently nonsense! The first law is concerned with the force on one body, not on a "system" consisting of more than one body. Moreover, the whole argument depends on the fact that the "interposed" object is rigid, so that it keeps A and B separated, and of course an analysis of rigid bodies presupposes the third law. Finally, we might note that the same argument could just as well be made to work for repulsive forces:



What's even more amazing is that Newton actually described an experiment made to test this idea, using vessels floating on water instead of an air trough to reduce friction:

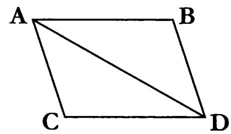
> I have tested this with a lodestone and iron. If these are placed in separate vessels that touch each other and float side by side in still water, neither one will drive the other forward, but because of the equality of the attraction in both directions they will sustain their mutual endeavors toward each other, and at last, having attained equilibrium, they will be at rest.

Considering the great importance that Newton attached to accurate experiments, his claim to have performed this experiment presumably should be taken at face value. But what a truly strange negative experiment this must have been! Would any one really expect that the lodestone and iron would continue moving together forever, with a motion that is always accelerated?!! On the other hand, what a revealing experiment we obtain when we remove the "interposed" object.

The parallelogram law. Some notion of the parallelogram law, at least in terms of the composition of *motions*, seems to date back at least to Aristotle (cf. Dugas [1; pg. 21]): "Let a moving body be simultaneously actuated by two motions that are such that the distances traveled in the same time are in a constant proportion. Then it will move along the diagonal of a parallelogram which has as sides two lines whose lengths are in this constant relation to each other."

And here is Newton's statement, and proof:

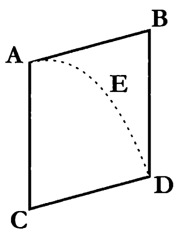
> *A body acted on by [two] forces acting jointly describes the diagonal of a parallelogram in the same time in which it would describe the sides if the forces were acting separately.*
>
> Let a body in a given time, by force M alone impressed in A, be carried with uniform motion from A to B, and, by force N alone impressed in the same place, be carried from A to C; then complete the parallelogram ABDC, and by both forces the body will be carried in the same time along the diagonal from A to D. For, since force N acts along the line AC parallel to BD, this force, by law 2, will make no change at all in the velocity toward the line BD which is generated by the other force. Therefore, the body will reach the line BD in the same time whether force N is impressed or not, and so at the end of that time will be found somewhere on the line BD. By the same argument, at the end of the same time it will be found somewhere on the line CD, and accordingly it is necessarily found at the intersection D of both lines. And, by law 1, it will go with [uniform] rectilinear motion from A to D.



Even before we reach any questionable steps, we see from the very first phrases that Newton is framing this proof in terms of impulsive forces, since he states that the forces M and N individually produce a *uniform* motion on the object. The remaining part of the argument, with its claim that the force N "will make no change at all in the velocity toward the line BD which is generated by the other force", really requires for its justification certain remarks that Newton makes after the statement of the Second Law; the text, as quoted at the bottom of page 12, is immediately followed by an amplifying paragraph that concludes " . . . if [the force] was in an oblique direction [to the direction of a moving body], [it] is combined obliquely and compounded with it according to the directions of both motions." In this completed form, the statement of the Second Law practically includes the parallelogram law! At best, however, it simply demonstrates this result if we regard the body as having already been set in motion by force M, with the second force N applied a bit later. Essentially Newton observes that the result holds when N is applied shortly after M, or *visa versa*, and concludes that it holds when they are applied at the same time. Of course, one can say that this is a very reasonable assumption, but that just replaces one axiom with another.[1]

Moreover, it's certainly not clear how one would apply this assumption to the case of continuous forces, and although Newton's proof is formulated in terms of impulsive forces, he clearly means to apply it to continuous forces also. In fact, in his Scholium he mentions Galileo's observations on the parabolic shape of a projectile's path as an illustration of this rule for compounding forces, and even goes so far as to provide a little picture:

> For example, let body A by the motion of projection alone describe the straight line AB in a given time, and by the motion of falling alone describe the vertical distance AC in the same time; then complete the parallelogram ABDC, and by the compounded motion the body will be found in place D at the end of the time; and the curved line AED which the body will describe will be a parabola which the straight line AB touches at A and whose ordinate BD is as AB^2.



Here, of course, we are considering, on the one hand, an impulsive force, which gives the object its uniform horizontal motion, and, on the other hand, the force of gravity, which gives the object its non-uniform vertical motion. And indeed this really illustrates only that the action of a force on an object is independent of the object's uniform velocity, which was Galileo's basic observation.

Newton's questionable proof of the parallelogram law apparently stimulated the search for more convincing arguments, a somewhat quixotic enterprise, since we are trying to provide a mathematical "proof" for a non-mathematical question. It is probably better to rephrase it as an investigation into whether the parallelogram law can be deduced from other assumptions that we might consider more basic. Daniel Bernoulli offered one of the first such demonstrations in 1726. Other proofs appear in Laplace's great *Traité de Mécanique Céleste* of 1799 and in Poisson's *Traité de Mécanique* of 1833, and a short proof was offered by Hamilton in 1841, to name just some of the more illustrious contributors to this question.

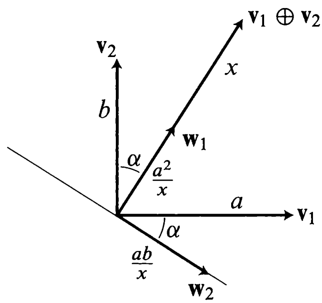
[1] Pourciau [5; pp. 161–163] and [6] gives a rather different interpretation to the whole historical problem that we are discussing, involving an investigation of what Newton "really meant" by his statement of the second law.

All these proofs are shrouded in a somewhat impenetrable veil of unstated assumptions, making their decipherment rather difficult, especially since even mathematical results were often stated rather vaguely in those days. More significantly, all these proofs have one feature in common: they use, but carefully do not state, the one essential hypothesis without which no conclusions can possibly be drawn.

To illustrate this point, we will indicate the first part of Bernoulli's proof. To simplify the discussion, let us restrict ourselves to $\mathbb{R}^2$, so that we are only considering forces in one plane. As we already mentioned long ago, in the footnote on page 29, the first basic assumption is that two forces $\mathbf{v}, \mathbf{w}$ acting together have the same effect as some other force. Thus, we are assuming that for each pair $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$ we have another element $\mathbf{v} \oplus \mathbf{w} \in \mathbb{R}^2$. We presumably shouldn't object to assuming that $\mathbf{v} \oplus \mathbf{w} = \mathbf{w} \oplus \mathbf{v}$ and also equations like $\mathbf{v} \oplus \mathbf{v} = 2\mathbf{v}$ (compare Problem 1-25). We will also need an hypothesis that expresses our experience that the laws of nature are invariant under orthogonal maps:

> If $T: \mathbb{R}^2 \to \mathbb{R}^2$ is any orthogonal map, for the usual vector space structure of $\mathbb{R}^2$, then
>
> $$T(\mathbf{v} \oplus \mathbf{w}) = T(\mathbf{v}) \oplus T(\mathbf{w}) \qquad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbb{R}^2.$$

Bernoulli begins by considering two perpendicular vectors, $\mathbf{v}_1$ of length a, and $\mathbf{v}_2$ of length b, with x being the length of $\mathbf{v}_1 \oplus \mathbf{v}_2$. Let $\mathbf{w}_2$ be the vector on the line perpendicular to $\mathbf{v}_1 \oplus \mathbf{v}_2$ with length $\frac{ab}{x}$, and let $\mathbf{w}_1$ be the vector



along $\mathbf{v}_1 \oplus \mathbf{v}_2$ of length $\frac{a^2}{x}$. Since

$$\begin{aligned}
\text{length } \mathbf{w}_1 &= \tfrac{a}{x} \cdot \text{length } \mathbf{v}_1 \\
\text{length } \mathbf{w}_2 &= \tfrac{a}{x} \cdot \text{length } \mathbf{v}_2 \\
\text{length } \mathbf{v}_1 &= \tfrac{a}{x} \cdot \text{length}(\mathbf{v}_1 \oplus \mathbf{v}_2),
\end{aligned}$$

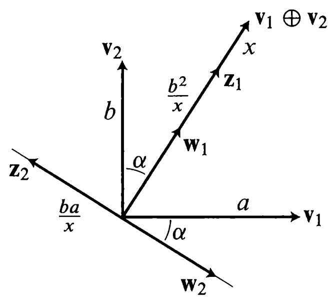
and the angle α from $\mathbf{w}_2$ to $\mathbf{v}_1$ equals the angle from $\mathbf{w}_1$ to $\mathbf{v}_2$, there is an orthogonal map T—involving a rotation through the angle α, together with a reflection—such that

$$\begin{aligned} \tfrac{a}{x}T(\mathbf{v}_1) &= \mathbf{w}_1 \\ \tfrac{a}{x}T(\mathbf{v}_2) &= \mathbf{w}_2 \\ \tfrac{a}{x}T(\mathbf{v}_1 \oplus \mathbf{v}_2) &= \mathbf{v}_1. \end{aligned}$$

Consequently, our hypothesis of invariance under orthogonal maps implies that

$$(1) \qquad \mathbf{w}_1 \oplus \mathbf{w}_2 = \mathbf{v}_1.$$

But similarly, we can consider $\mathbf{z}_1$ and $\mathbf{z}_2$ of lengths $\frac{ba}{x} = \frac{ab}{x}$ and $\frac{b^2}{x}$, respectively, and conclude that we have



$$(2) \qquad \mathbf{z}_1 \oplus \mathbf{z}_2 = \mathbf{v}_2.$$

Since $\mathbf{w}_2 = -\mathbf{z}_2$, equations (1) and (2) give

$$(3) \qquad \mathbf{w}_1 \oplus \mathbf{z}_1 = \mathbf{v}_1 \oplus \mathbf{v}_2.$$

But $\mathbf{w}_1$ and $\mathbf{z}_1$ lie along $\mathbf{v}_1 \oplus \mathbf{v}_2$, so the length of $\mathbf{v}_1 \oplus \mathbf{v}_2$ is the sum of the lengths of $\mathbf{w}_1$ and $\mathbf{z}_1$, which means that

$$\frac{a^2}{x} + \frac{b^2}{x} = x \implies a^2 + b^2 = x^2 \implies x = \sqrt{a^2 + b^2},$$

and thus $\mathbf{v}_1 \oplus \mathbf{v}_2 = a\mathbf{e}_1 \oplus b\mathbf{e}_2$ has length $\sqrt{a^2 + b^2}$, which is the length of $a\mathbf{e}_1 + b\mathbf{e}_2$.

Thus, Bernoulli has demonstrated that $\mathbf{v}_1 \oplus \mathbf{v}_2$ has precisely the length you would expect it to have, in the special case that $\mathbf{v}_1$ and $\mathbf{v}_2$ are perpendicular. He then proceeds by involved arguments to prove the complete result, for the general case.

The one point that is usually ignored is that in our quick trip from equations (1) and (2) to (3), we had to use associativity of $\oplus$, which is likewise used in all the other proofs that have been fashioned. But if we assume associativity of $\oplus$, then everything is essentially trivial: Consider the map

$$(a, b) = a\mathbf{e}_1 + b\mathbf{e}_2 \longmapsto a\mathbf{e}_1 \oplus b\mathbf{e}_2.$$

If $\oplus$ is associative, then this map will be linear. But it takes $\mathbf{e}_1$ to $\mathbf{e}_1$ and $\mathbf{e}_2$ to $\mathbf{e}_2$, so it must be the identity. Q.E.D. The somewhat convoluted discussion in Mach [1; pp. 55–57] essentially boils down to the same point. Note also that there is no reasonable way to verify associativity experimentally without already being able to measure what $\mathbf{v} \oplus \mathbf{w}$ is.

Newton at the hands of the scholars. As pointed out in Addendum 4A, almost none of Newton's contemporaries realized the significance of Proposition 41 of the Principia, which amounts to the modern solution to the problem of inverse square forces, while the geometric proof that he first supplied confused them by its brevity. In fact, the particular form in which Newton chose to present his proof has spawned all sorts of scholarly arguments, which began in his day and have continued into ours. As we mentioned in Chapter 2, in the first edition of the Principia (1687), the Corollary on page 62 was simply stated, without the remaining two explanatory sentences. These were added to the second edition (1713), and we have a letter of 1709 in which Newton instructs the editor to supply them.

This date is significant because in 1710 Johann Bernoulli criticized the Corollary on various grounds, including the assertion that it basically amounted to assuming the converse of a proposition on the basis of the proposition itself. Skipping over the acrimonious disputes that arose,[1] we merely note that in 1719, Bernoulli wrote Newton an apologetic, not to say obsequious, letter that said in part

> Gladly I believe what you say about the addition to Corollary 1, Proposition 13, Book One of your incomparable work, the *Principia*, that this was certainly done before these disputes began, nor have I any

[1] For a thorough treatment, see Guicciardini [1]; by the way, pg. 546 of this paper points out that Newton definitely knew how to evaluate the integrals required for a direct proof.

> doubts that the demonstration of the inverse proposition, which you have merely stated in the first edition of the work, was yours; I only ... wished that someone would give a [direct proof]. This indeed, which I would not have said to your displeasure, I think was first put forward by me, as least so far as I know at present.

Actually, Bernoulli was doubly wrong on this account, as Newton had already provided the proof in Proposition 41 of the Principia, nor did Bernoulli even have priority for a published specifically analytic proof, as pointed out in Guicciardini [1].

Bernoulli's letter may be found in Newton [1, Vol. 7, p. 77], where the real reason for writing it is revealed a bit later.

Now one would think that this would have settled the matter! But Newton has continued to be faulted by various physicists. Wintner [1] offhandedly credited Bernoulli with being the first to prove that the paths are conics, and more recent challenges have led to all sorts of embarrassing scholarly cat fights,[1] which should have been quite unnecessary since the third edition makes the argument quite clear to any one with mathematical instincts.

It might also be mentioned that Arnold addressed this question in a very different way, by arguing that Newton knew "in essence" a short theorem—with a neat one-paragraph proof, Arnold [3; pg. 32]—which immediately provides the desired result. Arnold might have been overly generous in attributing knowledge of this result to Newton, or perhaps he was simply singularly in tune with Newton's mode of thought.

---

[1] See Cohen and Whitman [1; pp. 135-136] for many references; Pourciau [1] is cited as a corrective antidote, and further details are added in Pourciau [2].

PART II

BUILDING ON THE FOUNDATIONS

CHAPTER 8
OSCILLATIONS

> The investigations that are to follow will teach us nothing new about the principles of mechanics. So great, however, is the significance of oscillation processes for physics and engineering that their separate systematic treatment is deemed essential.
>
> — Sommerfeld, *Mechanics*

Thus opens the third chapter, on oscillation problems, in the famous book on *Mechanics* by Arnold Sommerfeld. Actually, various results about oscillations, and about the equations defining them, are needed in succeeding chapters, but we will proceed somewhat in the same spirit, first discussing some specific interesting mechanical systems, then progressing to topics of greater generality, which can often be applied to mechanics itself.

One sort of oscillation is already familiar to us. In our analysis of the pendulum moving through a small angle in Problem 1-21, we have already encountered the equation

$$x'' + \omega^2 x = 0,$$

with oscillatory solutions

$$x(t) = a\cos\omega t + b\sin\omega t \qquad \begin{cases} a = x(0), \\ b = x'(0)/\omega. \end{cases}$$

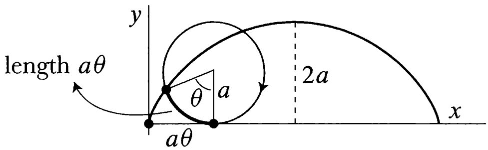
This is known as "simple harmonic motion", and we should note that it can also be written in terms of the *amplitude* A and the *phase* ϕ, as

$$x(t) = A\sin(\omega t + \phi), \qquad A = \sqrt{a^2 + b^2}, \quad \tan\phi = \frac{a}{b}.$$

More systematically, we can seek a solution in terms of the exponential function $x(t) = e^{\lambda t}$, giving the equation $\lambda^2 + \omega^2 = 0 \implies \lambda = \pm i\omega$, noting that $Ce^{i\omega t} + De^{-i\omega t}$ will be real if C and D are conjugate; the simplest choice $C = D = a/2$ for real a gives $a\cos\omega t$, while $C = -ib/2 = -D$ for real b gives $b\sin\omega t$.

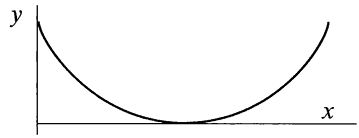
The oscillating motion repeats itself over a *period* T of $2\pi/\omega$, so that the *frequency* $\nu = 1/T$ is $\nu = \omega/2\pi$. The term *circular frequency* is often used for $\omega = 2\pi\nu$.

Huygens' cycloidal pendulum. In Problem 1-21 we had to restrict ourselves to oscillations through small angles because the pendulum is not *isochronous* (its period is not constant), but merely close to isochronous for oscillations through small angles, which can be rather inconvenient in the design of pendulum clocks. This problem was handled in an ingenious way by Huygens, who has been mentioned in Chapters 1 and 7 and who will also make an important appearance



cycloid

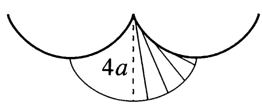
$$x = a(\theta - \sin\theta),$$
$$y = a(1 - \cos\theta).$$



upside-down cycloid

$$x = a(\theta - \sin\theta),$$
$$y = a(1 + \cos\theta).$$

later on. Huygens discovered that an upside-down cycloid is *tautochronous*—if a particle starts at rest from any point, the time taken to slide down to the bottom is always the same. So a pendulum bob sliding on a perfectly smooth cycloid would have the same period no matter how large or small the angle through which the pendulum bob slid. Huygens managed to obtain the equivalent of this frictionless situation by means of another wonderful fact he discovered about cycloids: A cycloid has an "involute" that is another cycloid of the same size: If we tie a string of length $4a$ to the top vertex of a cycloid and pull it taut against

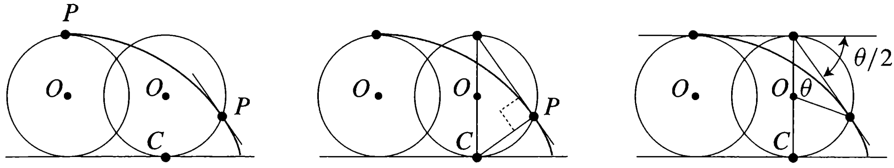


the cycloid, then as we pull it out tangent to the cycloid we obtain a congruent cycloid, basically just like a simple pendulum swinging along a circular arc. It was Huygens who introduced the notion of involute, as well as the related notion of evolute, treated geometrically in terms of envelopes, see Addendum B and Problem 2.

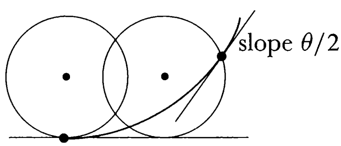
Though we now have many alternatives to pendulum clocks for accurate measurements of time, Huygens' investigation inspired Abel's pioneering work, described briefly in Addendum A, in the subject of integral equations.

Huygens originally proved the cycloid was tautochronous by complicated geometric arguments, which can be found in English translation in Huygens [1] or on Ian Bruce's wonderful web site `17centurymaths.com`. The analytic proofs one encounters nowadays are somewhat opaque, but we can give a proof that is probably close to the spirit of Huygens' arguments by using a fact about cycloids that is most easily proved geometrically.

In the figure below, (a) shows the position that the top point P of the circle has moved into after the circle has rolled a certain amount. Since the circle



is rolling, we know from the Proposition on page 222 that "up to first order" the motion of P is now simply rotation about C, which means that the tangent line to the cycloid at P is perpendicular to CP, as in (b); and since an angle inscribed in a semi-circle is right angle, and conversely, this means that the tangent line goes through the point directly above C. Finally, from (c) we easily see that if the circle has rotated by the angle θ to get to this position, then the angle between the tangent and the horizontal line above the circles is just $\theta/2$. Turning this picture upside down, this means that for an upside-down cycloid,



after the circle rotates through an angle of θ from the bottom, the tangent line has slope $\theta/2$; an awkward analytic derivation can be found in Problem 1.

Now if we let $s(\theta)$ be the length of the cycloid as a function of θ, then from the parameterization on the previous page we get

$$\begin{aligned} s'(\theta)^2 &= x'(\theta)^2 + y'(\theta)^2 = 2a^2(1-\cos\theta) \\ &= 4a^2\sin^2(\theta/2) \qquad \text{by the half angle formulas.} \end{aligned}$$

So the length from the bottom—reached after rotation through the angle π—to the point where the circle has rotated through an additional angle of θ, is

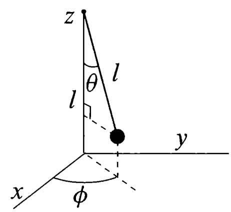
$$s(\theta) = \int_{\pi}^{\pi+\theta} s'(\theta)\,d\theta = \int_{\pi}^{\pi+\theta} 2a\sin(\theta/2)\,d\theta = 4a\sin(\theta/2).$$

But since the slope of the tangent line is $\theta/2$ at this point, a particle sliding down the cycloid, with $s(t)$ its distance from the bottom, satisfies the equation

$$s''(t) = -g\sin(\theta/2) = -\frac{g}{4a}\cdot s(t),$$

and we have simple harmonic motion with a period of $2\pi\sqrt{4a/g} = 4\pi\sqrt{a/g}$, independent of the the point from which the particle begins sliding; the time of descent is always $2\pi\sqrt{a/g}$.

The spherical pendulum. The spherical pendulum, introduced in Problem 3-5, is a very valuable example, which will be used to illustrate important points later on, especially in Chapters 12 and 21. An exact analysis involves elliptic functions, just as for the ordinary pendulum, but we can give an analysis that describes the basic features of the motion, which will have some interesting



parallels in Chapter 9. Regarding our pendulum bob as a particle $c(t) = (x(t), y(t), z(t))$ of mass m, Problem 3-5 gives

(a) $$xy' - yx' = C$$

for a constant C, while conservation of energy gives

(b) $$\tfrac{1}{2}mv^2 + mgz = E.$$

In terms of the spherical coordinates θ and ϕ, with

$$\begin{aligned} x &= l\cos\phi\sin\theta & & & x' &= l\theta'\cos\phi\cos\theta - l\phi'\sin\phi\sin\theta, \\ y &= l\sin\phi\sin\theta & &\Longrightarrow & y' &= l\theta'\sin\phi\cos\theta + l\phi'\cos\phi\sin\theta, \\ z &= l(1-\cos\theta) & & & z' &= l\theta'\sin\theta, \end{aligned}$$

we find that

$$v^2 = x'^2 + y'^2 + z'^2 = l^2(\theta'^2 + \phi'^2\sin^2\theta)$$

and we can write equations (a) and (b) as

(a′) $$l^2\phi'\sin^2\theta = C$$

(b′) $$\tfrac{1}{2}ml^2(\theta'^2 + \phi'^2\sin^2\theta) + mgl(1-\cos\theta) = E.$$

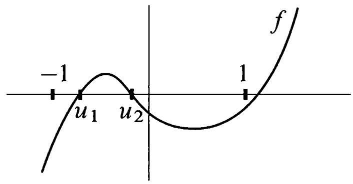
The substitution

$$u = \cos\theta, \qquad \theta' = \frac{-u'}{\sqrt{1-u^2}}$$

leads to

$$u'^2 = \frac{2}{ml^2}(1-u^2)(E - mgl(1-u)) - \frac{C^2}{l^4} = f(u), \quad \text{say,}$$

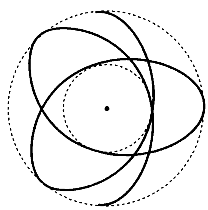
which means that $u = \cos\theta$ can only have values where the cubic $f(u) \geq 0$. If the pendulum isn't swinging in a plane, then $C \neq 0$, so $f(u) < 0$ for $u = \pm 1$,



and thus u lies in some interval (u_1, u_2) with $-1 < u_1 < u_2 < 1$ (the figure, with $u_2 < 0$, is for a pendulum whose swing remains below the horizontal). Differentiation of our equation for u'^2 leads to a second order equation

$$2u'' = f' \circ u \quad \text{or, in Leibnizian notation,} \quad 2\frac{d^2u}{dt^2} = f'(u),$$

of the very form that we encountered in Chapter 4 (see page 128), and the pendulum, when viewed from above, exhibits the same characteristics as an elliptical orbit, with the height $l(1 - \cos\theta) = l(1 - u)$ of the pendulum bob



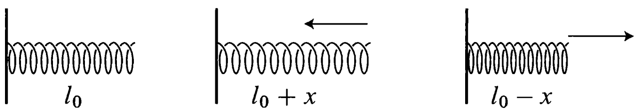
varying between $l - u_1$ and $l - u_2$. (Problem 1-20 covers the case $u_1 = u_2$.)

If, as in the case of the regular pendulum, we want to consider small oscillations of the spherical pendulum, the coordinates ϕ and θ don't seem very promising for our analysis, since they don't play a symmetric role. Indeed, θ really wasn't a good choice for the regular pendulum problem, for we had to replace the $\sin\theta$ by θ in Problem 1-21 in order to get a linear equation, and as Problems 1-22 and 3-16 indicate, it will be easier simply to use the coordinates x and y instead. The method of Problem 3-16 (with z now playing the role of y in that problem) leads to the equation

$$x'' + \frac{g}{l}x + y'' + \frac{g}{l}y = 0,$$

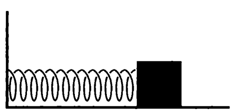
and we can then simply solve $x'' + (g/l)x = 0$ and $y'' + (g/l)y = 0$ separately, each as linear combinations of $\sin\omega t$ and $\cos\omega t$ for $\omega^2 = g/l$, and obtain any given initial conditions by an appropriate combination of these solutions. Problem 2-4 implies that the (x, y) component of the path of the pendulum bob is a small ellipse, which might seem like a poor approximation to the rotating elliptical path previously described. Unlike the case of the ordinary pendulum, where we definitely had an oscillating motion, which we then approximated to first order by a harmonic oscillation, in the case of the spherical pendulum we have made a first order approximation to the whole problem at the outset, before identifying a specific oscillatory motion to which our approximations should apply, so we are only studying small oscillations of an approximation, not harmonic oscillatory approximations to actual small motions!

Springs. In addition to the beautiful oddity of the Huygens pendulum, oscillations with constant period are also produced by springs, which obey Hooke's law, briefly mentioned on page 47 as well as in Addendum 6B. For a spring with one end fixed and unstretched length l_0, there is a "spring constant" $\kappa > 0$ so that when the string is stretched to length l the force on the end of the spring is $-\kappa x$ for $x = l - l_0$. This is true only for $|x|$ for which $l_0 + x$ is within a



certain "elastic limit". However, this range is far greater than the range for small oscillations of a pendulum, and within this range the proportionality is quite accurate (though spring stretching is actually quite complicated, involving a twisting of the spring around its axis, and a reshaping of the spring; the stretching of a wire as discussed on page 253 is simpler, at least conceptually.)

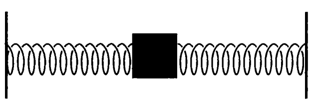
For an object of mass m attached to a spring of negligible mass and sliding on a frictionless surface, moving it further from its equilibrium point sets up a



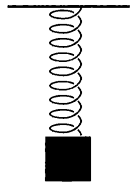
motion described by the equation

$$x'' = -Kx, \qquad K = \kappa/m.$$

Or we could eliminate the problem of friction by using two springs of the same construction to suspend the weight between two walls, where K will now



have twice the value, if we neglect the sagging due to gravity, or work in a weightless environment. We could also simply hang the weight from one end of the spring, which is suspended from the ceiling at the other end; in this case, we should measure x as the displacement from the equilibrium point, rather



than from the unstretched length of the spring. In all cases we end up with an equation of the form

$$x'' + Kx = 0, \qquad K > 0,$$

whose solutions we mentioned at the beginning of the chapter.

For

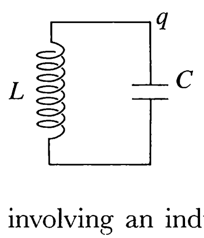
$$\begin{aligned} x(t) &= A\sin(\omega t + \phi) \\ x'(t) &= A\omega\cos(\omega t + \phi) \end{aligned} \qquad \omega = \sqrt{\kappa/m},$$

if we choose the potential energy U to be 0 at $x = 0$, or $\sin(\omega t + \phi) = 0$, and hence $\cos(\omega t + \phi) = \pm 1$, the total energy E at this time is the kinetic energy

$$\tfrac{1}{2}mv^2 = \tfrac{1}{2}m|A\omega\cos(\omega t + \phi)|^2 = \tfrac{1}{2}mA^2\omega^2 = \tfrac{1}{2}A^2\kappa,$$

so that $E = \frac{1}{2}A^2\kappa$ at all times. Conversely (Problem 3), the method of deriving conservation of energy can be used to obtain our solution for $x(t)$.

Harmonic oscillations. Our spring example is about the only mechanical one giving rise to simple harmonic oscillation not restricted to small oscillations, but simple harmonic oscillation occurs in other important physical systems. In



particular, for a simple circuit involving an inductance L attached across the plates of a capacitance C, the charge q satisfies

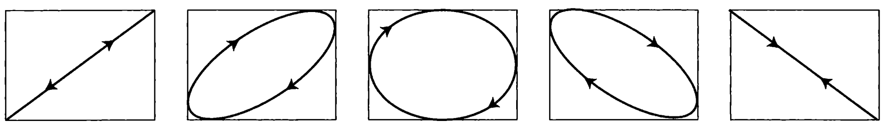
$$Lq'' + q/C = 0,$$

so that it varies sinusoidally, over a large range of values, which is why it so easy to get a beautiful sine curve on an oscilloscope, by sweeping vertically with such an oscillating voltage while sweeping horizontally with uniform speed. In fact, many investigations of simple harmonic oscillation are made with electrical examples in mind.

By using the output from two different circuits to sweep horizontally and vertically on an oscilloscope, we can easily get a system whose two coordinates each exhibit simple harmonic oscillation,

$$
\begin{aligned}
x_1(t) &= A_1 \cos(\omega_1 t + \phi_1) \\
x_2(t) &= A_2 \cos(\omega_2 t + \phi_2),
\end{aligned}
$$

producing so-called *Lissajous figures*. When $\omega_1 = \omega_2$ the resulting figure, inside the rectangle $[-A_1, A_1] \times [-A_2, A_2]$, is always a (possibly degenerate) ellipse (Problem 2-4 again). The figure below shows several cases, where, setting

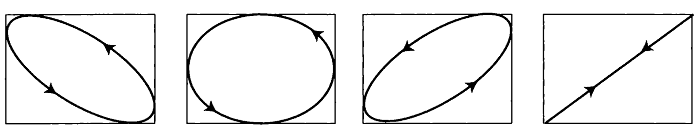


$\delta = 0$ $\quad \dfrac{\pi}{4}$ $\quad \dfrac{\pi}{2}$ $\quad \dfrac{3\pi}{4}$ $\quad \pi$

$\delta = \phi_2 - \phi_1$, we have shifted the time parameter so that the equations can be written in terms of the *phase difference* δ as

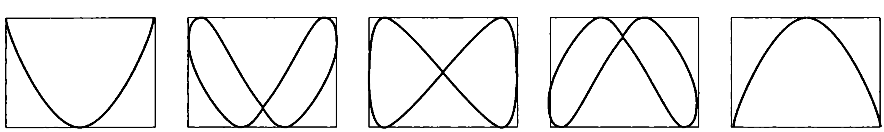
$$
\begin{aligned}
x_1(t) &= A_1 \cos(\omega t) \\
x_2(t) &= A_2 \cos(\omega t + \delta);
\end{aligned}
$$

the arrows on the paths indicate the direction of the path for increasing t. The next figure shows the results as δ increase to 2π, and the bottom figure shows

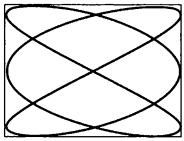


$\dfrac{5\pi}{4}$ $\quad \dfrac{3\pi}{2}$ $\quad \dfrac{7\pi}{4}$ $\quad 2\pi$

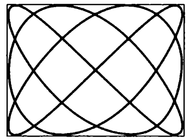
several results for $\omega_2 = 2\omega_1$, where the first and last curves are parabolas.



The shape of the Lissajous figure is very sensitive to the ratio of ω_1 and ω_2. In general, if ω_1 and ω_2 are commensurable, then their ratio is the ratio of the numbers of tangencies of the figure along a horizontal and a vertical side of the rectangle, except for the cases where the figure enters a corner. If ω_2 is just



2:3



4:3

a bit larger than ω_1, then the segments for $[0, 2\pi]$, $[2\pi, 4\pi]$, ... , are close to ellipses, but they keep rotating, so the Lissajous curve slowly rotates around, and if ω_1 and ω_2 are not commensurable the Lissajous figure is not periodic, and its image is dense in the rectangle. Lissajous figures on an oscilloscope thus give a good test of whether two frequencies are the same. Lissajous himself, before the days of oscilloscopes, tested whether two tuning forks had the same frequency by directing a narrow beam of light onto a tiny mirror glued to one tine of the first vibrating tuning fork, which reflected it to a mirror on the second vibrating tuning fork, which in turn reflected it to a screen.

Damped oscillations. In practice, of course, springs never provide truly harmonic oscillation. Even if we ignore the fact that springs are not massless, there is always some outside force, like the friction of the moving weight on the floor, or the resistance of the surrounding air or a fluid, that tends to slow the oscillation down. Moreover, there will generally be internal factors, like the fact that the spring heats up, that act similarly. It is customary to consider the case where the total "damping" force is proportional to the velocity. The frictional force of a moving weight on a floor is *not* proportional to velocity, but a resistance proportional to velocity does describe fairly well the case of air or fluid friction if the motion is slow enough, and internal factors also often act this way, as in the case of the tuning fork and the rubber band on page 298. Perhaps most important of all, the oscillatory behavior of the charge q in electrical circuits often has this character.

Instead of our equation $x'' + \omega^2 x = 0$, we will now write

$$x'' + 2\rho x' + \omega_0{}^2 x = 0, \qquad \rho > 0,$$

where ω_0 is the "natural" circular frequency that we would have without the damping force, and the factor 2 is inserted to simplify some algebra. If we try for a complex solution that is a multiple of $x(t) = e^{\omega t}$ we obtain

$$\omega^2 + 2\rho\omega + \omega_0{}^2 = 0$$

with the two roots

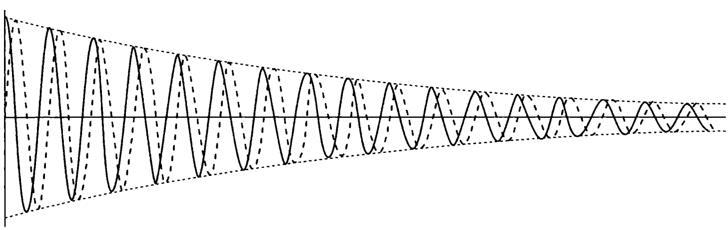
$$\omega_1 = -\rho + \sqrt{\rho^2 - \omega_0{}^2},$$
$$\omega_2 = -\rho - \sqrt{\rho^2 - \omega_0{}^2},$$

and the nature of the solution depends on the sign of $\rho^2 - \omega_0{}^2$.

$\rho < \omega_0$, "underdamped". Letting $\omega = \sqrt{\omega_0{}^2 - \rho^2}$, we have

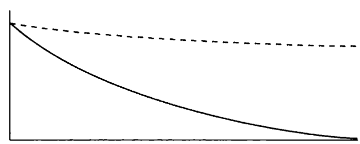
$$x(t) = e^{-\rho t}(Ae^{i\omega t} + Be^{-i\omega t}),$$

giving the solutions $e^{-\rho t}(a\cos\omega t + b\sin\omega t)$. It is customary to speak of this as an "oscillation" of circular frequency ω; the zeroes of the solutions are spaced apart by π/ω, although the maxima and minima are not equally spaced between them. The figure below shows the basic solutions involving cos and sin alone.



$\rho > \omega_0$, "overdamped". Now ω_1 and ω_2 are real, negative, and distinct, and the solutions are linear combinations

$$x(t) = ae^{\omega_1 t} + be^{\omega_2 t}$$



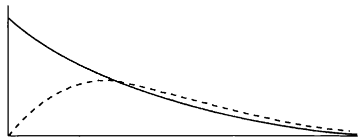
of two exponential decays. When $\rho^2 - \omega_0{}^2$ is large, the root ω_1 is close to 0, and there will be solutions with a component $ae^{\omega_1 t}$ that decay only very slowly. As we make $\rho^2 - \omega_0{}^2$ smaller this becomes less pronounced, which gives special importance to the final case:

$\rho = \omega_0$, "critically damped". Then we have only one root $\omega = -\rho$, giving us only the solutions $x(t) = ae^{-\rho t}$. The standard way to guess a second solution is to consider the underdamped equation with ${\omega_0}^2 - \rho^2 = \varepsilon$ for small ε, and note that the averaged solution

$$\frac{1}{2i\varepsilon}e^{-\rho t}[e^{i\varepsilon t} - e^{-i\varepsilon t}] \quad \text{approaches} \quad te^{-\rho t} \quad \text{as } \varepsilon \to 0,$$

revealing the general solution

$$x(t) = ae^{-\rho t} + bte^{-\rho t}.$$



In many mechanisms, one wants a quick return to a steady position after an initial displacement. For example, an electrical meter should give a steady reading shortly after it has been connected to a circuit or a switch has been closed, hydraulic and pneumatic spring returns for doors need to close the door reasonably quickly without hitting the door frame so hard that the door bounces back, shock absorbers need to return a car bumped by the road to its initial position without causing the car to oscillate up and down. In such cases, the mechanisms are designed to have a damping constant just a little larger than critical damping.

Returning to the underdamped case, if the damping is small,

$$x(t) = e^{-\rho t}A\cos(\omega t + \phi), \qquad \rho \ll \omega,$$

many oscillations will occur in a period of time during which the $e^{\rho t}$ term varies only slightly, so at any time t we practically have simple harmonic motion of amplitude $Ae^{-\rho t}$ with energy $E(t) = \frac{1}{2}\kappa A^2 e^{-2\rho t} = E(0)e^{-2\rho t}$, decaying exponentially. While $\omega/\pi \approx \omega_0/\pi$ is the reciprocal of the "semi-period" (the time between two zeros of the decaying oscillation), 2ρ is basically the reciprocal of the time it takes for the energy to decay by a factor of $1/e$. The quotient

$$Q = \frac{\omega_0}{2\rho}$$

is a "dimensionless" number[1] called the *quality* of the oscillation; an oscillation with high Q loses very little energy per oscillation.

[1] Alternatively, all terms of the equation $mx'' + m{\omega_0}^2x + 2m\rho x' = 0$ must have the dimensions of force, MLT^{-2}. For $m{\omega_0}^2x$, where mx has the dimensions ML, this means that ${\omega_0}^2$ must have dimensions T^{-2}, so ω_0 has dimensions T^{-1}. Similarly, for $2m\rho x'$, where mx' has dimensions MLT^{-1}, this means that ρ must have dimensions T^{-1}.

As an example[1] of the significance of Q, a standard tuning fork (A above middle C) has a frequency of 440 cycles per second. Using a decibel meter for a rough approximation, the intensity of sound was found to decrease by a factor of 5 in 4 seconds. This means that $4 \times (2\rho) = \log 5 \approx 1.6$, so

$$\begin{aligned}2\rho &\approx 0.4\\ Q &\approx \frac{2\pi \times 440}{0.4} \approx 700.\end{aligned}$$

By contrast, a paperweight suspended from a sturdy rubber band was found to have a period of 1.2 seconds and the amplitude of oscillation decreased by a factor of 2 after three periods, giving

$$\begin{aligned}2\rho &\approx .39\\ Q &\approx \frac{\omega}{2\rho} \approx \frac{2\pi/T}{0.39} \approx \frac{2\pi/1.2}{0.39} \approx 13.\end{aligned}$$

In both cases, the air resistance makes only a small contribution to the damping factor—most of the energy loss occurs internally, showing up as heating of the metal or the rubber during the vibrations. The damping factor 2ρ is nearly the same in both cases, but the tuning fork has a much greater Q because it oscillates much more rapidly, so it has a much lower loss of energy *per cycle*, which is what Q measures.

Forced oscillations. The equation

$$x''(t) + \omega_0{}^2 x(t) = F(t)$$

describes a situation where an external "driving" force F is being applied, in addition to the force of the spring or other appropriate feature of our system that provides the $\omega_0{}^2 x(t)$ term; note that here F already has the mass m of our object divided out, so it actually has the dimensions of acceleration. The general solution of this inhomogeneous equation is a linear combination of any particular solution and the solutions of the homogeneous equation $x'' + \omega_0{}^2 x = 0$. Although a solution can be found for arbitrary F by the method of "variation of parameters" of elementary differential equations, the main case that interests us is when the driving force F is itself oscillating, so that we have the equation

$$x''(t) + \omega_0{}^2 x(t) = c \sin \bar{\omega} t$$

for some $\bar{\omega}$, the circular frequency of the driving force.

---

[1] Kleppner and Kolenkow [1]

If we try for a particular solution of the form

$$x(t) = \boldsymbol{c} \sin \bar{\omega} t,$$

we find that

$$\boldsymbol{c} = \frac{c}{\omega_0{}^2 - \bar{\omega}^2},$$

and our general solution is

$$x(t) = a \cos \omega_0 t + b \sin \omega_0 t + \boldsymbol{c} \sin \bar{\omega} t,$$

where a and b are determined by the initial conditions, while $\boldsymbol{c}$ is a constant.

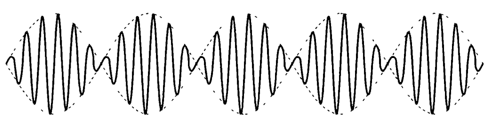
Note that if we write the general solution for $\bar{\omega} = \omega_0 + \varepsilon$ as

$$x(t) = Ae^{i\omega_0 t} + Be^{i(\omega_0+\varepsilon)t} = [A + Be^{i\varepsilon t}]e^{i\omega_0 t},$$

then for $\varepsilon \ll \omega_0$ the factor $A + Be^{i\varepsilon t}$ will vary only slightly over the period $2\pi/\omega_0$ of $e^{i\omega_0 t}$, so we have something like oscillations of period $2\pi/\omega_0$ with varying amplitude $|A + Be^{i\varepsilon t}|$. Writing $A = ae^{i\alpha}$, $B = be^{i\beta}$, we find that

$$|A + Be^{i\varepsilon t}|^2 = a^2 + b^2 + 2ab \cos(\varepsilon t + \beta - \alpha),$$

so this amplitude varies periodically with frequence ε, giving the phenomenon known as *beats.*



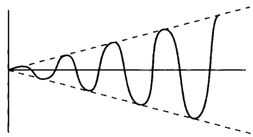
For $\bar{\omega} = \omega_0$ our solution makes no sense, giving an "infinite" amplitude. A specific solution for $\bar{\omega} = \omega_0$ might be discovered, analogously to the case of critically damped oscillations, by considering the solution

$$x(t) = \frac{c}{\omega_0{}^2 - \bar{\omega}^2}(\sin \omega_0 t - \sin \bar{\omega} t),$$

where the limit as $\bar{\omega} \to \omega_0$ is

$$x(t) = -\frac{c}{2\omega_0} t \cos \omega_0 t.$$

This solution still has disconcerting features, as its amplitude approaches ∞ as $t \to \infty$. This anomaly arises not only because our equations won't even hold for



these large oscillations, but also because we have so far ignored the fact that physically there is always some damping present.

Damped forced oscillations. We therefore consider the equation

$$x''(t) + 2\rho x'(t) + {\omega_0}^2 x(t) = F(t),$$

where again we are mainly interested in the case

$$x''(t) + 2\rho x'(t) + {\omega_0}^2 x(t) = C\cos\bar{\omega}t;$$

we can also write this in the more convenient complex form

$$x''(t) + 2\rho x'(t) + {\omega_0}^2 x(t) = Ce^{i\bar{\omega}t}.$$

We now try for a solution of the form $x(t) = (ce^{i\phi})e^{i\bar{\omega}t}$. We need that

$$ce^{i\phi} = \frac{C}{{\omega_0}^2 - \bar{\omega}^2 + 2i\rho\bar{\omega}} = \frac{C({\omega_0}^2 - \bar{\omega}^2 - 2i\rho\bar{\omega})}{({\omega_0}^2 - \bar{\omega}^2)^2 + 4\bar{\omega}^2\rho^2},$$

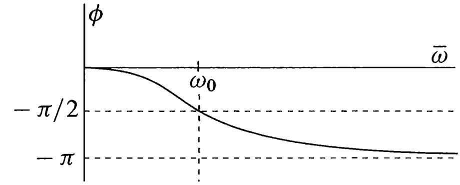
so, denoting the denominator by $\Delta > 0$, we have

$$\left\{\begin{aligned} c\cos\phi &= C({\omega_0}^2 - \bar{\omega}^2)/\Delta \\ c\sin\phi &= -2c\rho\bar{\omega}/\Delta \end{aligned}\right\} \implies c = \frac{C}{\sqrt{\Delta}}, \quad \tan\phi = \frac{-2\rho\bar{\omega}}{{\omega_0}^2 - \bar{\omega}^2}.$$

The real part of the solution $x(t) = ce^{i\phi}e^{i\bar{\omega}t}$ then gives the particular solution

$$c\cos(\bar{\omega}t + \phi).$$

Any ambiguity in determining ϕ from the formula for $\tan\phi$ is resolved by the specific formula for $c\sin\phi$, which shows that $\phi \le 0$ for all $\bar{\omega} > 0$, with the value of $\boldsymbol{\phi}$ starting at 0 when $\bar{\omega} = 0$, reaching $-\pi/2$ at $\bar{\omega} = \omega_0$, and approaching $-\pi$ as $\bar{\omega} \to \infty$.



In particular, consider the underdamped case $\rho < \omega_0$, and let $\omega = \sqrt{\omega_0{}^2 - \rho^2}$, as on page 296. The general solution is

$$x(t) = e^{-\rho t}(a \cos \omega t + b \sin \omega t) + \boldsymbol{c} \cos(\bar{\omega} t + \boldsymbol{\phi}),$$

or, if we also write the solution of the homogenous part in terms of amplitude and phase, as

$$x(t) = Ae^{-\rho t} \cos(\omega t + \phi) + \boldsymbol{c} \cos(\bar{\omega} t + \boldsymbol{\phi}),$$

where a and b, or A and ϕ, are determined by the initial conditions and $\boldsymbol{c}$ and $\boldsymbol{\phi}$ are constants. The first term, which dies out as t becomes large, is called the *transient*, while the second term is the *steady state* solution. Since $\boldsymbol{\phi} < 0$, the steady state solution always lags behind the driving force; it is exactly $\pi/2$ behind when $\bar{\omega} = \omega_0$, and comes close to being π behind as $\bar{\omega}$ increases.

If we plot the value of $\boldsymbol{c}$ against $\bar{\omega}$ (for some fixed C), we get a graph like that in (a) of the figure below, with the maximum at $\bar{\omega} = \sqrt{\omega_0{}^2 - \rho^2}$. Part (b) of the

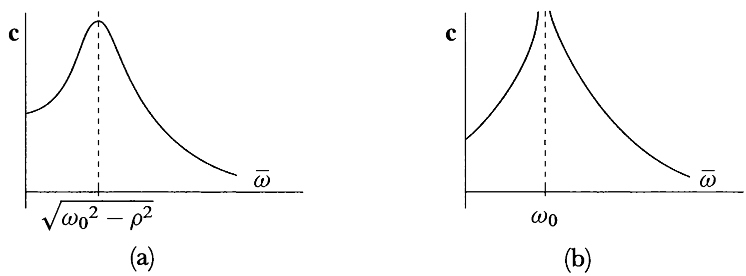
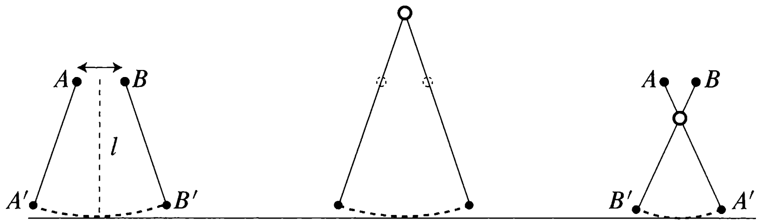


figure shows the graph we would have for $\rho = 0$, as on page 299; the formula actually gives negative values for $\bar{\omega} > \omega_0$, but amplitude is by definition positive, so this must correspond to a phase shift of magnitude π, and consideration of the damped case shows that we should consider it to be a lagging shift $-\pi$, rather than a leading one $+\pi$.

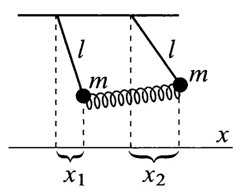
The phase shift can be demonstrated with a pendulum of length l, and thus natural circular frequency $\omega_0 = \sqrt{l/g}$, whose suspension point is being moved horizontally along AB in harmonic motion with circular frequency $\bar{\omega}$ as in (a) of the figure below, since its equation, from Problem 1-18, is precisely an undamped forced oscillation. In the case $\bar{\omega} < \omega_0$, once the steady state has essentially been



(a) (b) (c)

reached the pendulum bob follows the direction of the suspension point: when the suspension point is all the way to the left (A) or right (B), the same is true of the lower end of the pendulum (at A' or B'). This lower ends moves exactly like the end of a pendulum having length $> l$ that is attached to a higher suspension point, as in part (b). Part (c) shows the case where $\bar{\omega} > \omega_0$: the lower end swings right when the suspension point is moving left and left when the suspension point is moving right, and moves like the end of a pendulum of length $< l$ suspended from a lower point.

Coupled oscillators. Many physical systems can be thought of as a collection of oscillators influencing each other. The possible behavior of such systems can be quite complex, even in one of the simplest such systems, consisting of two



pendulums of the same mass m and pendulum length l, connected by a spring, with spring constant κ, whose unstretched length is the distance between the two pendulums when they are both vertical. Letting $K = \kappa/m$ and $\omega_0 = \sqrt{g/l}$, and restricting our attention to a linear approximation of the actual motion, as in the case of the spherical pendulum, we can write the equations for the displacements x_1 and x_2 of the two pendulums from their respective vertical positions as the pair of equations

$$\begin{aligned} x_1'' + {\omega_0}^2 x_1 &= -K(x_1 - x_2) \\ x_2'' + {\omega_0}^2 x_2 &= -K(x_2 - x_1). \end{aligned}$$

Although there are various mathematical tricks to try for solving this pair of equations, it's really easiest to consider the two physically obvious solutions

(a) (b)

where the pendulums either move in sync (a), with x_1 always equal to x_2 or in anti-sync (b), with x_1 always equal to $-x_2$.

Motion in sync naturally leads to identical equations $x_i'' + {\omega_0}^2 x_i = 0$, so that both pendulums have circular frequency ω_0, with

$$\begin{aligned} x_1(t) &= a\cos\omega_0 t + b\sin\omega_0 t \\ x_2(t) &= a\cos\omega_0 t + b\sin\omega_0 t. \end{aligned}$$

Motion in anti-sync, $x_2 = -x_1$, also leads to identical equations,

$$\begin{aligned} x_1'' + {\omega_0}^2 x_1 &= -K(x_1 - x_2) = -2Kx_1 \\ x_2'' + {\omega_0}^2 x_2 &= -K(x_2 - x_1) = -2Kx_2, \end{aligned}$$

so that both pendulums have circular frequency

$$\omega_0^+ = \sqrt{{\omega_0}^2 + 2K} \qquad \left(\approx \omega_0 + \tfrac{K}{\omega_0} \text{ for } K \ll \omega_0\right),$$

giving

$$\begin{aligned} x_1(t) &= c\cos\omega_0^+ t + d\sin\omega_0^+ t \\ x_2(t) &= -c\cos\omega_0^+ t - d\sin\omega_0^+ t. \end{aligned}$$

It is to be expected that $\omega_0^+ > \omega_0$, since the stretched spring speeds up the oscillations.

These solutions, each with both pendulums having the *same period*, are called *normal* modes of the system. Combinations of the two,

$$\begin{aligned} x_1(t) &= a\cos\omega_0 t + b\sin\omega_0 t + c\cos\omega_0^+ t + d\sin\omega_0^+ t \\ x_2(t) &= a\cos\omega_0 t + b\sin\omega_0 t - c\cos\omega_0^+ t - d\sin\omega_0^+ t, \end{aligned}$$

with four arbitrary constants a, b, c, d, will give us a solution (x_1, x_2) with any desired initial conditions for $x_i(0), x_i'(0)$.

In contrast to the special cases we began with, consider the asymmetrical case where we start with the second pendulum hanging straight down, and the first displaced by a certain amount,

$$x_1(0) = C, \quad x_1'(0) = 0; \qquad x_2(0) = 0, \quad x_2'(0) = 0.$$

These initial conditions give

$$\begin{aligned} x_1(0) &= C = a + c & \qquad x_1'(0) &= 0 = -b\omega_0 - d\omega_0^+ \\ x_2(0) &= \ 0 \ = a - c & x_2'(0) &= 0 = -b\omega_0 + d\omega_0^+, \end{aligned}$$

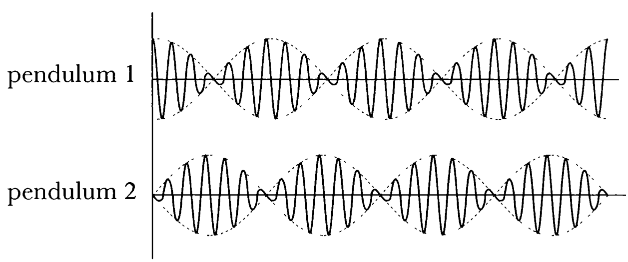
so that $a = c = \frac{1}{2}C$ and $b = d = 0$, and we have

$$\begin{aligned} x_1(t) &= \tfrac{1}{2}C(\cos\omega_0 t + \cos\omega_0^+ t) \\ x_2(t) &= \tfrac{1}{2}C(\cos\omega_0 t - \cos\omega_0^+ t), \end{aligned}$$

which can be written as

$$\begin{aligned} x_1(t) &= C\cos\frac{\omega_0 - \omega_0^+}{2}t \cdot \cos\frac{\omega_0 + \omega_0^+}{2}t \\ x_2(t) &= -C\sin\frac{\omega_0 - \omega_0^+}{2}t \cdot \sin\frac{\omega_0 + \omega_0^+}{2}t. \end{aligned}$$

For "weak coupling", when K/ω_0 is small, and hence $\omega_0^+ - \omega_0$ is small, the first factors in the expressions for $x_1(t)$ and $x_2(t)$ vary slowly with time, and we again have beats. What we observe is that as pendulum 1 begins to swing, the



amplitude of the swing decreases; at the same time, pendulum 2 slowly begins to swing, with increasing amplitude. This process continues until pendulum 1 momentarily comes to a stop, at which point pendulum 2 is now swinging with the original amplitude of pendulum 1 and then the process reverses, until we once again have pendulum 1 swinging with maximum amplitude and pendulum 2 at rest. Energy is continually transferred back and forth from one pendulum to the other, with the total energy remaining the same, except for frictional losses.

Instead of solving our original equations

$$x_1'' + \omega_0^2 x_1 = -K(x_1 - x_2)$$
$$x_2'' + \omega_0^2 x_2 = -K(x_2 - x_1),$$

on the basis of the two physically obvious solutions, we could have used the equivalent mathematical trick of adding and subtracting them to get equations for $z_1 = x_1 - x_2$ and $z_2 = x_1 + x_2$, and then expressing the results back in terms of x_1 and x_2.

With the advantage of hindsight, we might simply look for normal modes to begin with. This is the approach we will use for the case were the pendulums are not identical, so that we have masses m_1 and m_2 and lengths l_1 and l_2.

Setting

$$K_i = \frac{\kappa}{m_i} \qquad \omega_i^2 = \frac{g}{l_i},$$

we now have the equations

$$x_1''(t) + \omega_1^2 x_1(t) = -K_1(x_1 - x_2)$$
$$x_2''(t) + \omega_2^2 x_2(t) = -K_2(x_2 - x_1),$$

and we look for solutions that are the real parts of complex solutions

$$x_1(t) = Ae^{i\lambda t}, \quad x_2 = Be^{i\lambda t},$$

with the same circular frequency ω. We then obtain

$$A(\omega_1^2 - \omega^2 + K_1) = K_1 B$$
$$B(\omega_2^2 - \omega^2 + K_2) = K_2 A.$$

This leads to

$$\text{(a)} \qquad \frac{B}{A} = \frac{\omega_1^2 - \lambda^2 + K_1}{K_1} = \frac{K_2}{\omega_2^2 - \lambda^2 + K_2},$$

and thus to

$$[\lambda^2 - (\omega_1^2 + K_1)][\lambda^2 - (\omega_2^2 + K_2)] = K_1 K_2,$$

which is a quadratic equation in λ^2, having roots that we will call $\omega_1{}^2$ and $\omega_2{}^2$. If r_1 and r_2 are the ratios B/A arising from (a) for $\lambda^2 = \omega_1{}^2$ and $\lambda^2 = \omega_2{}^2$, respectively, then we can write the general solution as

$$\begin{aligned} x_1(t) &= a\cos\omega_1 t + b\sin\omega_1 t + c\cos\omega_2 t + d\sin\omega_2 t \\ x_2(t) &= r_1 a\cos\omega_1 t + r_1 b\sin\omega_1 t + r_2 c\cos\omega_2 t + r_2 d\sin\omega_2 t. \end{aligned}$$

With the same initial conditions as before,

$$x_1(0) = C \qquad x_1{}'(0) = 0 \qquad x_2(0) = 0 \qquad x_2{}'(0) = 0,$$

we again find $b = d = 0$ and

$$a = \frac{r_2}{r_2 - r_1}C, \qquad c = \frac{-r_1}{r_2 - r_1}C,$$

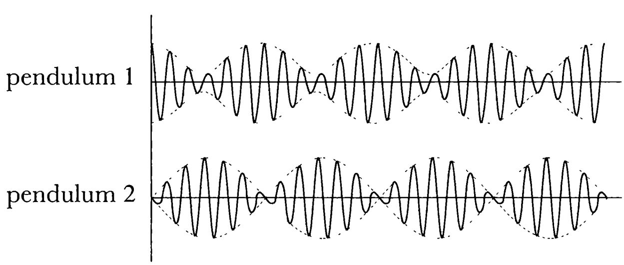
giving two equations of somewhat different form,

$$\begin{aligned} x_1(t) &= \frac{C}{r_2 - r_1}(r_2\cos\omega_1 t - r_1\cos\omega_2 t) \\ x_2(t) &= \frac{C}{r_2 - r_1} r_1 r_2(\cos\omega_1 t - \cos\omega_2 t). \end{aligned}$$

As before, the second can be written

$$x_2(t) = -\frac{2r_1 r_2 C}{r_2 - r_1}\sin\frac{\omega_2 - \omega_1}{2}t \cdot \sin\frac{\omega_1 + \omega_2}{t},$$

with zeros at $t = 2\pi n(\omega_2 - \omega_1)$. But x_1 is not zero at the times when x_2 has a maximum, so the energy is never completely transfered from the first pendulum to the second.



Generalization of our considerations to N harmonic oscillators, all interacting in a linear way, is easy, mainly because we speak only in generalities. We are considering the N equations

$$x_k''(t) + \omega_k{}^2 x_k(t) = \sum_{l=1}^{N} a_{kl} x_l, \qquad k = 1, \ldots, N,$$

where $a_{kl} = a_{lk}$ because of the third law. For a normal mode

$$x_k(t) = c_k e^{i\lambda}, \qquad k = 1, \ldots, N,$$

we must have

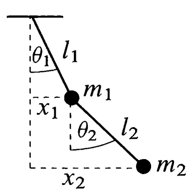
$$-\lambda^2 c_k = \sum_{l=1}^{N} a_{kl} c_l - c_k \omega_k{}^2, \qquad k = 1, \ldots, N,$$

or in matrix form

$$\lambda^2 \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix} = \left[\begin{pmatrix} -\omega_1{}^2 & & 0 \\ & \ddots & \\ 0 & & -\omega_N{}^2 \end{pmatrix} - \begin{pmatrix} a_{11} & \ldots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \ldots & a_{NN} \end{pmatrix} \right] \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix},$$

so that λ^2 is an eigenvalue of the matrix in brackets. Since this matrix is symmetric, it has a basis of eigenvectors with real eigenvalues. Of course, an eigenvalue λ^2 can be negative, so that $\lambda = ib$ for $b > 0$, giving solutions that are multiples of e^b, rather than oscillations. For example, if our original equations on page 302 had $\omega_0 = 1$, but $K = -1$ instead of $K = 1$ (corresponding to a strange spring that keeps pushing the pendulums further and further apart), then our formula for ω_0^+ on page 303 wouldn't make sense. Thus, there is a basis of eigenvectors, with positive eigenvalues each leading to a normal mode, and negative eigenvalues leading to exponentials, which we may regard as a sort of normal mode also. Since the corresponding $(c_1, \ldots, c_N)$ are linearly independent, every solution can be written as a linear combination of these normal modes; some eigenvalues may have multiplicities, leading to normal modes with the same circular frequencies or exponential solutions with the same exponent. In Chapter 12 we will see that this result obtains for much more general systems.

The double pendulum. Another sort of coupling is exhibited by the double pendulum. We will be especially interested in the case where $l_1 = l_2$ and $m_2 \ll m_1$,



where a heavy pendulum, like a chandelier, has a light pendulum of nearly the same length suspended from it. As described in Sommerfeld [2], after a sharp blow to the heavy bob "the light bob will be set in vigorous motion, which suddenly subsides and stays at zero for a short time. At this instant one perceives that the heavy bob, which had previously remained practically at rest, now starts oscillating with noticeable amplitude. This oscillation soon ceases, however, whereupon in its turn the light pendulum again begins to move with considerable vigor, and so forth."

Once again, we are considering only a linear approximation to the actual motion, involving small θ_i; the present analysis, involving some additional complications, may be compared with that to be found at the end of Chapter 12. We have the approximations

$$\theta_1 \approx \sin\theta_1 = \frac{x_1}{l_1}, \quad \theta_2 \approx \sin\theta_2 = \frac{x_2 - x_1}{l_2}$$

$$\sin(\theta_2 - \theta_1) \approx \theta_2 - \theta_1 \approx \frac{x_2 - x_1}{l_2} - \frac{x_1}{l_1}$$

$$\cos\theta_1 \approx 1, \quad \cos\theta_2 \approx 1, \quad \cos(\theta_1 - \theta_2) \approx 1.$$

The lower pendulum is acted on only by gravity, but the upper pendulum is also affected by the tension on the string holding the lower pendulum, which, as on page 210, is

$$m_2 g\cos\theta_2 - m_2 g l_2 \dot{\theta}_2{}^2.$$

Following Sommerfeld, we will drop the term $\dot{\theta}_2{}^2$, as being small of second order, since, supposedly, $\dot{\theta}_2$ is of the same order as θ_2 (compare the treatment at the end of Chapter 12). The horizontal component of this tension $m_2 g\cos\theta_2$ is $-m_2 g\cos\theta_2 \sin(\theta_1 - \theta_2)$, so we have the equations

$$m_1\ddot{x}_1 = -m_1\frac{g}{l_1}x_1 + m_2 g\left(\frac{x_2 - x_1}{l_2} - \frac{x_1}{l_1}\right)$$

$$m_2\ddot{x}_2 = -m_2\frac{g}{l_2}(x_2 - x_1),$$

or, setting $\mu = m_2/m_1$,

$$(*) \qquad \begin{aligned} \ddot{x}_1 + \left(\frac{g}{l_1} + \mu\frac{g}{l_2} + \mu\frac{g}{l_1}\right)x_1 &= \mu\frac{g}{l_2}x_2 \\ \ddot{x}_2 + \frac{g}{l_2}x_2 &= \frac{g}{l_2}x_1. \end{aligned}$$

If we now take the case $l_1 = l_2 = l$, and set $\omega_0 = \sqrt{g/l}$ we obtain

$$\ddot{x}_1 + \omega_0{}^2(1 + 2\mu)x_1 = \mu\omega_0{}^2 x_2$$

$$\ddot{x}_2 + \omega_0{}^2 x_2 = \omega_0{}^2 x_1,$$

similar to the equation in the middle of page 305. When we look for normal modes

$$x_1(t) = Ae^{i\lambda t}, \qquad x_2(t) = Be^{i\lambda t},$$

we now obtain the equations

$$B(\omega_0{}^2 - \lambda^2) = A\omega_0{}^2$$

$$A(\omega_0{}^2(1+2\mu) - \lambda^2) = B\mu\omega_0{}^2,$$

and, as on page 305, we get

$$\frac{B}{A} = \frac{\omega_0{}^2}{\omega_0{}^2 - \lambda^2} = \frac{\omega_0{}^2(1+2\mu) - \lambda^2}{\mu\omega_0{}^2}$$

leading to

$$(\lambda^2 - \omega_0{}^2)^2 + 2\mu\omega_0{}^2(\omega_0{}^2 - \lambda^2) = \mu\omega_0{}^4,$$

a quadratic equation in λ^2 with roots that we will call $\lambda^2 = \omega_1{}^2$ and $\lambda^2 = \omega_2{}^2$. Writing the solutions of the quadratic equation in terms of $\sqrt{\mu}$, with higher powers of $\sqrt{\mu}$ dropped, we find that

$$\omega_1,\ \omega_2 = \omega_0(1 \pm \tfrac{1}{2}\sqrt{\mu}).$$

We now have the general solution

$$x_1(t) = r_1 a \cos\omega_1 t + r_1 b \sin\omega_1 t + r_2 c \cos\omega_2 t + r_2 d \sin\omega_2 t$$

$$x_2(t) = a\cos\omega_1 t + b\sin\omega_1 t + c\cos\omega_2 t + d\sin\omega_2 t$$

if we define r_1 and r_2 as on page 306, and we have approximately

$$r_1 = -\sqrt{\mu}, \quad r_2 = \sqrt{\mu} \qquad \Longrightarrow \qquad r_2 - r_1 = 2\sqrt{\mu}.$$

Now the initial conditions, from a sharp blow to the heavy bob,

$$x_1(0) = 0 \qquad x_1'(0) = C \qquad x_2(0) = 0 \qquad x_2'(0) = 0$$

lead to

$$x_1 = \frac{C}{r_1 - r_2}\left(\frac{r_1}{\omega_1}\sin\omega_1 t - \frac{r_2}{\omega_2}\sin\omega_2 t\right)$$

$$x_2 = \frac{C}{r_1 - r_2}\left(\frac{1}{\omega_1}\sin\omega_1 t - \frac{1}{\omega_2}\sin\omega_2 t\right),$$

which gives, using our approximate values for r_1 and r_2,

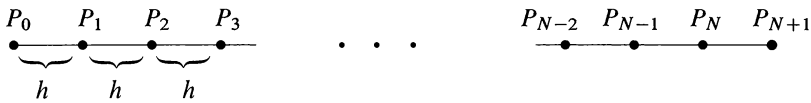
$$x_1' = \frac{C}{2}(\cos\omega_1 t + \cos\omega_2 t)$$

$$x_2' = \frac{C}{2\sqrt{\mu}}(-\cos\omega_1 t + \cos\omega_2 t),$$

with the velocity of the light lower bob $1/\sqrt{\mu}$ times greater than that of the heavy upper bob. Our equations can also be written in a form similar to those on page 304 to show the beats and interchange of energy.

The vibrating string. Returning to the consideration of oscillators coupled by springs, we are going to consider a "continuous" example, essentially involving an (almost) infinite number of coupled oscillators. In complete generality this would encompass the discussion of vibrations in continuous media, but the vibrating string, where both the motions and the couplings of the particles will be quite restricted, may be regarded as a gentle introduction to such problems.

We think of a string under tension (e.g., a violin string) as a system of particles numbered $P_0, P_1, \ldots, P_N, P_{N+1}$, each of mass m, arrayed along a straight line

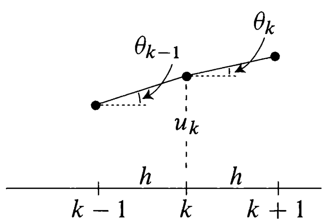


of length L, with the particles P_0 and P_{N+1} restrained to a fixed position. We think of adjacent particles as being connected by identical springs (a rough representation of intermolecular forces) of length $h = L/(N+1)$, greater than the relaxed length of the spring, so that there is a uniform force drawing the particles towards each other, leading to the tension τ of the string.

We assume that our particles move only in a plane, and then ask for the possible motions when each particle is oscillating by a small amount about its initial

point. It might seem rather quixotic to approach this question in terms of coupled oscillators, since Fourier series essentially provides all the information we need about vibrating strings. But, aside from the fact that the initial steps are the same as those we would take to obtain the proper differential equations for the continuous case, there are some interesting differences between the continuous case and the discrete case.

Suppose first that our particles are moving only vertically, with particle P_k at height u_k above the initial horizontal line; we should write $u_k(t)$ to indicate the dependence on the time t, but for the moment we are only considering what happens at one particular time. The line from P_k to P_{k+1} makes an angle θ_k with the horizontal, which is small if our vibrations are small (and N is large).



The first thing we need to note is that the distance from P_k to P_{k+1} is now

$$\frac{h}{\cos\theta_k} = \frac{h}{1 - \frac{1}{2}\theta_k{}^2 + \cdots} = h(1 + \tfrac{1}{2}\theta_k{}^2 + \cdots),$$

so, up to first order, the distance remains the same, and thus, up to first order the tension remains τ throughout.

By the same token, the horizontal component of the force on P_k is

$$-\tau\cos\theta_{k-1} + \tau\cos\theta_k = \tfrac{1}{2}\tau(\theta_{k-1}{}^2 - \theta_k{}^2 + \cdots),$$

which is 0 up to first order, somewhat justifying our original supposition that the particles move only vertically.

Having fudged our way through these preliminaries, we now consider the force on P_k,

$$\begin{aligned} F_k &= -\tau\sin\theta_{k-1} + \tau\sin\theta_k \\ &\approx -\frac{\tau}{h}(u_k - u_{k-1}) + \frac{\tau}{h}(u_{k+1} - u_k). \end{aligned}$$

For the various functions $u_k(t)$, this gives

$$\text{(a)} \qquad u_k{}''(t) = -\frac{\tau}{mh}\big(u_k(t) - u_{k-1}(t)\big) + \frac{\tau}{mh}\big(u_{k+1}(t) - u_k(t)\big),$$

which can be written as

$$u_k{}''(t) + 2\omega_0{}^2 u_k(t) = \omega_0{}^2\big(u_{k-1}(t) + u_{k+1}(t)\big), \qquad \omega_0{}^2 = \frac{\tau}{mh},$$

in the form of the equations on page 307, where now all $\omega_k{}^2 = 2\omega_0{}^2$ and $a_{kl} = \omega_0{}^2$ for $l = k+1$ and $a_{kl} = -\omega_0{}^2$ for $l = k-1$, with all other $a_{kl} = 0$.

Finding normal modes as real parts of

$$u_k(t) = c_k e^{i\omega t} \qquad c_0 = c_{N+1} = 0,$$

by solving the equations

$$(-\omega^2 + 2\omega_0{}^2)c_k - \omega_0{}^2(c_{k-1} + c_{k+1}) = 0$$

or by finding the eigenvalues of the corresponding matrix, does not look like a particularly inviting task. If we were clever enough, we might notice the following:

(1) Our equations can be written as

$$\text{(a}'\text{)} \qquad \frac{c_{k-1} + c_{k+1}}{c_k} = \frac{-\omega^2 + 2\omega_0{}^2}{\omega_0{}^2},$$

so that the left side should be constant for a normal mode.

(2) We have

$$\sin(k-1)\theta + \sin(k+1)\theta = 2\sin k\theta \cos\theta$$

or

$$\frac{\sin(k-1)\theta + \sin(k+1)\theta}{\sin k\theta} = 2\cos\theta,$$

so we will have this constant relation if we take

$$c_k = \sin k\theta$$

for any θ. Since we want $c_0 = c_{N+1} = 0$, we just need $(N+1)\theta$ to be a multiple of π, leading to

$$\theta = \frac{n\pi}{N+1}, \qquad n = 1, 2, 3, \ldots, N,$$

and thus

$$\frac{c_{k-1} + c_{k+1}}{c_k} = 2\cos\left(\frac{n\pi}{N+1}\right).$$

By (a′), the corresponding solution $\omega = \omega_n$ is given by

$$2\cos\left(\frac{n\pi}{N+1}\right) = \frac{-\omega_n{}^2 + 2\omega_0{}^2}{\omega_0{}^2},$$

and thus

$$\begin{aligned}
\omega_n{}^2 &= 2\omega_0{}^2\left[1 - \cos\left(\frac{n\pi}{N+1}\right)\right] \\
&= 4\omega_0{}^2 \sin^2\left(\frac{n\pi}{2(N+1)}\right), \\
\omega_n &= 2\omega_0 \sin\left(\frac{n\pi}{2(N+1)}\right).
\end{aligned}$$

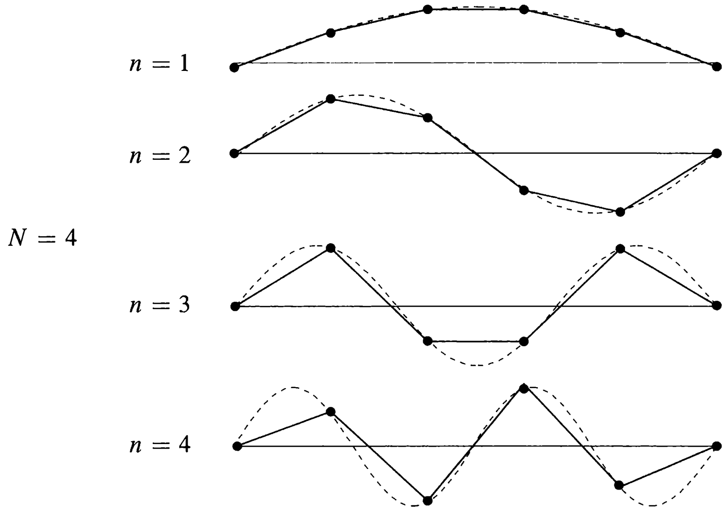
The general real solution of our equations $u_k(t) = c_k e^{i\omega t}$ for this n will be

$$(*_n) \qquad u_k = \sin\left(\frac{kn\pi}{N+1}\right)[a\cos\omega_n t + b\sin\omega_n t].$$

For the particle P_k, at distance $x = kh = kL/(N+1)$ from the initial point, the sine factor is just

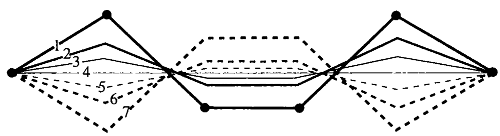
$$\sin\left(\frac{n\pi x}{L}\right),$$

so for $t = 0$, the particles all lie on a times this sine curve. In the lowest mode the



particles are all on the same side of the horizontal axis; in the highest mode they alternate sides.

The factor $a \cos \omega_n t + b \sin \omega_n t = A \cos(\omega_n t + \phi)$ for some A and ϕ then indicates how this configuration changes amplitude with time. In the figure below, where we have taken $\phi = 0$, amounting merely to a shift of the time



coordinate, we see 7 stages of a half-cycle, with the particles moving from their positions at $t = 0$ to their positions at $t = \pi/\omega_n$.

If our figure had an additional t axis, one could "see" this motion more clearly; over the period $T_n = 2\pi/\omega_n$, the whole configuration would pass through the positions 1–7 and then back to position 1. In the previous figure, the sine curves and their inscribed piecewise-linear curves are plotted against distance, rather than time, and the corresponding "period", during which a curve goes through a complete cycle, is now called the *wave length* λ. The reciprocal $1/\lambda$, the *wave number*, is the analogue of the frequency $\nu = 1/T$, so the analogue of the circular frequency $\omega = 2\pi\nu$ is $\kappa = 2\pi/\lambda$, the *angular wave number*.

Our solution $(*_n)$ has a wave length $\lambda_n = 2L/n$, so the corresponding angular wave number is

$$\kappa_n = \frac{2\pi}{\lambda_n} = \frac{n\pi}{L} = \frac{n\pi}{(N+1)h}$$

and our formula for ω_n can be written

$$\text{(A)} \qquad \omega_n = 2\sqrt{\frac{\tau}{mh}} \cdot \sin\left(\frac{\kappa_n h}{2}\right).$$

To approximate a continuous string of density ρ, we want to choose m and h so that $m/h = \rho$, and thus $\sqrt{\tau/mh} = \sqrt{\tau/\rho} \cdot 1/h$, and we can write

$$\omega_n = 2\sqrt{\frac{\tau}{\rho}} \cdot \frac{1}{h} \cdot \sin\left(\frac{\kappa_n h}{2}\right).$$

We normally think of having $N \gg n$, in which case $\kappa_n h$ will be very small. Since $1 > \sin\theta/\theta \to 1$ as $\theta \to 0$, we have

$$(\text{A}_{\approx}) \qquad \omega_n \lessapprox 2\sqrt{\frac{\tau}{\rho}} \cdot \frac{1}{h} \cdot \frac{k_n h}{2} = \sqrt{\frac{\tau}{\rho}} \cdot \kappa_n,$$

and we might expect that we will have exact equality for the continuous case, which we want to consider next, naturally hoping that everything will turn out to be easier, and we won't have to be so clever.

In equation (a) on page 311 we consider $u(x,t)$ for arbitrary x in an interval $[0, L]$, and see what we get as $h \to 0$, or $N \to \infty$, keeping $m/h = \rho$. Instead of considering particles P_{k-1}, P_k, P_{k+1}, we simply consider arbitrary points $x-h,\ x,\ x+h \in [0, L]$, and write

$$\frac{\partial^2 u}{\partial t^2}(x,t) = \frac{\tau}{\rho} \cdot \lim_{h\to 0} \frac{u(x+h,t) + u(x-h,t) - 2u(x,t)}{h^2}.$$

But an easy application of Taylor's theorem shows that

$$\lim_{h\to 0} \frac{[f(x+h) - f(x-h) - 2f(x)]}{h^2} = f''(x),$$

so we arrive at the classical 1-dimensional wave equation,

$$\frac{\partial^2 u}{\partial t^2}(x,t) = v^2 \frac{\partial^2 u}{\partial x^2}(x,t), \qquad v^2 = \frac{\tau}{\rho}.$$

Now we have to be clever again, but we only need the standard cleverness of "separation of variables", seeking a solution of the form

$$u(x,t) = X(x)T(t).$$

We find that $X(x)T''(t) = v^2 X''(x)T(t)$ or

$$\frac{T''(t)}{T(t)} = v^2 \frac{X''(x)}{X(x)}.$$

So the two sides must be a constant K. The equation $v^2 X''/X = K$ won't have a non-zero solution with $X(0) = X(L) = 0$ for $K \geq 0$, so we can set $K = -\omega^2$, giving

$$\begin{aligned} T''(t) + \omega^2 T &= 0 \\ X''(x) + \frac{\omega^2}{v^2} X &= 0. \end{aligned}$$

The solutions to the first are simply

$$T(t) = a\cos(\omega t) + b\sin(\omega t),$$

and for the second we will write

$$X(x) = \sin\left(\frac{\omega x}{v} + \phi\right).$$

From $X(0) = 0$ it follows that $\sin\phi = 0$, so $X(x) = \pm\sin(\omega x/v)$, and then from $X(L) = 0$ that ω must be one of the numbers

$$\omega_n = \frac{n\pi v}{L}, \qquad n = 1, 2, 3, \ldots;$$

for each such ω_n we have the normal modes

$$(**_n) \qquad u(x,t) = \sin\left(\frac{n\pi}{L}x\right)[a_n \cos\omega_n t + b_n \sin\omega_n t],$$

for constants a_n and b_n. Naturally, we can use the modes with $b_n = 0$ to work backwards to the clever choice (a′) on page 312.

The Fourier series for any continuous curve f with $f(0) = f(L) = 0$ is an (infinite) linear combination of the terms $\sin\frac{n\pi}{L}x$, so any desired initial condition $u(x,0)$ can be determined by an appropriate linear combination of these normal modes by appropriate choice of the a_n. Moreover, any desired initial condition $\partial u/\partial t(x,0)$ can be obtained also, by the proper choice of the b_n. So we have the general solution for our equation with the initial conditions $u(0) = u(L) = 0$.

Notice that now the relationship between $\omega_n = n\pi v/L$ and $\kappa_n = n\pi/L$, with $v^2 = \tau/\rho$, is simply

$$\omega_n = \sqrt{\tau/\rho}\cdot\kappa_n.$$

Comparing with ($A_\approx$) on page 314, we would be led to conclude that for a physical string, ω_n should be a little smaller. In actuality, all sorts of other factors may come into play. In our model of a string as a collection of N particles, the only force considered was the tension resulting from the increase of distance between particles, with no account taken of the force resulting from the slight difference between the angles that a particle makes with its neighbors (basically, the "stiffness" of the string). For a piano string, this turns out to be proportional to ${\kappa_n}^4$, and the relation between ω_n and κ_n is approximately

$${\omega_n}^2 = \sqrt{\tau/\rho}\cdot{\kappa_n}^2 + \alpha{\kappa_n}^4,$$

for a positive constant α, a measure of the stiffness. So in this case, ω_n grows faster than the idealized case (the "harmonics" of a piano string are a little sharp, rather than a little flat).

Although our analysis of the wave equation followed an obvious path from the analysis of the discrete case, a completely different approach is also possible. With a multiplicative change of coordinates, and t replaced by y, the classical 1-dimensional wave equation can be written as

$$u_{xx} = u_{yy},$$

with subscripts now denoting partial derivatives. This is actually the prototypical example of a hyperbolic equation, which has the "normal form"

$$u_{xx} - u_{yy} + \cdots = 0$$

where $\cdots$ denotes terms not involving second derivatives. There is an alternative normal form

$$u_{xy} + \cdots = 0,$$

corresponding to the possibility of writing the equation for a hyperbola in the form $xy = 1$, which can be obtained simply by considering the function

$$v(\xi,\eta) = u\left(\frac{\xi+\eta}{2}, \frac{\xi-\eta}{2}\right).$$

In the study of partial differential equations, rather than seeking solutions for functions with prescribed values at the end points of an interval, we more often consider arbitrary initial values for a hyperbolic equation on an open interval, or even the whole x-axis, and in the case of the wave equation, this alternative form gives a complete solution: for the function v just defined, the wave equation for u becomes simply

$$v_{\xi\eta} = 0,$$

with the general solution

$$v(\xi,\eta) = f(\xi) + g(\eta),$$

leading to

$$u(x,y) = v(x+y, x-y) = f(x+y) + g(x-y)$$

for arbitrary functions f and g, and it is not hard (Problem 5) to determine the functions f and g in terms of the initial conditions, and thus write out $u(x, y)$ as an explicit formula.

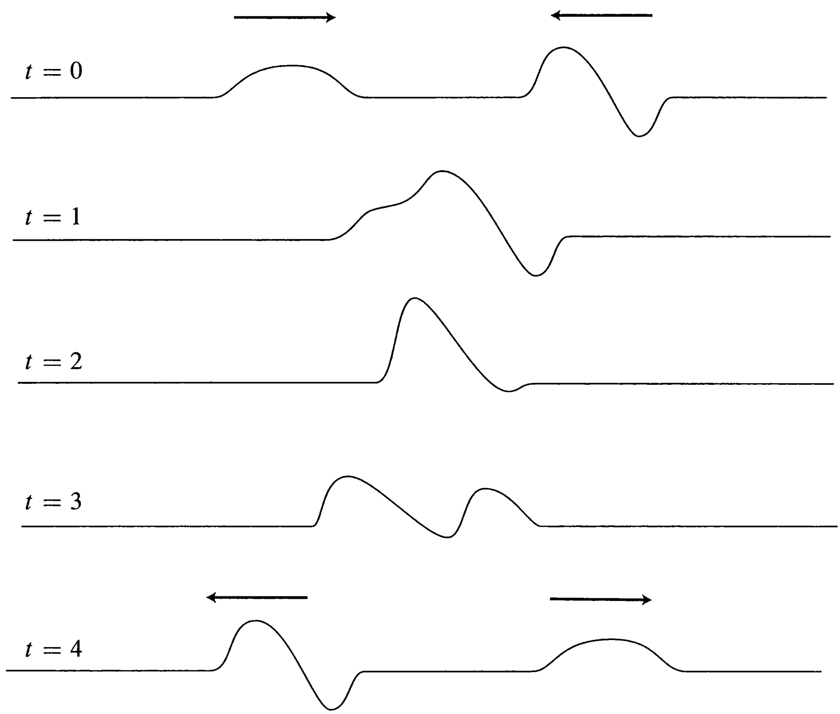
When we write our equation with the additional constant,

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial x^2},$$

we have

$$u(x,t) = f(x+vt) + g(x-vt),$$

which is the sum of two arbitrary "waves", the first wave moving to the left with velocity v, the second moving to the right with velocity v.



Although we can indeed express any "standing wave" $u(x,t)$ with $u(0,t) = 0 = u(L,t)$ in this way (Problem 6), the "moving waves" defined by this more general solution are the ones that play the most important role in all sorts of physical phenomenon. Problem 5 provides some information about these waves that will be used later on in various places in this volume, but a more complete discussion will have to be deferred for now.

ADDENDUM 8A

ABEL'S INTEGRAL EQUATION

As opposed to calculus of variation problems, where the Euler equations give a straightforward way of finding necessary conditions for a solution, with questions of sufficiency leading to more involved considerations, most solutions of the tautochrone problem demonstrate that the cycloid is a solution without establishing that it is the only one.

In 1823, Abel considered a more general problem, one of the first examples of an integral equation. If a particle of mass $m = 1$ starts from rest sliding down the upside-down cycloid at a point of height Y, with potential energy gY, then conservation of energy shows that when it reaches a point of height y we have

$$\frac{ds}{dt} = -\sqrt{2g(Y-y)}.$$

If $\tau(y)$ is the time at which the particle sliding along the curve has height y, and $\sigma(y)$ is the length of the curve from the bottom to the point at height y, then

$$-\sqrt{2g(Y-y)} = \frac{ds}{dt} = \frac{ds/dy}{dt/dy} = \frac{\sigma'(y)}{\tau'(y)},$$

so the time $\phi(Y)$ of descent from height Y to the bottom, $y = 0$, satisfies

$$\phi(Y) = \int_Y^0 \tau'(y)\,dy = -\frac{1}{\sqrt{2g}}\int_Y^0 \frac{\sigma'(y)\,dy}{\sqrt{Y-y}} = \frac{1}{\sqrt{2g}}\int_0^Y \frac{\sigma'(y)\,dy}{\sqrt{Y-y}}.$$

The tautochrone problem involves the case where the function ϕ is a simply a constant, $\phi(Y) = T$, and Abel considered the more general problem, for a given function ϕ, of solving for the function σ satisfying

$$\phi(Y) = \int_0^Y \frac{\sigma'(y)\,dy}{\sqrt{Y-y}} \tag{1}$$

or even more generally

$$\phi(Y) = \int_0^Y \frac{\sigma'(y)\,dy}{(Y-y)^n} \qquad 0 < n < 1.$$

Abel showed that the solution is

$$\sigma(y) = \frac{\sin n\pi}{\pi} \int_0^y \frac{\phi(Y)\,dY}{(y-Y)^{1-n}},$$

which for the case of $n = \frac{1}{2}$, or $\sqrt{Y-y}$, gives

$$\sigma(y) = \frac{1}{\pi} \int_0^y \frac{\phi(Y)\,dY}{\sqrt{y-Y}}. \tag{2}$$

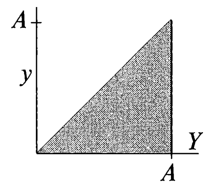
Abel's proof (Abel [1]) is a complex chain of computations starting from the Γ function! (pun intended); nowadays it would be written in terms of the Laplace transform of the convolution of two functions. An annotated English translation can be found in Smith [1], which also explains how an earlier paper by Abel gave a solution in terms of "fractional calculus" (see Miller and Ross [1] for an introduction to this subject).

A simple proof,[1] sticking to the case $\sqrt{Y-y}$ for convenience, may be given as follows. Consider

$$\int_0^A \frac{\phi(Y)\,dY}{\sqrt{A-Y}} = \int_0^A \left(\int_0^Y \frac{\sigma'(y)\,dy}{\sqrt{Y-y}} \right) \frac{dY}{\sqrt{A-Y}}.$$

This is the integral of the function

$$\frac{\sigma'(y)}{\sqrt{(Y-y)(A-Y)}}$$ over a triangular region



expressed as an iterated integral, first along the y-axis, and then along the Y-axis. Reversing the order of integration, we get

$$\int_0^A \frac{\phi(Y)\,dY}{\sqrt{A-Y}} = \int_0^A \sigma'(y) \left(\int_y^A \frac{dY}{\sqrt{(Y-y)(A-Y)}} \right) dy.$$

[1] Adopted from Landau and Lifschitz [1], where it is used—though without any mention of Abel's equation—to consider a question about oscillation periods, which appears here as Problem 4.

The inner integral can be calculated to have the value π, whatever values y and A may have.[1] This means that the right side is simply $\pi\sigma(A)$. Finally replacing A by y, we get the desired result (2).

For the tautochrone problem, $\phi = T = 1/\sqrt{2g}$, Abel's solution gives

$$\sigma(y) = \frac{\sqrt{2g}\,T}{\pi}\int_0^y \frac{dY}{\sqrt{y-Y}} = \frac{2\sqrt{2g}\,T}{\pi}\sqrt{y}$$

$$\sigma'(y) = \frac{\sqrt{2g}\,T}{\pi}\frac{1}{\sqrt{y}}.$$

Assuming our curve starts at $(0,0)$, we can then write

$$\sigma'(y) = \frac{ds}{dy} = -\sqrt{1+\left(\frac{dx}{dy}\right)^2}$$

$$\frac{dx}{dy} = \sqrt{\sigma'^2(y)-1}$$

$$x = \int_0^y \sqrt{\frac{2gT^2}{\pi^2 y}-1}\;dy + 0 \qquad \text{(since } x=0 \text{ at } y=0\text{)}.$$

The substitution

$$y = 2a\sin^2 u, \qquad a = \frac{gT^2}{\pi^2}$$

changes this to

$$x = 4a\int_0^{\vartheta}\cos^2 u\,du \qquad \text{for } \vartheta = \arcsin\sqrt{y/2a},$$

giving

$$x = 2a\left(\vartheta + \tfrac{1}{2}\sin 2\vartheta\right)$$
$$y = 2a\sin^2\vartheta,$$

which is a cycloid parameterized by $\theta = 2\vartheta$.

---

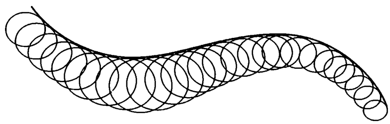
[1] Numerous approaches are possible, including an elementary integration. Note that the substitution $u = (Y-y)/(Y-A)$ reduces the integral to $\int_0^1 du/(u(1-u))^{\frac{1}{2}} = \int_0^1 u^{-\frac{1}{2}}(1-u)^{-\frac{1}{2}}\,du$, and $\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,du = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$, which is the formula that Abel starts from, and which could be used to adapt this proof for the general case.

ADDENDUM 8B

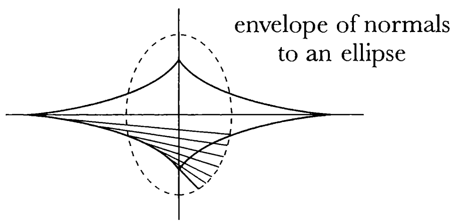
ENVELOPES

Envelopes, which have been mentioned in this chapter, will play an important role in Chapter 15 and later chapters, so we will give a brief discussion of this topic, which is often slighted in modern treatments of differential geometry.

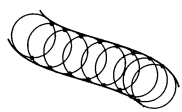
We'll begin by considering the simple case of a 1-parameter family $\bar{\alpha}$ of curves in the plane, given by $\bar{\alpha}(u) = t \mapsto \alpha(u,t)$ for some C^∞ function $\alpha\colon [0,1] \times [0,1] \to \mathbb{R}^2$. An **envelope** of this family is defined to be a curve c which is *not* a member of this family but which is tangent to some member of the family at every point. Unfortunately, it often turns out that the envelope of a perfectly



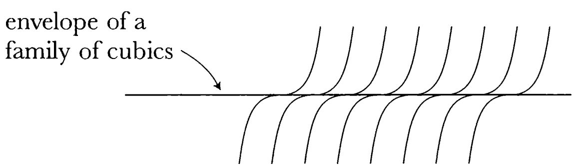
nice family of curves has a cusp or something worse, as shown below, but we won't be worrying too much about this.



The classical way of finding the envelope of α was very geometric. For each u, we let $c(u)$ be the limit, as $\varepsilon \to 0$, of the intersection of $\bar{\alpha}(u)$ and $\bar{\alpha}(u+\varepsilon)$: the

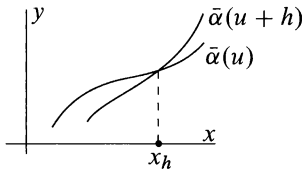


envelope consists of the "intersections of members of the family with another member infinitely close to it". The picture below shows that this idea can run into some serious difficulties. Nevertheless, it often works out rather well in



particular cases, and even in the general case it leads us to the proper analytic condition, when we argue as follows.

Let us consider first the case where our curves $\bar{\alpha}(u)$ are all expressed as the graphs of functions; thus there is a function $(u,x) \mapsto f(u,x)$ such that $\alpha(u,t) = (t, f(u,t))$. Suppose that the curve $\bar{\alpha}(u)$ and the curve $\bar{\alpha}(u+h)$



intersect at the point

$$(x_h, f(u,x_h)) = (x_h, f(u+h, x_h)).$$

Then we have

$$0 = \frac{f(u+h,x_h) - f(u,x_h)}{h}.$$

Assuming that x_h approaches a number $x(u)$ as $h \to 0$, we find that $x(u)$ must be a point for which

$$(*) \qquad D_1 f(u, x(u)) = 0.$$

If we find the points $x(u)$ for all u, then the envelope should be the curve consisting of all points $(x(u), f(u,x(u)))$.

If we are given a general family $\bar{\alpha}$, not necessarily expressed as graphs of functions, then we can introduce the function f in two steps. We first determine $t(u,x)$ so that

$$(1) \qquad \alpha_1(u, t(u,x)) = x,$$

and then define

$$(2) \qquad f(u,x) = \alpha_2(u, t(u,x)).$$

Then equation $(*)$ becomes

$$(3) \qquad 0 = D_1\alpha_2(u,t(u,x)) + D_2\alpha_2(u,t(u,x)) \cdot D_1 t(u,x),$$

while equation (1) gives

$$D_1\alpha_1(u,t(u,x)) + D_2\alpha_1(u,t(u,x)) \cdot D_1 t(u,x) = 0,$$

$$D_1 t(u,x) = -\frac{D_1\alpha_1}{D_2\alpha_1 1}(u,t(u,x)).$$

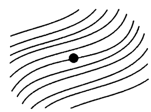
Substituting this into (3), we obtain

$$[D_1\alpha_2 \cdot D_2\alpha_1 \; - \; D_1\alpha_1 1 \cdot D_2\alpha_2 2](u,t(u,x)) = 0.$$

Thus we find that the envelope should consist of points $\alpha(u,t)$ where (u,t) satisfies

$$(**) \qquad \det\big(D_i\alpha_j(u,t)\big) = 0.$$

Now even without resorting to the motivating geometric construction, it is clear that if there is an envelope of the family $\bar{\alpha}$, then it must be a subset of the points $\alpha(u,t)$ for which (u,t) satisfies $(**)$. For, if the determinant in $(**)$ is non-zero, then α is an immersion at (u,t), and the curves $\bar{\alpha}(u)$ form a foliation of a neighborhood of $\alpha(u,t)$; consequently, the only curve through $\alpha(u,t)$ which



is tangent to some curve of the family at each point is $\bar{\alpha}(u)$ itself, which means that $\alpha(u,t)$ cannot be a point of an envelope. Problem 2 (d) gives an example of the use of this criterion.

Envelopes of a family of surfaces in $\mathbb{R}^3$, which will be our main interest later on, are obviously best handled in this general way, and we will basically need the concept, rather than any particular methods of calculation. The special case of the envelope of a family of planes in $\mathbb{R}^3$ is discussed in DG, Vol. 3, Chap. 3, Addendum, from which most of the foregoing material was taken.

ADDENDUM 8C

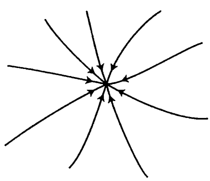
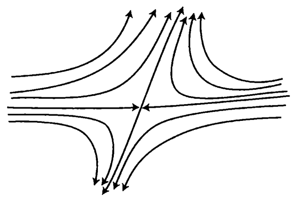
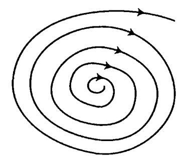
STABILITY OF SOLUTIONS OF DIFFERENTIAL EQUATIONS

To describe the question of stability for solutions of a first order equation, we'll begin by discussing the 2-dimensional case, which is the one we will apply in Addendum 10A, in order to draw some illustrative pictures, merely indicating the situation, without giving proofs, for which a reference is given at the end.

So we consider a first order equation

$$(c_1{}', c_2{}')(t) = \big(f(c_1(t), c_2(t)),\ g(c_1(t), c_2(t))\big)$$

corresponding to the vector field X with components $(f(x,y), g(x,y))$ at (x,y), where the integral curves near a 0 can have numerous different arrangements.

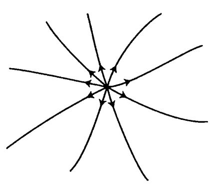
In order to get information about the integral curves of the vector field X near the 0 point, we "linearize", by considering the Jacobian matrix

$$A = \begin{pmatrix} \partial f/\partial x & \partial g/\partial x \\ \partial f/\partial y & \partial g/\partial y \end{pmatrix}$$

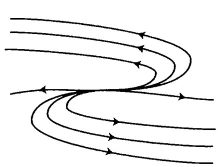
at the point in question, which we'll consider to be the origin, for convenience. The determination of stability will not be affected by a linear change of coordinates in the plane, which will change A to a conjugate BAB^{-1}, so it's useful to examine how the possible canonical forms for A correspond to pictures of the integral curves.

If the characteristic polynomial of A has a double root λ, then we have two possible canonical forms. The accompanying pictures show the case $\lambda > 0$,

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$



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



while the arrows will be reversed for $\lambda < 0$.

For the first canonical form, our equation is approximated by

$$\begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

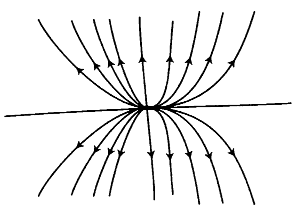
having solutions with components $ae^{\lambda t}$, $be^{\lambda t}$. The solutions near the 0 solution at $(0, 0)$ will have the basic characteristics of this solution: if $\lambda > 0$, then in time $1/\lambda$ they increase by a factor of about e, so the 0 solution is not stable; on the other hand, it is stable for $\lambda < 0$.

For the second canonical form our equation is approximated by

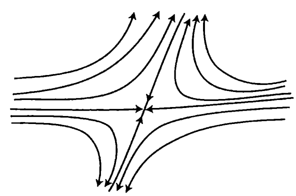
$$\begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$
$$= \begin{pmatrix} \lambda c_1 + c_2 \\ \lambda c_2 \end{pmatrix},$$

having solutions with components $ate^{\lambda t} + be^{\lambda t}$, $be^{\lambda t}$, and again if $\lambda > 0$, then one component of solutions near the 0 solution increase by a factor of about e in time $1/\lambda$.

When there are distinct real roots λ_1 and λ_2, we have the following pictures,



$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$



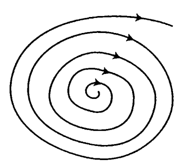
$0 < \lambda_1 < \lambda_2$ $\qquad\qquad$ $\lambda_2 < 0 < \lambda_1$

with the arrows reversed when the signs of λ_1 and λ_2 are reversed. In the second case (a "saddle point"), either c_1 or c_2 will involve a positive exponential, so that the 0 solution is always unstable.

When A is singular ($\lambda = 0$ for a double root, or at least one of λ_1 and λ_2 =0 for distinct roots), no conclusion is possible.

When there are complex roots $u + iv$ and $u - iv$, the canonical form is

$$\begin{pmatrix} u & v \\ -v & u \end{pmatrix} \quad u > 0$$



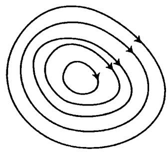
with the arrows in the picture reversed for $u < 0$. Solving the equations $(c_1{}', c_2{}') = (uc_1 + vc_2, -vc_1 + uc_2)$ by setting

$$z = c_1 + ic_2 \implies z' = (u - iv)z \implies z = e^u e^{-iv},$$

we have $(c_1(t), c_2(t)) = e^{ut}(\cos vt, \sin vvt)$, and the 0 solution is unstable for $u > 0$, and stable for $u < 0$.

The case $u = 0$, where we have pure complex roots iv and $-iv$, is again indeterminate. Although we might have a picture with closed curves, the curves

$$\begin{pmatrix} 0 & v \\ -v & 0 \end{pmatrix}$$



could also spiral outwards. Since our approximating equation is $c_1{}' = vc_2$, $c_2{}' = -vc_1$, with solutions $c_1(t) = \sin(vt)$, $c_2(t) = \cos(vt)$, when we do have closed curves their period is close to $2\pi/v$.

The general result, in n dimensions, is that stability is assured if the real parts of the eigenvalues of the Jacobian matrix at the 0 of the vector field are all negative, and instability will always occur if the real part of any eigenvalue is positive. A modern proof may be found in Palais and Palais [1; pp. 47, 53ff.], where the result is first proved for linear equations and then extended to the general case by approximating to a linear equation.

PROBLEMS

1. The parameterization of the upside-down cycloid shows that

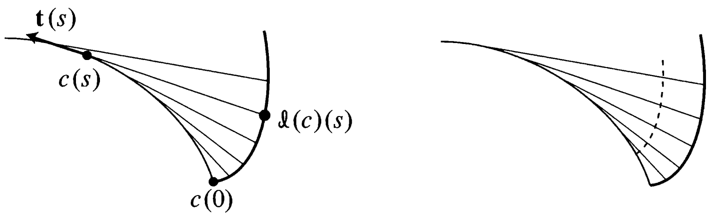
$$\frac{dy}{dx} = \frac{dy}{d\theta} \div \frac{dx}{d\theta} = \frac{-\sin\theta}{1-\cos\theta},$$

or, for rotation from the bottom, so that we replace θ by $\pi+\theta$, the tangent line makes an angle α with

$$\tan\alpha = \frac{\sin\theta}{1+\cos\theta}.$$

Check that $\alpha = \theta/2$, though it is not so clear how one would notice this!

2. We described the involute of a curve in terms of pulling out a string wound around the curve, and the actual definition is simply a translation of this descrip-



tion. If c is a curve parameterized by arclength s, with unit tangent vector $\mathbf{t}(s)$, we define $\mathcal{I}(c)$, the *involute* of c, by

$$\mathcal{I}(c)(s) = c(s) - s\mathbf{t}(s).$$

The curve c really has infinitely many different involutes, depending on which point we choose as $c(0)$; they are "parallel" curves, intersecting the tangents at a constant distance from each other.

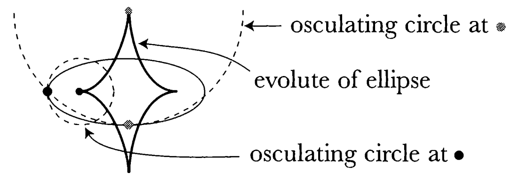
(a) Show that for $\gamma = \mathcal{I}(c)$ we have $\langle \gamma'(s), \mathbf{t}(s)\rangle = 0$, as drawn in the picture. Note that s is generally not the arclength parameterization for γ.
(b) For the cycloid

$$\begin{aligned} x &= \theta - \sin\theta \\ y &= 1 - \cos\theta, \end{aligned}$$

(taking $a = 1$ for simplicity), check that the tangent vector has length $2\sin(\theta/2)$. Choosing an initial point so that we simply have $s(\theta) = -4\cos(\theta/2)$, check that the involute is another cycloid of the same size and shape.

(c) Although not relevant to the discussion of the cycloidal pendulum, another property of involutes may be mentioned. Recall[1] that for an arclength parameterized curve c with tangent vector $\mathbf{t}$ and normal vector $\mathbf{n}$, the osculating circle of c at s is the circle tangent to c at $c(s)$ with radius $1/\kappa(s)$, where $\kappa = \pm|c''(s)|$ is the curvature; the center of the osculating circle is thus at $c(s) + \mathbf{n}(s)/\kappa(s)$. We define $\mathcal{E}(c)$, the *evolute* of c, as the curve traced out by these centers,

$$\mathcal{E}(c)(s) = c(s) + \frac{\mathbf{n}(s)}{\kappa(s)}.$$



For $\gamma = \mathcal{I}(c)$, consider $\mathcal{E}(\gamma) = \mathcal{E}(\mathcal{I}(c))$, defined by

$$\begin{aligned} \mathcal{E}(\gamma)(s) &= \gamma(s) + \frac{\mathbf{n}_\gamma(s)}{\kappa_\gamma(s)} \\ &= c(s) - s\mathbf{t}(s) + \frac{\mathbf{n}_\gamma(s)}{\kappa_\gamma(s)}. \end{aligned}$$

Show that this is simply $c(s)$ (remember that s is not the arclength parameterization of γ), so that we have

$$\mathcal{E}(\mathcal{I}(c)) = c.$$

Of course, it is great fun to try to prove all this geometrically, as Huygens did, with the center of the osculating circle defined as the limiting position of normals near a point, or alternatively, with the evolute defined as the envelope of the normals to the curve.

(d) For an arclength parameterized curve $c(s) = (c_1(s), c_2(s))$, with normal $\mathbf{n}(s) = (-c_2{}'(s), c_1{}'(s))$, use $(**)$ on page 322 to show that the envelope of the normals consists of points $c(s) + \mathbf{n}(s)/\kappa(s)$.

3. (a) Multiplying the equation $x'' + \omega^2 x = 0$ by x', conclude that

$$x'^2 + \omega^2 x^2 = 2E$$

for a constant E.

(b) Taking the initial conditions $x(0) = a$, $x'(0) = 0$, deduce that

$$\frac{dx}{dt} = \omega\sqrt{a^2 - x^2},$$

---

[1] E.g., DG, Vol. 2, Chap. 1.

and then that

$$\omega t = \arcsin\left(\frac{a}{x} - \frac{\pi}{2}\right) \implies x = a\sin\left(\omega t + \frac{\pi}{2}\right) = a\cos\omega t.$$

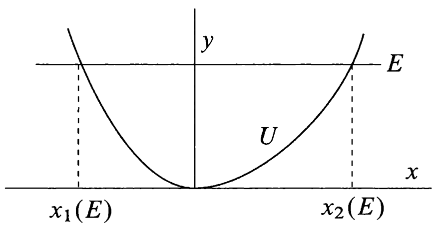
4. In Addendum 4B, we considered the case of one-dimensional motion with a potential energy function U [the $\tilde{V}$ that arose in that Addendum], so that we have the equation of motion

$$\tfrac{1}{2}x'^2 - U(x) = E.$$

If U has a relative minimum point, where the energy is E_0 say, then for a range of $E > E_0$, the solution with energy E will oscillate between two values $x_1(E)$ and $x_2(E)$, and the period $P(E)$ of this oscillation is given by

$$(*) \qquad P(E) = \int_{x_1(E)}^{x_2(E)} dt = \sqrt{2}\int_{x_1(E)}^{x_2(E)} \frac{dx}{\sqrt{E - U(x)}}.$$

(a) For convenience, assume that the minimum value of U occurs at $x = 0$, and has the value $U(0) = 0$; the functions x_1 and x_2 are then the inverses of U



restricted to the positive and negative x axes, and for the above integral the substitution $y = U(x)$ can be written as $x = x_2(y)$ for $x > 0$ and $x = x_1(y)$ for $x < 0$. Conclude that

$$P(E) = \sqrt{2}\int_0^E \frac{\sigma'(y)\,dy}{\sqrt{E-y}} \qquad \text{for } \sigma = x_2{}' - x_1{}'.$$

(b) Use Abel's result from Addendum A to show that

$$x_2(y) - x_1(y) = \frac{1}{\sqrt{2}\pi}\int_0^y \frac{P(Y)\,dY}{\sqrt{y-Y}}.$$

(c) Conclude that we can choose the shape of the graph of U arbitrarily for $x > 0$, and then determine the shape for $x < 0$ so that $(*)$ holds for the given function P.

(d) For a symmetric graph, $U(x) = U(-x)$, we must have

$$x_2(y) = -x_1(y) = \frac{1}{2\sqrt{2}\,\pi}\int_0^y \frac{P(Y)\,dY}{\sqrt{y-Y}}.$$

502 **5.** (a) For the solution

$$u(x,y) = f(x+vt) + g(x-vt)$$

on page 317, with given initial conditions

$$\begin{aligned} u(x,0) &= \phi(x) \\ u_t(x,0) &= \psi(x), \end{aligned}$$

find equations for $f'(x)$ and $g'(x)$, and then for $f(x)$ and $g(x)$ in terms of ϕ and integrals involving ψ, and conclude that we have *d'Alembert's formula*

$$u(x,t) = \frac{\phi(x+vt)+\phi(x-vt)}{2} + \frac{1}{2v}\int_{x-vt}^{x+vt} \psi(s)\,ds.$$

(b) Consider the wave equation only for positive x and t:

$$u_{tt} - v^2 u_{xx} \qquad x,t > 0$$

with initial conditions

$$\begin{aligned} &u(x,0) = \phi(x), \qquad u_t(x,0) = \psi(x) \quad \text{for } x > 0 \\ &u(0,t) = 0 \quad \text{for } t \geq 0, \end{aligned}$$

where $\phi(0) = 0 = \psi(0)$. Suppose we extend ϕ and ψ as *odd* functions on $\mathbb{R}$. Show that the solution is again given by d'Alembert's formula.

6. For a solution $u(x,t) = f(x+vt) + g(x-vt)$ of the wave equation, we have $u(x,0) = f(x) + g(x)$.

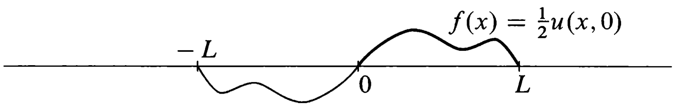
(a) If $u(0,t) = 0$ for all t, then $g(x) = -f(-x)$, so that

$$u(x,t) = f(x+vt) - f(-x+vt).$$

(b) To obtain

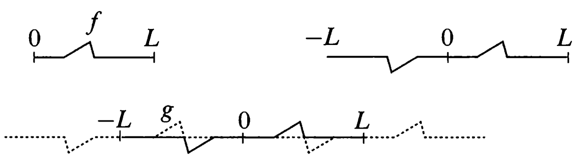
$$u(x,0) = f(x) - f(-x),$$

we can take $f(x) = \frac{1}{2}u(x,0)$ on $[0, L]$, and then use this equation to define f on $[-L, 0]$.

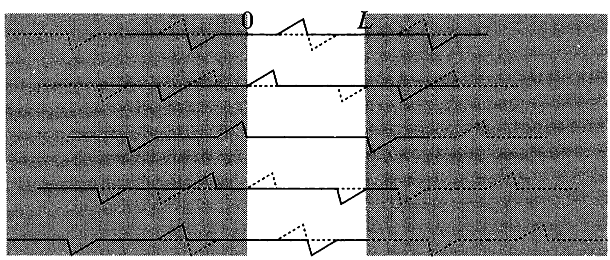


(c) If we extend f to be periodic, with period $2L$ on the whole line, then we also have the other boundary condition $u(L,t) = 0$ for all t.

If we start with the function f shown below, and then extended it in this way,



and then define g, then we can trace the movement of f left and g right. Since the actual graphs of f and g would both be solid lines, it would look as if the two waves were being reflected from 0 and L.



(d) Apply this to the two special cases where $u(x,0)$ is either $\sin\big([n\pi/L]x\big)$ or $\cos\big([n\pi/L]x\big)$; these special cases, which might have been noted independently, then imply the result in general, by using the Fourier series expansion.

CHAPTER 9

RIGID BODY MOTION

Euler was one of the first to analyze the motion of rigid bodies in any detail, and he wrote his equations in terms of a rotating, non-inertial, coordinate system. Aside from the use of such coordinate systems in the study of rigid body motion, the basic preparation will also serve as an introduction to the material in Chapter 10.

Rotating coordinate systems. For a one-parameter family of rotations $B(t)$ of $\mathbb{R}^3$, at time t we can consider the orthonormal basis $(\mathbf{u}_1(t), \mathbf{u}_2(t), \mathbf{u}_3(t))$ defined by $\mathbf{u}_i(t) = B(t)(\mathbf{e}_i)$ for the standard orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. These can be taken as the unit vectors of a new coordinate system at time t, so that in this way we obtain a "rotating coordinate system".

Now given a curve $\mathbf{r}$ in $\mathbb{R}^3$, with

$$\mathbf{r} = r_1 \cdot \mathbf{e}_1 + r_2 \cdot \mathbf{e}_2 + r_3 \cdot \mathbf{e}_3, \qquad \text{or simply } \mathbf{r} = (r_1, r_2, r_3),$$

we want to consider the components of this curve when written with respect to this rotating coordinate system, that is we want to write

$$\text{(a)} \qquad \begin{aligned} \mathbf{r} &= \rho_1 \cdot \mathbf{u}_1 + \rho_2 \cdot \mathbf{u}_2 + \rho_3 \cdot \mathbf{u}_3 \\ &= \rho_1 \cdot B(\mathbf{e}_1) + \rho_2 \cdot B(\mathbf{e}_2) + \rho_3 \cdot B(\mathbf{e}_3), \end{aligned}$$

where we are using abbreviated notation: $B(\mathbf{e}_i)$ stands for $t \mapsto B(t)(\mathbf{e}_i)$.

An observer rotating along with these coordinate systems will regard these rotating coordinates as simply being a standard set of coordinates, and for such an observer the ρ_i are the coordinates of $\mathbf{r}$ in this standard set of coordinates, and $\mathbf{r}$ will simply be described as the curve

$$\text{(b)} \qquad \boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3) \in \mathbb{R}^3.$$

To relate these two descriptions, we simply note that comparison of (a) and (b) gives the following equation in $\mathbb{R}^3$,

$$\mathbf{r} = B(\boldsymbol{\rho}), \qquad \text{again using abbreviated notation.}$$

Continuing to use abbreviated notation for convenience, so that, for example, $B(\boldsymbol{\rho}')$ stands for $t \mapsto B(t)(\boldsymbol{\rho}'(t))$, we now have

$$\begin{aligned} \mathbf{r}' &= B(\boldsymbol{\rho}') + B'(\boldsymbol{\rho}) \\ &= B(\boldsymbol{\rho}') + B'B^{-1}(\mathbf{r}), \end{aligned}$$

and if we introduce $\boldsymbol{\omega}$, giving the components of the skew-symmetric matrix $B'B^{-1}$, as on page 186, we can write

$$\text{(1)} \qquad \mathbf{r}' = B(\boldsymbol{\rho}') + \boldsymbol{\omega} \times \mathbf{r}.$$

To interpret the term $B(\boldsymbol{\rho}')$, we note that since

$$\boldsymbol{\rho}' = (\rho_1{}', \rho_2{}', \rho_3{}') = \rho_1{}' \cdot \mathbf{e}_1 + \rho_2{}' \cdot \mathbf{e}_2 + \rho_3{}' \cdot \mathbf{e}_3,$$

we have

$$B(\boldsymbol{\rho}') = \rho_1{}' \cdot \mathbf{u}_1 + \rho_2{}' \cdot \mathbf{u}_2 + \rho_3{}' \cdot \mathbf{u}_3.$$

This means that $B(\boldsymbol{\rho}')$ is just what the observer would compute for the derivative of the curve by taking the components of $\mathbf{r}$ in the rotating coordinates, and then simply differentiating these components, what one might call the "rotating observer's derivative". We will denote it by $\mathbf{r}$', so that we can write

$$\text{(2)} \qquad \mathbf{r}' = \mathbf{r}\text{'} + \boldsymbol{\omega} \times \mathbf{r}.$$

Thus, the derivative $\mathbf{r}'$ is the sum of the "rotating observer's derivative" $\mathbf{r}$' and a correction term $\boldsymbol{\omega} \times \mathbf{r}$. [Physicists often write something like $\frac{d\mathbf{r}}{dt} = \frac{\delta \mathbf{r}}{\delta t} + \boldsymbol{\omega} \times \mathbf{r}$, with $\frac{\delta}{\delta t}$ indicating the operation where only the components in the rotating coordinate system are being differentiated.]

The Euler equations. In Chapter 10 we will consider rotating coordinate systems in general, but in this chapter we are mainly concerned with a rigid body rotating about some fixed point, and we will usually take the rotating coordinate axes to be the principle axes of inertia of the body at all times, so that the $\boldsymbol{\omega}$ of equation (1) is the same as the angular velocity $\boldsymbol{\omega}$ of the rotating body. In this particular context the rotating observer's derivative $\mathbf{r}$' is often called the "body derivative" [and physicists often write something like $\left(\frac{d\mathbf{r}}{dt}\right)_{\text{space}} = \left(\frac{d\mathbf{r}}{dt}\right)_{\text{body}} + \boldsymbol{\omega} \times \mathbf{r}$].

In particular, we can now apply equation (2) to the angular momentum curve $\mathbf{L}$ of our rotating body to get

$$(\mathbf{E}_\tau) \qquad \boldsymbol{\tau} = \mathbf{L}' = \mathbf{L}\text{'} + \boldsymbol{\omega} \times \mathbf{L}.$$

When there are no external forces, so that $\boldsymbol{\tau} = 0$, we then have

$$(\mathbf{E}) \qquad \mathbf{L}\text{'} = \mathbf{L} \times \boldsymbol{\omega}.$$

Since the rotating coordinate axes are now the principle axes of inertia of the body at all times, the components L_1, L_2, L_3 of the vector $\mathbf{L}$ will just be

$\omega_1 I_1$, $\omega_2 I_2$, $\omega_3 I_3$, where the constants I_1, I_2, I_3 are the principal moments of inertia. When there are no external forces, by taking the components of our vector equation (**E**) in the rotating coordinate system we obtain the standard form of the *Euler equations*:

$$\text{(E)} \qquad \begin{aligned} I_1\omega_1' &= (I_2 - I_3)\omega_2\omega_3 \\ I_2\omega_2' &= (I_3 - I_1)\omega_3\omega_1 \\ I_3\omega_3' &= (I_1 - I_2)\omega_1\omega_2. \end{aligned}$$

For the case of external forces, equation ($\mathbf{E}_\tau$) yields corresponding equations

$$(\mathrm{E}_\tau) \qquad \begin{aligned} \tau_1 &= I_1\omega_1' + (I_3 - I_2)\omega_2\omega_3 \\ \tau_2 &= I_2\omega_2' + (I_1 - I_3)\omega_3\omega_1 \\ \tau_3 &= I_3\omega_3' + (I_2 - I_1)\omega_1\omega_2. \end{aligned}$$

As a particular consequence of the Euler equations (E), we can reprove the result on page 193 concerning rotation of a rigid body about an axis when there are no external forces, though we must assume (as proved on page 192), that the angular velocity is constant: This means that the ω_i are all constants, so that the right hand side of each equation of (E) is 0; consequently, if the I_i are all distinct, we have $0 = \omega_2\omega_3 = \omega_3\omega_1 = \omega_1\omega_2$, which means that if one of the ω_i is non-zero, the other two must be zero, and thus our rotation is around the principle axis corresponding to the non-zero ω_i.

We can obtain more sensitive information by using conservation of angular momentum $\mathbf{L}$, and in particular of its norm $L = |\mathbf{L}|$, together with conservation of kinetic energy T, writing L and T in terms of the components $L_i = \omega_i I_i$ of $\mathbf{L}$. From the equation for T_{rot} on page 194,

$$\text{(T)} \qquad 2T = \langle \mathbf{I}(\boldsymbol{\omega}), \boldsymbol{\omega} \rangle,$$

we have

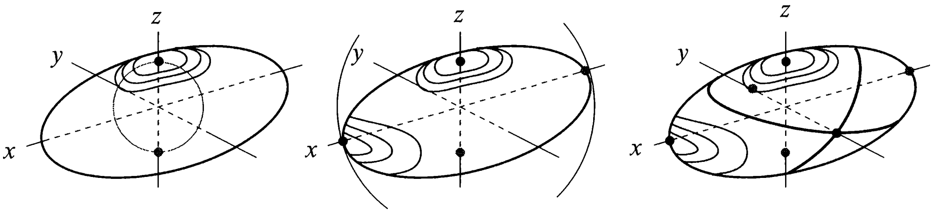
$$(1) \qquad 2T = \frac{{L_1}^2}{I_1} + \frac{{L_2}^2}{I_2} + \frac{{L_3}^2}{I_3}$$

$$(2) \qquad L^2 = {L_1}^2 + {L_2}^2 + {L_3}^2.$$

So $\mathbf{L}$ lies in the intersection of an ellipsoid and a sphere (in the rotating coordinate system). To be specific, let us say that $I_1 < I_2 < I_3$, so that the semiaxes of the ellipsoid are

$$\sqrt{2TI_1} < \sqrt{2TI_2} < \sqrt{2TI_3}.$$

Then (L_1, L_2, L_3) lies on the ellipsoid that appears in the figure below as the surface $x^2/I_1 + y^2/I_2 + z^2/I_3 = 2T$. Part (a) of the figure shows the smallest sphere, of radius $\sqrt{2TI_1}$, that intersects the ellipsoid. The intersection consists of the two points at the end of the smallest axis, and as the radius of the sphere is increased we obtain a family of curves, a few of which are shown. Similarly, (b) shows the largest sphere, of radius $\sqrt{2TI_3}$, intersecting the ellipsoid



(a) (b) (c)

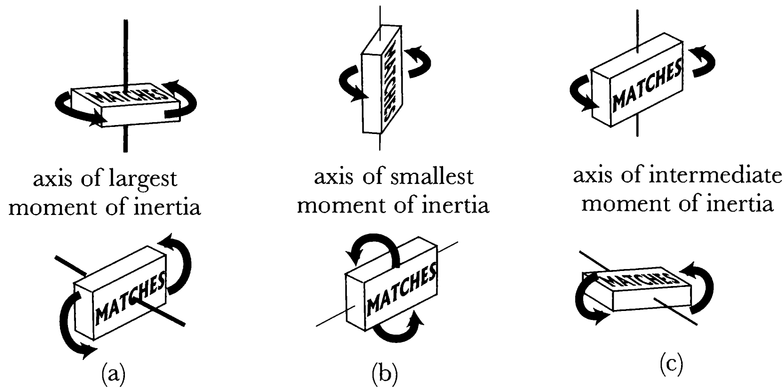
in the two points at the end of the largest axis, with some of the curves obtained as the radius of the sphere is *de*creased added in. The situation for the sphere of radius $\sqrt{2TI_2}$ is completely different, however. If we set $L^2 = 2TI_2$ in (2), we can solve (1) and (2) to get

$$L_3{}^2 = \frac{I_1 - I_2}{I_2 - I_3}\frac{I_1}{I_3}L_1{}^2 = \alpha^2 L_1{}^2, \quad \text{say.}$$

This means that the intersection of the ellipsoid and the sphere is the same as the intersection of the ellipsoid with the planes $z = \pm\alpha x$, and thus two ellipses (c), with the remainder of the surface of the ellipsoid accounted for by the intersections with the curves from the first two families. (In the symmetric case when two moments of inertia are equal, the situation is rather different: if $I_1 = I_2$, say, then the intersections are simply the circles parallel to the z-axis.)

As an application of all this information we note the following. If our body is rotating about an axis with moment of inertia I_1 or I_3, and is then nudged a bit, the path of $\mathbf{L}$ will be changed from its constant path to one of the nearby small paths; more precisely, since L might change, we should say that its path will change to a small path on a nearby ellipsoid. So the body, though no longer rotating about an axis, will still stay close to its original position, and in this sense rotations about these axes are stable. On the other hand, if it is rotating about the axis with moment of inertia I_2, then the slightest nudge will send it onto a path that rapidly moves away (except in the very special case that it moves onto one of the two ellipses). Problem 3 gives an analytic treatment.

Our result can be demonstrated effectively with a rectangular solid of three substantially unequal dimensions, like a filled match box of the sort used to hold camping or kitchen matches. Whatever the initial position of the box, it is easy to get it to spin (a) around the shortest axis as it falls, this axis being the one with the largest moment of inertia. Similarly, although a bit more care may



be required, the box can be made to spin (b) around the largest axis, with the smallest moment of inertia. On the other hand, attempts to get the box to spin around the other axis (c) almost always result in an unwieldy tumbling motion.

For those who are sports maniacs, rather than pyromaniacs, stability about the longest axis can be demonstrated more sportingly with a tennis racket, as you jauntily throw it into the air, with a well-executed spin, and deftly snatch it back as it falls with its axis of spin still pointing in the same direction.

Poinsot's geometric description. For the general rotating body with one fixed point, in the absence of external torque, we can use equations (1) and (2) on page 334 to express ω_1 and ω_3 in terms of ω_2, and substitute into the middle Euler equation to get an equation for $\omega_2{}'$, leading (Problem 1) to solutions in terms of elliptic integrals, but they are not especially revealing. A description due to Poinsot, who preferred geometric constructions to analytic equations, is sometimes invoked instead.

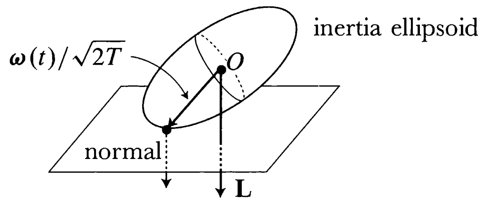
As a body rotates, the various vectors $\boldsymbol{\omega}(t)$ satisfy the equation (T) on page 194, or in the form on page 334,

$$2T = \langle \mathbf{I}(\boldsymbol{\omega}), \boldsymbol{\omega} \rangle,$$

which means that the vectors $\boldsymbol{\omega}(t)/\sqrt{2T}$ lie on the inertia ellipsoid. Since the equation of the inertia ellipsoid in body coordinates can be written as $1 = I_1x^2 + I_2y^2 + I_3z^2 = F(x, y, z)$, say, the normal at such a point on the inertia ellipsoid has the directions of the vector $(\partial F/\partial x, \partial F/\partial y, \partial F/\partial z)$ at $\boldsymbol{\omega}(t)/\sqrt{2T}$, and is thus a multiple of

$$(2I_1\omega_1(t),\ 2I_2\omega_2(t),\ 2I_3\omega_3(t)) = (2L_1(t),\ 2L_2(t),\ 2L_3(t)),$$

so the normal at $\boldsymbol{\omega}(t)/\sqrt{2T}$ is parallel to $\mathbf{L}$. This means that in our standard coordinate system, the normal of the inertia ellipsoid at $\boldsymbol{\omega}(t)/\sqrt{2T}$ has the constant direction of $\mathbf{L}$, so the tangent plane at any $\boldsymbol{\omega}(t)/\sqrt{2T}$ is perpendicular

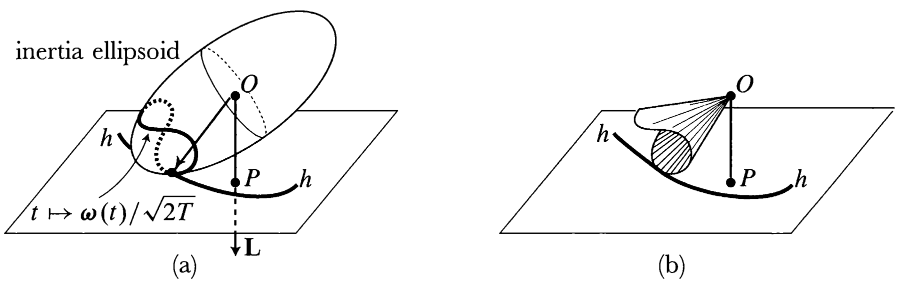


to $\mathbf{L}$. Moreover, remembering that $\mathbf{L} = \mathbf{I}(\boldsymbol{\omega})$ (as on page 188), we find that the distance from the fixed point to this plane is

$$\frac{\langle \boldsymbol{\omega}(t)/\sqrt{2T}, \mathbf{L}\rangle}{L} = \frac{\langle \boldsymbol{\omega}(t), \mathbf{I}(\boldsymbol{\omega}(t))\rangle}{L\sqrt{2T}} = \frac{2T}{L\sqrt{2T}} = \frac{\sqrt{2T}}{L}.$$

Since T and L are constants, these planes are all the *same* plane, known as the **invariable plane**.

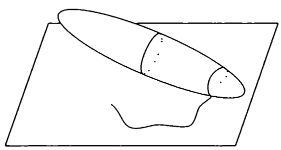
In other words, as our body moves, its inertia ellipsoid revolves around the fixed point O in such a way that at time t the point $\boldsymbol{\omega}(t)/\sqrt{2T}$ is tangent to



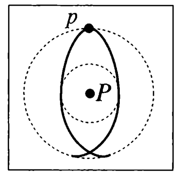
the invariable plane at some point $h(t)$, as in (a). In addition, condition (3) of the Proposition on page 222 holds for $B(t)$ [with the C of condition (3) being the cross-product with $\boldsymbol{\omega}(t)$], so the ellipsoid is *rolling* along the curve h; more conveniently, as in (b), we can simply visualize a cone rolling about the fixed point O.

Poinsot [1] called the curve $t \mapsto \boldsymbol{\omega}(t)/\sqrt{2T}$ the *polhode* of the moving body, from the Greek *πόλος* = pole, terminal and *οδος* = road, route. The figure on page 335 shows the ellipsoid $\langle \mathbf{I}(\boldsymbol{\omega}), \boldsymbol{\omega}\rangle = 2T$, a multiple of the inertia ellipsoid, with the paths of $\mathbf{L}$; these are not simply multiples of the polhodes, but since $L_i = \omega_i I_i$, the polhodes have the same general arrangement on the inertia ellipsoid (see also the remarks at the beginning of Addendum B).

Poinsot called the path h traced out on the invariable plane the *herpolhode*, so that one can intone: The polhode rolls without slipping on the herpolhode lying in the invariable plane. The name "herpolhode" was used to suggest a snakelike appearance, as Poinsot drew it in one of his diagrams, like the one in (a) of the figure below. Actually, as in (b), the herpolhode, when viewed



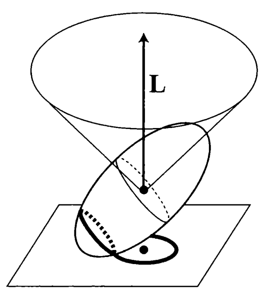
(a) herpolhode à la Poinsot



(b) an actual herpolhode as viewed from above

from above, is always concave with respect to the point P at which $\mathbf{L}$ intersects the invariable plane—the interesting, but computationally complicated, proof is outlined in Addendum B. After the inertia ellipsoid has rolled over the entire polhode, which is a closed curve, we are back to the initial conditions, except that the body has turned around the $\mathbf{L}$ axis, so the herpolhode is made up of repeating sections, just like the orbits studied in Chapter 4; in the above figure, the point p comes from a symmetry point of the polhode, and the two "loose ends" correspond to the symmetry point opposite it. If the angle through which the body turns happens to be commensurable with 2π, then the whole motion of the body will be periodic; otherwise, the herpolhode will fill up a dense region of an annulus around P.

Poinsot's result has been used to provide various elaborate descriptions of the rotation of a rigid body about a fixed point in the absence of external forces, but the only one likely to be mentioned nowadays is for the special symmetric case where we have $I_1 = I_2$, say. In this case the polhodes are circles, and the



herpolhode for any particular rolling motion is a circle in the invariant plane, and the axis of the inertia ellipsoid revolves at a constant rate around $\mathbf{L}$.

The free symmetric top, in body coordinates. Mechanics books usually refer to a body with $I_1 = I_2$ as a "symmetric top". By "free" we mean that there are no external forces, unlike the more familiar "heavy top" of everyday experience, to be discussed later, where the external force of gravity creates a torque around the point on which the top spins. Before giving an analytic treatment of the free symmetric top mirroring the geometric analysis given by Poinsot's method, we will first consider the analysis in body coordinates.

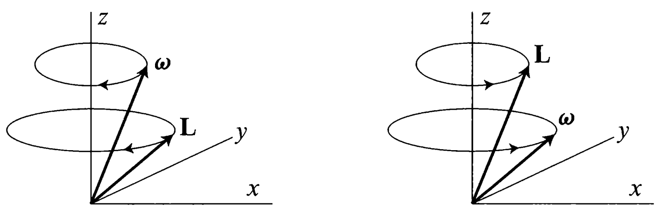
For the free symmetric top, we immediately get $\omega_3' = 0$ from the third Euler equation, so ω_3 is a constant, and the first two equations thus can be written in terms of the constant

$$\lambda = \frac{\omega_3(I_3 - I_1)}{I_1}$$

as

$$\left.\begin{aligned} \omega_1' &= -\lambda\omega_2 \\ \omega_2' &= \lambda\omega_1 \end{aligned}\right\} \quad \text{with the obvious}^1 \text{ solutions} \quad \left\{\begin{aligned} \omega_1 &= A\cos\lambda t \\ \omega_2 &= A\sin\lambda t \end{aligned}\right.$$

for a constant A (more generally, we should write $\lambda t + b$). This shows that $\boldsymbol{\omega}$ rotates around the symmetry axis with a constant angular frequency λ. Since $\mathbf{L}$ has the components $L_i = I_i\omega_i$, it follows that $\mathbf{L}$ rotates in a similar way.



$I_3 < I_1$ $\qquad\qquad$ $I_3 > I_1$

Euler applied these results to the earth, which is an oblate spheroid with

$$\frac{I_1 - I_3}{I_1} \approx -.003,$$

to predict the *Euler precession* of the "geometric North Pole", where the axis of rotation of the earth intersects its northern hemisphere, about the "celestial North Pole", which we may take as the axis through the center of the earth

[1] Disdaining the obvious, one can differentiate the first equation to get a second order equation for ω_1, which will be one for harmonic motion; or one can write $(\omega_1 + i\omega_2)' = i\lambda(\omega_1 + i\omega_2)$, with the solution $\omega_1 + \omega_2 = A\exp(i\lambda t)$.

pointing to the "fixed stars" (the axis from the center of the earth to the North Star will do).[1] The angular velocity should be given by λ on page 339,

$$\lambda = \omega_3 \frac{I_3 - I_1}{I_1} \approx \frac{\omega_3}{300} \approx \frac{|\boldsymbol{\omega}|}{300} \quad \text{since } \boldsymbol{\omega}_3 \text{ is practically the same as } |\boldsymbol{\omega}|.$$

If we use the day as our unit of time, then $|\boldsymbol{\omega}| = 2\pi$, so the period should be

$$\approx 2\pi/\lambda = 300 \text{ days}.$$

Euler, using the measurements available in 1755, predicted a period of 355 days, but searches by several astronomers for motions with a period close to this were unsuccessful. In 1891, Chandler, looking for motions with possibly quite different periods, reported a motion with a period of about 14 months, now known as the Chandler wobble. This was initially received with considerable scepticism, until the difference was explained as due to the non-rigidity of the earth, with, as usual, several other phenomena eventually adding even more complications to the whole picture.

The free symmetric top, in inertial coordinates. The body coordinates are the obvious choice for examining the Euler precession, but for observing everyday objects we want to consider the symmetric top in inertial coordinates (we can consider a top in space, or a top thrown into the air, so that the gravitational force merely causes the whole top to descend, or an arrangement like that of a gyroscope, cf. page 355 ff.). Although one can presumably derive the view in inertial coordinates from the view in the body coordinates, it's much easier to note directly that the equations

$$\mathbf{L} = I_1(\omega_1\mathbf{u}_1 + \omega_2\mathbf{u}_2) + I_3\omega_3\mathbf{u}_3$$
$$\boldsymbol{\omega} = \omega_1\mathbf{u}_1 + \omega_2\mathbf{u}_2 + \omega_3\mathbf{u}_3$$

lead to $\mathbf{L} = I_1\boldsymbol{\omega} + (I_3 - I_1)\omega_3\mathbf{u}_3 = I_1(\boldsymbol{\omega} + \lambda\mathbf{u}_3)$, and thus

$$\text{(a)} \qquad \boldsymbol{\omega} = \frac{\mathbf{L}}{I_1} - \lambda\mathbf{u}_3.$$

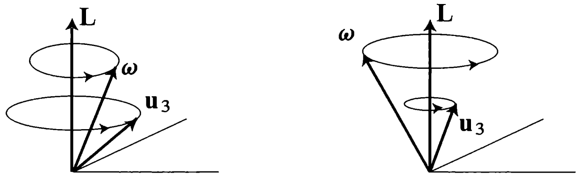
Remembering that $\mathbf{u}_3' = 0$, we have

$$\mathbf{u}_3{}' = \frac{\mathbf{L}}{I_1} \times \mathbf{u}_3$$

and thus $\mathbf{u}_3$ is rotating around the fixed vector $\mathbf{L}$ with frequency L/I_1. This is known as the *regular precession* of the top. Since (a) shows that $\boldsymbol{\omega}$, $\mathbf{L}$ and $\mathbf{u}_3$

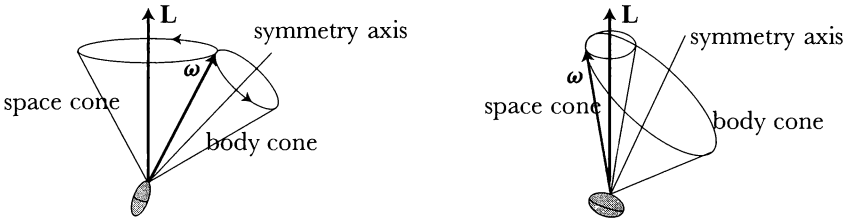
[1] This precession is totally distinct from the astronomical "precession of the equinoxes", with a period of 26,000 years, that we will mention later.

are coplanar, it follows that $\boldsymbol{\omega}$ also rotates around $\mathbf{L}$ with frequence L/I_1 (see Problem 2 for the reconciliation between this frequency and the frequency λ for the body coordinates).



$I_3 < I_1$ $\qquad$ $I_3 > I_1$

The view in the body coordinates and in the inertial coordinates are sometimes combined, for maximal confusion, in a figure showing the "space cone"

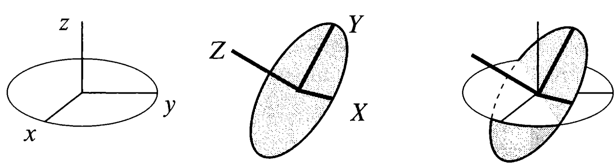


$I_3 < I_1$ $\qquad$ $I_3 > I_1$

around $\mathbf{L}$ swept out by $\boldsymbol{\omega}$, together with the "body cone" that $\boldsymbol{\omega}$ sweeps out around the symmetry axis in the body coordinates. Of course, in our inertial system the body cone will be moving around $\mathbf{L}$ also; in fact, since $\boldsymbol{\omega}$ is the instantaneous axis of rotation of the top, the body cone is rolling on the stationary space cone without slipping (invoking an appropriate interpretation of the Proposition on page 222). Note that for $I_3 > I_1$, the body cone is rolling on the space cone now situated inside it.

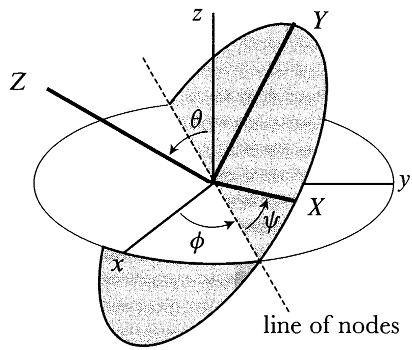
Euler angles. Our analysis so far has used the fact that the configuration space for a body with one fixed point is SO(3), without recourse to any specific coordinate systems. To go further, however, we will need to introduce such a coordinate system on SO(3), also due to Euler.

In the figure below, part (a) shows the standard x-y-z axes, part (b) the axes X-Y-Z that they are taken into by some rotation, and part (c) the two together.



(a) $\qquad$ (b) $\qquad$ (c)

The following enlarged version of (c) shows the intersection of the (x, y)-plane and the (X, Y)-plane, known as the *line of nodes*, a term borrowed from astronomy (cf. page 565). This diagram, of course, presumes that our rotation is not



simply a rotation around the z-axis. We can now specify our rotation by 3 coordinates: the counterclockwise angle ϕ of rotation about the z-axis that takes the positive x-axis to one ray of the line of nodes; the angle θ of rotation about the line of nodes that takes the positive z-axis to the positive Z-axis; and finally, the angle ψ of the rotation about the Z-axis that takes our ray on the line of nodes to the positive X-axis.[1] For these *Euler angle* coordinates to be well-defined, we must take $\theta \in (0, \pi)$, and $\phi, \psi \in (0, 2\pi)$.

Thus, a triple (ϕ, θ, ψ) determines a rotation B, namely the rotation that takes the standard x-y-z axes to the X-Y-Z axes having these values. If we let $\mathbf{Z}_\phi$, $\mathbf{X}_\theta$, $\mathbf{Z}_\psi$ be the rotations with the following matrices

$$\text{for } \mathbf{Z}_\phi \quad \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{for } \mathbf{X}_\theta \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \qquad \text{for } \mathbf{Z}_\psi \quad \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then we claim that the rotation B can be written as the product

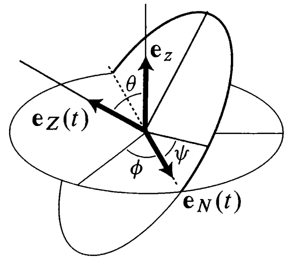
$$(*) \qquad B = \mathbf{Z}_\psi \mathbf{X}_\theta \mathbf{Z}_\phi.$$

In fact, after $\mathbf{Z}_\phi$, a rotation about the z-axis, the new first axis is the line of nodes, and $\mathbf{X}_\theta$ thus gives a rotation about this line, moving the z-axis to the new third axis Z, and $\mathbf{Z}_\psi$ then describes the rotation about this axis.

[1] Unfortunately, many variants appear in the literature; ϕ and ψ may be interchanged, angles may be measured in different directions, etc.

Functions $\phi(t)$, $\theta(t)$, $\psi(t)$ determine corresponding rotations $B(t)$, and the components of the skew-symmetric matrix $B'B^{-1}(t)$ then give us the vector $\boldsymbol{\omega}(t)$. We would like to determine these components directly in terms of $\phi(t)$, $\theta(t)$, $\psi(t)$. The easiest way to do this is to first express $\boldsymbol{\omega}$ in terms of the triad

$$\begin{aligned}\mathbf{e}_z &= \text{unit vector along the third axis } z,\\ \mathbf{e}_N(t) &= \text{unit vector along the line of nodes,}\\ \mathbf{e}_Z(t) &= \text{unit vector along the third body axis } Z,\end{aligned}$$



since, after all, these are the axes around which the angles (ϕ, θ, ψ) are measured. Note that

(i) If $\theta = \psi = 0$, with only ϕ changing, then $B(t)$ simply involves a rotation about the z-axis, so $\boldsymbol{\omega}(t)$ will simply be $\phi'(t) \cdot \mathbf{e}_z$.

(ii) If ϕ is constant and $\psi = 0$, with only θ changing, then $B(t)$ simply involves a rotation about the line of nodes, so $\boldsymbol{\omega}(t)$ will simply be $\theta'(t) \cdot \mathbf{e}_N(t)$.

(iii) If ϕ and θ are constant, with only ψ varying, then the matrix $B(t)$ simply involves a rotation about the Z-axis, so $\boldsymbol{\omega}(t) = \psi'(t) \cdot \mathbf{e}_Z(t)$.

On the basis of observations like these, the following result is usually considered to be obvious (cf. the note at the end of the proof).

1. LEMMA. The decomposition of $\boldsymbol{\omega}$ is

$$\boldsymbol{\omega}(t) = \phi'(t) \cdot \mathbf{e}_z + \theta'(t) \cdot \mathbf{e}_N(t) + \psi'(t) \cdot \mathbf{e}_Z(t).$$

PROOF. We have

$$B(t) = \mathbf{Z}_{\psi(t)}\mathbf{X}_{\theta(t)}\mathbf{Z}_{\phi(t)} = \boldsymbol{\Psi}(t)\boldsymbol{\Theta}(t)\boldsymbol{\Phi}(t), \quad \text{say,}$$

and the components of $\boldsymbol{\omega}$ are, up to sign, the off-diagonal components of the matrix of

$$\begin{aligned}B'B^{-1} &= (\boldsymbol{\Psi\Theta\Phi})'(\boldsymbol{\Psi\Theta\Phi})^{-1}\\ &= \big(\boldsymbol{\Psi\Theta\Phi}' + \boldsymbol{\Psi\Theta}'\boldsymbol{\Phi} + \boldsymbol{\Psi}'\boldsymbol{\Theta\Phi}\big)\boldsymbol{\Phi}^{-1}\boldsymbol{\Theta}^{-1}\boldsymbol{\Psi}^{-1}\\ &= (\boldsymbol{\Psi\Theta})(\boldsymbol{\Phi}'\boldsymbol{\Phi}^{-1})(\boldsymbol{\Psi\Theta})^{-1} + \boldsymbol{\Psi}(\boldsymbol{\Theta}'\boldsymbol{\Theta}^{-1})\boldsymbol{\Psi}^{-1} + \boldsymbol{\Psi}'\boldsymbol{\Psi}^{-1}.\end{aligned}$$

The third term involves rotations with ϕ and θ constant, since $\boldsymbol{\Psi}$ is our final rotation. Observation (iii) shows that this provides the proper ψ' term.

The second term involves rotations with ϕ constant, since the final rotation $\boldsymbol{\Psi}$ doesn't involve ϕ. Moreover, it also involves rotations with $\psi = 0$ because of the conjugation by $\boldsymbol{\Psi}$. This provides the proper θ' term, by observation (**ii**).

Similarly, the first term involves rotations with $\theta = \psi = 0$, because of the conjugation by $\boldsymbol{\Psi\Theta}$, providing the proper ϕ' term, by observation (**i**). ❖

Note: The idea behind this proof is encapsulated in Problem 4, showing that in certain cases "angular momentum vectors can be added", when this statement is properly formulated.

Now we just have to express $\mathbf{e}_N$ and $\mathbf{e}_z$ in terms of $\mathbf{e}_X$ and $\mathbf{e}_Y$, the unit vectors along the body axes X and Y. From our diagram for the Euler angles we easily see that

$$\mathbf{e}_N = (\cos\psi)\mathbf{e}_X - (\sin\psi)\mathbf{e}_Y. \tag{e$_1$}$$

The Z component of $\mathbf{e}_z$ is $\cos\theta$, while its component in the (X, Y)-plane, with length $\sin\theta$, can be decomposed with respect to the X and Y axes to get

$$\begin{aligned} \mathbf{e}_z &= (\cos\theta)\mathbf{e}_Z + \sin\theta\big[(\sin\psi)\mathbf{e}_X + (\cos\psi)\mathbf{e}_Y\big] \\ &= (\cos\theta)\mathbf{e}_Z + (\sin\theta\sin\psi)\mathbf{e}_X + (\sin\theta\cos\psi)\mathbf{e}_Y \end{aligned} \tag{e$_2$}$$

or equivalently

$$\begin{aligned} \langle \mathbf{e}_X, \mathbf{e}_z\rangle &= \sin\theta\sin\psi \\ \langle \mathbf{e}_Y, \mathbf{e}_z\rangle &= \sin\theta\cos\psi \\ \langle \mathbf{e}_Z, \mathbf{e}_z\rangle &= \cos\theta. \end{aligned} \tag{e$_2{}'$}$$

Substituting (e$_1$) and (e$_2$) into $\boldsymbol{\omega} = \phi'\mathbf{e}_z + \theta'\mathbf{e}_N + \psi'\mathbf{e}_Z$, we get finally

$$\begin{aligned} \boldsymbol{\omega} = {} & \big(\phi'\sin\theta\sin\psi + \theta'\cos\psi\big)\,\mathbf{e}_X \\ & + \big(\phi'\sin\theta\cos\psi - \theta'\sin\psi\big)\,\mathbf{e}_Y + \big(\phi'\cos\theta + \psi'\big)\,\mathbf{e}_Z, \end{aligned}$$

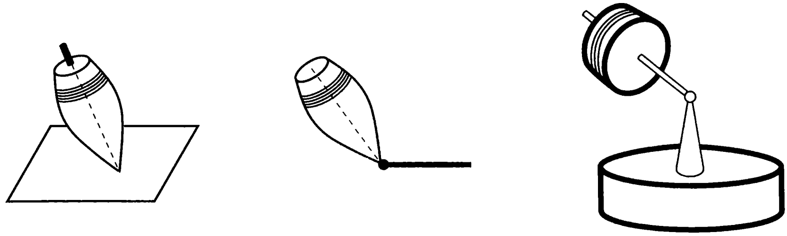
so that

$$\begin{aligned} \omega_1 &= \phi'\sin\theta\sin\psi + \theta'\cos\psi \\ \omega_2 &= \phi'\sin\theta\cos\psi - \theta'\sin\psi \\ \omega_3 &= \phi'\cos\theta + \psi'. \end{aligned} \tag{ω}$$

These equations are sometimes called "Euler's geometric equations", in contrast to "Euler's dynamic equations" (E) on page 334.

In the next section we will apply these equations to analyze the motion of a rigid body with one point fixed, but since they are purely geometrical, depending only on the fixed directions for the x-y-z axes, we can also apply them when the (x, y, z)-plane is moving parallel to itself through some point around which we wish to consider our rotation, a situation that we will sometimes encounter.

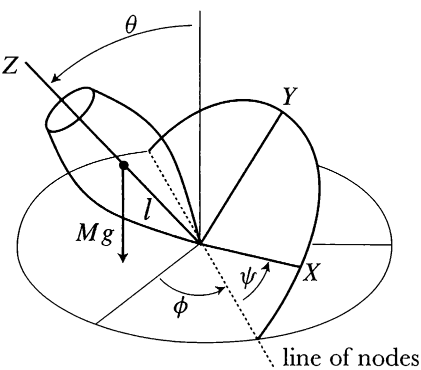
The heavy symmetrical top. We are now ready to consider a symmetrical top, with one point fixed, acted upon by gravity. We will draw the top as a symmetric body, although we really only require that its inertia ellipsoid is symmetrical, and that its center of gravity lies on the rotation axis. The picture we usually



(a) Toy Top (b) Swivel Top (c) Lab Top

have in mind is the (toy) top on a non-slippery surface (a), perhaps with grooves around which to wind a string, which can be pulled off rapidly to impart a rapid rotation. Our theoretical top might best be realized as a body having its fixed point attached to a swiveling joint (b). Textbooks often picture the top as in (c), where a heavy wheel spins on an axis attached to a joint supported on a heavy base. The position of our top is then described by a rotation about the fixed point (to determine it uniquely, we should imagine that some point not on the axis has been marked).

If the mass of the top is M, the total effect of gravity is equivalent to a single force of magnitude Mg acting on the center of mass, which we're assuming is on the rotation axis, at some distance l from the fixed point.[1]



[1] Of course, there is also an upward force on the fixed point of the top from the surface on which it spins, and horizontal frictional forces of that surface, all of which keep the point fixed; here (and previously in this Chapter) we are implicitly using the analysis of Chapter 6, with our configuration space being SO(3).

Substituting the first two formulas of (ω) into the formula for the kinetic energy,

$$T = \frac{1}{2} I_1 ({\omega_1}^2 + {\omega_2}^2) + \frac{1}{2} I_3 {\omega_3}^2,$$

gives

$$T = \frac{I_1}{2}(\theta'^2 + \phi'^2 \sin^2\theta) + \frac{I_3}{2}{\omega_3}^2,$$

while the potential energy is simply $V = Mgl\cos\theta$, so conservation of energy gives

$$E = \frac{I_1}{2}(\theta'^2 + \phi'^2 \sin^2\theta) + \frac{I_3}{2}{\omega_3}^2 + Mgl\cos\theta$$

for a constant E. We have not yet applied the third formula of (ω), for the following reason. Note that the torque of the downward force of gravity is along the line of nodes, while the Z-axis is perpendicular to this line, which means that $\langle \mathbf{L}, \mathbf{e}_Z \rangle = I_3\omega_3$ is a constant (hence the top is always spinning with constant angular velocity around its symmetry axis), and we can write our equation in terms of the new constant $\tilde{E} = E - \frac{1}{2} I_3 {\omega_3}^2$ as

$$\text{(A)} \qquad \tilde{E} = \frac{I_1}{2}(\theta'^2 + \phi'^2 \sin^2\theta) + Mgl\cos\theta$$

[compare to equation (b′) on page 290].

When we do apply the third equation of (ω), we find that

$$\text{(B)} \qquad I_3\omega_3 = I_3(\phi'\cos\theta + \psi') = I_1 a$$

for a suitable constant a.

Note, moreover, that since the z-axis is also perpendicular to the line of nodes, the component $\langle \mathbf{L}, \mathbf{e}_z \rangle = \langle I_1\omega_1\mathbf{e}_X + I_1\omega_2\mathbf{e}_Y + I_3\omega_3\mathbf{e}_Z,\ \mathbf{e}_z \rangle$ is also constant. Using the values of ω_i from (ω), together with ($\mathrm{e}_2{}'$), we find that

$$\text{(C)} \qquad (I_1 \sin^2\theta + I_3\cos^2\theta)\phi' + I_3\psi'\cos\theta = I_1 b$$

for a suitable constant b.

Equations (B) and (C) are often not treated in this direct way, but are instead deduced from considerations about the Lagrangian (see Chapter 12), which provide the equations "automatically", without even worrying about what they signify. In any case, however, (A), (B), (C) now provide everything we need.

We use (B) to write

$$I_3\psi' = I_1 a - I_3\phi' \cos\theta, \tag{1}$$

and substitute into (C) to obtain

$$\phi' = \frac{b - a\cos\theta}{\sin^2\theta} \tag{2}$$

[compare to equation (a′) on page 290]. Thus ϕ is known once θ is determined; moreover, by substituting (2) back into (1) we obtain

$$\psi' = \frac{I_1 a}{I_3} - \cos\theta \cdot \frac{b - a\cos\theta}{\sin^2\theta}, \tag{2'}$$

which shows that ψ is also known once θ is determined.

To obtain an equation for θ, we substitute (2) into (A), ending up with

$$\sin^2\theta \cdot \theta'^2 = \sin^2\theta(\alpha - \beta\cos\theta) - (b - a\cos\theta)^2 \tag{3}$$

where the constants α and β are defined as

$$\alpha = \frac{2\tilde{E}}{I_1}, \qquad \beta = \frac{2Mgl}{I_1}.$$

Finally, if we set $u = \cos\theta$, then (3) reduces to

$$u'^2 = f(u) = (1 - u^2)(\alpha - \beta u) - (b - au)^2 \tag{$*$}$$

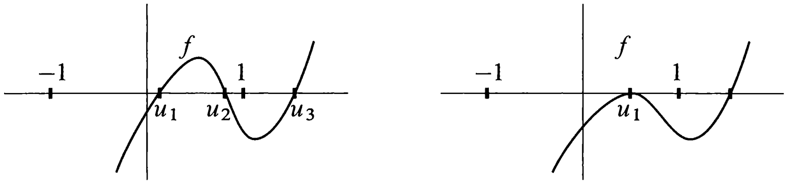
for a cubic polynomial $f(u)$ [and parts of the succeeding discussion will call to mind the discussion of the spherical pendulum].

At this point we could write $(*)$ as

$$t = \int \frac{du}{\sqrt{(1 - u^2)(\alpha - \beta u) - (b - au)^2}},$$

so that u could be written in terms of Jacobian elliptic functions, leading to a formula for θ, and thence by (2) and (2′) to formulas for ϕ and θ, theoretically solving the problem. But we can get a much better qualitative picture by considering the general properties of the cubic polynomial $f(u)$. Naturally, only the behavior of f for $u \in [-1, 1]$ is significant, since $u = \cos\theta$. Note that the coefficient of u^3 in $f(u)$ is β, which is positive.

We assume for now that 1 and -1 are not roots of f, implying that $a \neq \pm b$, leaving the contrary case to be considered later in the game. Since $f(\pm 1) = -(b \mp a)^2$, it follows that f is definitely negative at 1 and -1. But the f that arises for a top can't be negative on all of $[-1,1]$, since for any value of u occuring in $(*)$ we must have $f(u) \geq 0$. Generally we will have $f(u) > 0$ at



(a) (b)

some $u \in (-1,1)$, and then the graph of f looks something like (a), with f having two zeros $u_1 < u_2 \in [-1,1]$ (as well as a zero $u_3 > 1$ of no interest to us), so that u lies in the interval $[u_1, u_2]$. There is also the special situation where there is just one (double) zero, $u_1 = u_2$ in $[-1,1]$, as in (b), so that we always have $u = u_1$.

As in the case of the spherical pendulum on page 291, differentiation of equation $(*)$ leads to a second order equation

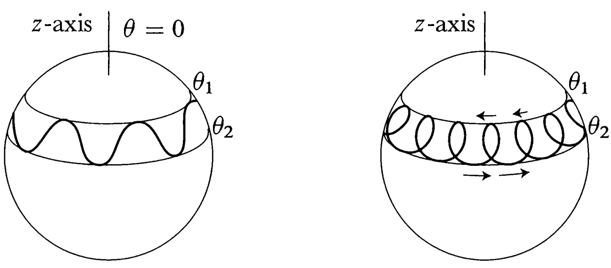
$$(*') \qquad 2u'' = f' \circ u \quad \text{or, in Leibnizian notation,} \quad 2\frac{d^2u}{dt^2} = f'(u),$$

of the form encountered in Chapter 4 (page 128), so if $\cos\theta$ is not constant, then it varies periodically between u_1 and u_2, the two places where $u' = 0$, and naturally θ varies similarly between $\theta_1 = \arccos(u_1)$ and $\theta_2 = \arccos(u_2)$, the two places where $\theta' = 0$, with the top rising and falling in a periodic fashion.

As for ϕ, since we have

$$\phi' = \frac{b - au}{1 - u^2} \qquad \text{with } 1 - u^2 > 0,$$

we immediately see that if $a/b \notin [u_1, u_2]$, then ϕ' is never 0, so ϕ varies monotonically, and the axis of the top traces out a curve on the sphere like that shown in (a). On the other hand, if $a/b \in (u_1, u_2)$, then the sign of ϕ' as θ varies



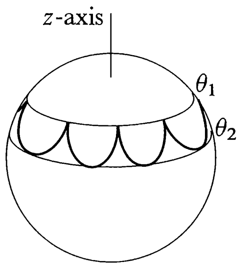
(a) (b)

from θ_1 to θ_2 is the opposite of the sign of ϕ' as θ varies from θ_2 back down to θ_1, so we obtain a curve with loops, as in (b).

Note that, although we drew the graph of f with $u_1 > 0$ and $u_2 > 0$, either or both could be negative, meaning that one or both of θ_1 and θ_2 can be greater than $\pi/2$, with the axis of the top pointing below the horizontal. Naturally, we would need a Swivel Top rather than a Toy Top to realize this situation.

Generally speaking, the motion of the top is determined by three periodic motions: (1) its fixed rate of rotation around its axis; (2) the change of θ, its *nutation*; and (3) the change of ϕ, its *precession*. The three periods will usually be distinct, and the top returns to its initial position only in the exceptional case where they are all commensurable.

The cuspidal case; fast tops. Our general discussion has left out some important special cases, among them the case of a double root of f, which we will defer to the next section. For now, we note that although we've seen that the axis of the top traverses a looping curve for $a/b \in (u_1, u_2)$, we still have to examine the possibility that a/b is an endpoint u_i of $[u_1, u_2]$, so that at some time we have $\theta = \theta_i$ and $\theta' = \phi' = 0$. At such a time, in equation (A) we have $\tilde{E} = Mgl\cos\theta_i$, while the other terms are always positive. Consequently, for $\tilde{E}$ to remain constant, $Mgl\cos\theta$ must begin to decrease, so θ must begin to increase, which means that we must have had $\theta = \theta_1$, rather than θ_2. Problem 5



shows that the curve traced out by the axis of the top is perpendicular to the circle $\theta = \theta_1$ at the cuspidal intersection points.

This seemingly exceptional situation occurs whenever we have a top held spinning with its axis stationary at inclination θ_1, and then simply release it at time $t = 0$, without imparting any other motion to it, so that $\theta'(0) = \phi'(0) = 0$. We could start the Lab Top spinning while holding the end of its axle, or, if our Toy Top was designed to spin independently around the protruding axle (some toy tops are made this way, designed so that the top can be set spinning by pushing down on the axle), we could set it spinning and then carefully place it on the surface at a given angle.

As our analysis shows, the top always starts falling (with ϕ increasing all the while), until it gets to the angle θ_2, at which point it starts rising until it gets back to the angle θ_1. But we can say quite a bit more.

Since $\phi'(0) = 0$, equation (2) for ϕ' gives

$$b = a\cos\theta_1 = au_1, \quad \text{say,}$$

and since we have $\tilde{E}(0) = Mgl\cos\theta_1$, our definition of the constants α and β in equation (3) gives

$$\alpha = \beta\cos\theta_1 = \beta u_1.$$

We then have

$$\begin{aligned} f(u) &= (1-u^2)(\alpha - \beta u) - (b - au)^2 \\ &= (1-u^2)\beta(u_1 - u) - a(u_1 - u)^2, \end{aligned}$$

which can be written as

$$(*_{\text{cusp}}) \qquad f(u) = (u_1 - u)\big[\beta(1-u^2) - a^2(u_1 - u)\big],$$

so that, in addition to the zero u_1 of f, the other zeros are the solutions of the quadratic equation

$$0 = \beta u^2 - a^2 u + (a^2 u_1 - \beta).$$

If we set

$$\lambda = \frac{a^2}{2\beta},$$

the solutions can be written as $\lambda \pm \sqrt{\lambda^2 - 2\lambda u_1 + 1}$. Since one of these must be the irrelevant solution that is greater than 1, the solution of interest is the one with a negative square root, so

$$\cos\theta_2 = \lambda - \sqrt{\lambda^2 - 2\lambda u_1 + 1} = \lambda - \lambda\left(1 - \frac{2\cos\theta_1}{\lambda} + \frac{1}{\lambda^2}\right)^{\frac{1}{2}}.$$

To evaluate λ explicitly, note that from the definition $\beta = 2Mgl/I_1$ and the value of a given by equation (B) we have

$$\lambda = \frac{(I_3\omega_3)^2}{4I_1 Mgl}.$$

Now consider what happens for large λ, which simply means that ω_3 is large, i.e., the top is spinning rapidly. We can use the binomial theorem to write

$$\cos\theta_2 = \lambda - \lambda\left(1 + \tfrac{1}{2}\left[-\frac{2\cos\theta_1}{\lambda} + \frac{1}{\lambda^2}\right] - \tfrac{1}{8}\left[\frac{4\cos^2\theta_1}{\lambda^2} + \cdots\right] + \cdots\right)$$

so

$$\cos\theta_2 - \cos\theta_1 \approx -\frac{\sin^2\theta_1}{2\lambda}.$$

In particular, $\theta_2 - \theta_1$ will be small, so that we also have

$$\cos\theta_2 - \cos\theta_1 \approx (\theta_2 - \theta_1)(-\sin\theta_1),$$

and therefore

$$\theta_2 - \theta_1 \approx \frac{\sin\theta_1}{2\lambda} = \frac{2I_1 Mgl}{(I_3\omega_3)^2}\sin\theta_1.$$

Thus, the faster the top is spinning, the *smaller the nutation.*

Moreover, since the nutation is small, we can approximate the equation $u'^2 = f(u)$ by replacing the term $(1-u^2)$ in $(*_{\text{cusp}})$ by $(1-u_1{}^2) = \sin^2\theta_1$. If we write the resulting equation $u'^2 = f(u)$ in terms of the new variable $x = u_1 - u$, with $x'^2 = u'^2 = f(u)$, then our equation becomes

$$x'^2 = x(\beta\sin^2\theta_1 - a^2 x),$$

and the solution of this differential equation for the initial condition $x(0) = 0$, equivalent to $u(0) = u_1$, is

$$x(t) = \frac{\beta\sin^2\theta_1}{2a^2}(1 - \cos at).$$

Since this is a constant times $(1 - \cos at)$, the frequency of nutation is approximately

$$a = \frac{I_3\omega_3}{I_1}.$$

Thus, the faster the top is spinning, the *greater the frequency of nutation.*

Finally, since

$$\phi' = \frac{a(u_1 - u)}{\sin^2\theta} \approx \frac{ax}{\sin^2\theta_1},$$

our formula for $x(t)$ gives

$$\phi' \approx \frac{\beta}{2a}(1 - \cos at),$$

which has an average value of

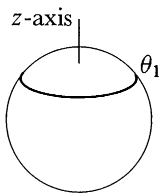
$$\frac{\beta}{2a} = \frac{Mgl}{I_3\omega_3}.$$

Thus, the faster the top is spinning the *smaller the rate of precession.*

To put all this another way, as a top slows down because of friction it will start precessing faster, and with a nutation of larger magnitude, though of smaller frequency.

If we start an actual top rapidly spinning at a fixed angle, and then release it without pushing it in any direction, the resulting very small nutation may be unnoticeable, especially as it is very likely to be damped out by friction because of its high frequency. So it can appear that the top simply starts precessing on its own, which is actually impossible, since the only force on it is gravity, which is perpendicular to the direction of the precession.[1] The point raised at the end of Chapter 5 would seem to be related to this fact, a matter that will be discussed later on, in the section on gyroscopes.

Precessing tops. There are certain situations where a top will truly exhibit precession without nutation. In the special case where there is only one zero of f in $[-1, 1]$, so that $u_1 = u_2$, we have $\theta = \theta_1$, a constant, and then ϕ' is also a constant ϕ'_1 (and ψ' is similarly a constant ψ'_1), so that we just get a curve circling at constant angular velocity around the parallel at θ_1.



Experimentally, we can obtain this situation by starting the Toy Top or Lab Top spinning, as in the previous section, and then giving the top a horizontal shove to initiate precession. To show mathematically that there are initial conditions that will lead to this case, we note that since u_1 is a double root, we have $f'(u_1) = 0$, and equation $(*')$ gives $u''(u_1) = 0$, implying that $\theta''(\theta_1) = 0$. If we write equation (3) on page 347 as

$$\theta'^2 = (\alpha - \beta \cos\theta) - \frac{(b - a\cos\theta)^2}{\sin^2\theta},$$

differentiate it, divide by θ', and then use the fact that $\theta''(\theta_1) = 0$, together with equation (2) on page 347, we find that

$$\frac{\beta}{2} = a\phi'_1 - (\phi'_1)^2 \cos\theta_1.$$

Using the definitions of a and β, we then get as the condition for precession

$$Mgl = \phi'_1\big(I_3\psi'_1 - (I_1 - I_3)\phi'_1 \cos\theta_1\big),$$

[1] Chapter 11 gives other examples of situations where the unobserved effects of friction serve to produce paradoxical results.

which is a quadratic equation for the constant ϕ_1':

$$\text{(P)} \qquad (I_1 - I_3)\cos\theta_1(\phi_1')^2 - (I_3\psi_1')\phi_1' + Mgl = 0.$$

If we want to specify initial conditions at $t = 0$,

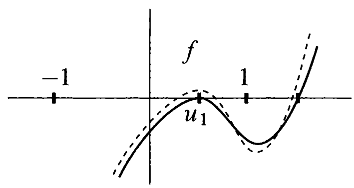
$$(\theta, \phi, \psi; \theta', \phi', \psi')(0) = (\theta_1, \phi(t_1), \psi(t_1); 0, \phi_1', \psi_1'),$$

then $\phi(t_1)$ and $\psi(t_1)$ can be chosen arbitrarily, and after choosing a value for ψ_1', we can find ϕ_1' satisfying (P), as long as we choose ψ_1' making the discriminant of (P) non-negative:

$$(I_3\psi_1')^2 \geq 4Mgl(I_1 - I_3)\cos\theta_1.$$

When the discriminant satisfies this condition with the strict inequality holding, there will be two different solutions for ϕ_1', the "slow" and "fast" precessions. Our equation (P) can never be satisfied for the initial condition $\phi_1' = 0$: a true procession is possible only when we start the top with an initial precessional velocity.

Notice that since the cubic f for a top always has at least one positive root, in the case of a precessing top any slight modification of f will have to have two



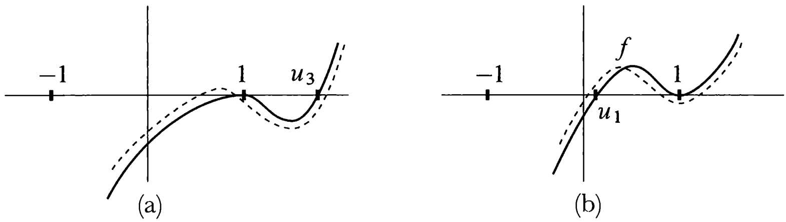
nearby roots, so the preccessing top is stable, in the sense that slight perturbations, like friction or a slight breeze, won't cause the top to suddenly start nutating through a large angle, in contrast to a situation that can arise in the next case that we will examine.

Sleeping tops. The case where 1 or -1 is a root of $f(u)$ arises for $b = \pm a$; we'll concentrate on the case $b = a$, since $b = -a$ is basically the same situation with the top upside down. The simplest example occurs for a "sleeping" top, spinning with its axis vertical. Since we always have $\theta = 0$ we also have $\theta' = 0$, and then equation (A) gives $\tilde{E} = Mgl\cos\theta$, and it follows that we also have $\alpha = \beta$.

As with the case of the precessing top, we want to investigate whether the motion of a sleeping top is stable. Although the Euler angles aren't well-defined for a vertical axis, we can still consider the cubic occuring in $(*)$,

$$\begin{aligned} f(u) &= (1-u^2)(\alpha - \beta u) - (b - au)^2 \\ &= (1-u^2)\beta(1-u) - a^2(1-u)^2 \\ &= (1-u)^2[\beta(1+u) - a^2], \end{aligned}$$

which now has 1 as a double root, like either (a) or (b) below. Under a slight perturbation, the cubic f will again change only slightly, but in case (a) there will now be two roots both close to 1, so that the axis remains close to vertical,



while in case (b) one of the roots will be close to the root u_1, and the top will suddenly start nutating through an angle θ close to $\arccos u_1$.

We will definitely have case (a) if $f''(1) < 0$ and case (b) if $f''(1) > 0$. We find that $f''(1) < 0$ is equivalent to $2\beta < a^2$, or

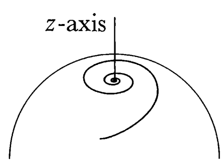
$$\omega_3{}^2 > \frac{4MglI_1}{I_3{}^2},$$

so we have stability if the top is spinning rapidly enough, with ω_3 satisfying this inequality, while the motion will definitely be unstable if ω_3 satisfies the reverse inequality. For the boundary case $\omega_3{}^2 = 4MglI_1/I_3{}^2$, we find that 1 is a triple root of f, which again implies stability.

The rising top. A cubic of type (b) for the sleeping top can also arise for a non-sleeping top whose initial conditions happen to give $a = b$ and $\alpha = \beta$, and then the angle θ can vary between $\arccos u_1$ and 0. However, θ will only approach 0 asymptotically as $t \to \infty$, since the solution of the differential equation $u'^2 = f(u)$ with the initial condition $u(t_0) = 1$ is simply the constant function $u(t) = 1$. On the other hand, since

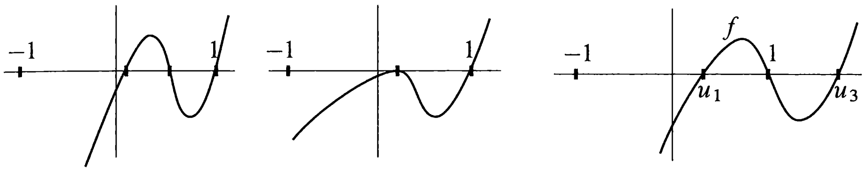
$$\lim_{\theta\to 0} \phi' = \lim_{\theta\to 0} \frac{a(1-\cos\theta)}{\sin^2\theta} = \frac{a}{2},$$

the angle ϕ grows infinitely large as $t \to \infty$, so the axis winds around the top pole of the sphere infinitely often. Of course, aside from the fact that the top

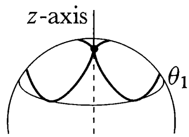


won't continue spinning forever, once the axis of the top gets close enough to the vertical, friction will make it indistinguishable from a sleeping top, so this motion will essentially just look like the time reversal of a slowly spinning sleeping top that dips down because of instability.

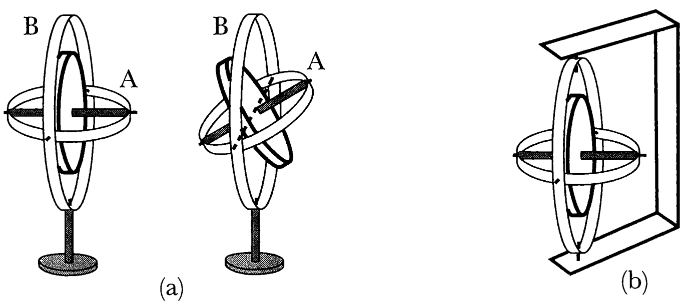
The polar cuspidal top. The final theoretical motion occurs when we consider the remaining case, where $a = b$, $\alpha \neq \beta$, with $f'(1) = -2(\alpha - \beta) \neq 0$, so that 1 is not a double root. The three different possibilities are shown below.



The first two graphs are the same as those appearing on page 348, except that the "extraneous" root x_3 is now 1, which doesn't change anything in the analysis. The third graph is also like the first graph on that page, but since our interval is now $[u_1, 1]$ with $a/b = 1 \in [u_1, 1]$, this behaves like the cuspidal case, except that the top circle now degenerates into the top pole.

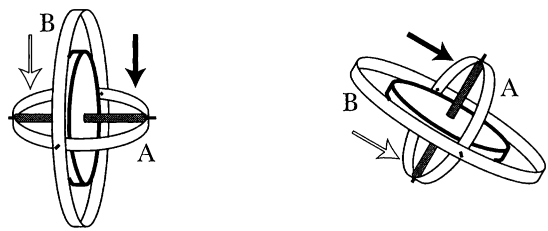


Gyroscopes. The basic features of souped-up laboratory or commercial gyroscopes are all illustrated by the toy gyroscope (a), where the axis of the wheel rotates in an inner "gimbal" A, which can rotate in an outer gimbal B, as shown in the second picture, where the dashed line is the imaginary axis connecting the two points where the gimbals intersect. The gimbal B, in turn, can rotate about the vertical axis, providing 3 degrees of freedom. For the sake of stability, an arrangement like (b) might be preferred.



This is customarily called a *Cardan suspension*, after Jerome Cardan (Girolamo Cardano, 1501–1576), famous for his book *Ars Magna* (1545), which presented the solutions to the cubic and quartic equations. However, it seems[1] that Cardan merely described, but never claimed to have invented, gimbals, which appear to have been around at least since 140 B.C. in China. They can be used in such mundane applications as cup holders in boats and moving vehicles, and were used in the 19th century to hold lamps in Roma caravans. The first gyroscope seems to have been constructed by the mathematician and astronomer Johann Bohnenberger in 1817; it was given its name, from the Greek *γύρος* = turn and *σκοπὸς* = view, by Foucault, whom we will meet again in Chapter 10.

The fixed point of a gyroscope is its center of mass, so the Mgl term that we used in the analysis of the top is irrelevant here; the net effect of the gimbals and stand is to have no force on the center of mass, so that gravity is irrelevant, and we can just as well imagine the gyroscope at any angle, providing a picture closer to our previous picture of the top. A situation equivalent to the heavy top occurs when we exert a force on the inner gimbal. A steady downward force

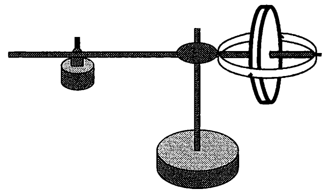


on gimbal A, indicated by the black arrow in the figure, produces a constant torque on the axle, which takes the place of the Mgl term in our analysis of the top, so the exact same analysis will apply (the downward force indicated by the white arrow would cause a torque corresponding to a $-Mgl$ term, so we we would have to turn our second picture upside down).

Since we usually start the gyroscope spinning, and then exert the force, this corresponds to the cuspidal case. When the gyroscope is attached to its stand, the precession involves the outer gimbal rotating, while the small nutation of the axis involves the inner gimbal oscillating. For a heavy, rapidly spinning gyroscope, this nutation is quickly damped, and might even not be perceived, so it can seem that a push in one direction simply causes motion in a perpendicular direction. The precession is sometimes demonstrated with apparatus like that in the figure below, where the balancing counterweight can be moved toward

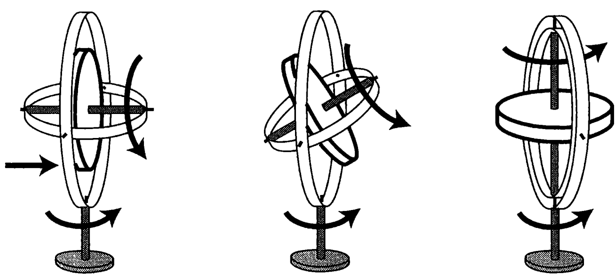
[1] See `findarticles.com/p/articles/mi_m1310/is_1988_Oct/ai_6955856` for free access to an article in the UNESCO Courier.

the tall central axis to produce precession in one direction, or further away to produce precession in the other direction.



The original gyroscope on its stand provides a good model for the experiment described at the end of Chapter 5. The back half of the inner gimbal A represents the person's arms, which can be rotated, principally by the shoulder muscles, within the back half of gimbal B, representing the rest of the person, with the stand representing the rotating bench. Although the surprising precession is the main point of such a demonstration, in most cases the nutation is also noticeable, though it is probably usually perceived simply as the person struggling to rotate the axis.

As we would suspect, when we push on the outer gimbal, the axis of the spinning wheel moves away from the horizontal, and the direction in which it moves, which we can figure out by the considerations of Chapter 5, can be

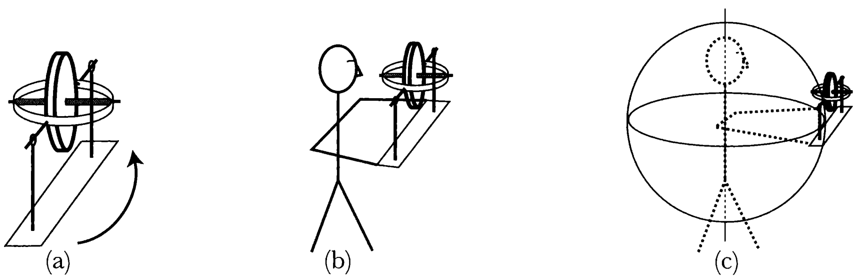


(a) (b) (c)

stated as a simple rule: Given the direction of rotation of the outer gimbal and the direction of rotation of the wheel (a), the axis of the spinning wheel moves so that the directions of spin tend to align (b), with the two directions being identical when the the axis is vertical (c). This situation does not correspond directly to the heavy top because a constant push on the outer gimbal does not produce a constant torque. What one observes when pushing on the outer gimbal is that the resistance tends to diminish, and the axis suddenly flips up to the vertical direction. Continuing to push on the outer gimbal then has no effect, but if we push the gimbal in the other direction, the gyroscope suddenly flips over and starts spinning in the opposite direction!

Of course, neither of these observed phenomena can occur exactly as just described. When the axis first reaches the vertical direction, it can't suddenly stop there. It must oscillate about the vertical in some way, though this oscillation is quickly damped by friction. Once the axis has stabilized in the vertical direction, continuing to push on the outer gimbal should theoretically have no effect at all. In practice, the axis will always wobble a bit, with the rotation of the outer gimbal simply helping to bring it back to vertical more quickly. However, if we now push the outer gimbal in the other direction, then as soon as the axis deviates the slightest bit from the vertical, the axis will have to move in the other direction, in order for the directions of spin to align, so the whole process is reversed, giving the impression that the gyroscope suddenly flips over.

The device in (a) of the figure below is substantially the same gyroscope, with the torque produced by rotating the base. Alternatively, as in (b) we can hold



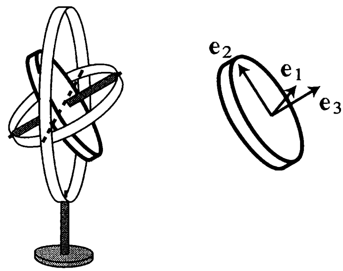
our arms outstretched in front of us, with one hand holding each end of the base, and then rotate ourselves, again causing the axle of the wheel to flip up to the vertical direction. This is actually a somewhat misleading description of the axle's behavior, since the direction in which it flips has nothing to do with "up" or "down"; what we should say is that it causes the axle to line up with the axis of the rotation, namely, our body.

Finally, part (c) of the figure shows our device positioned on the equator of the earth in such a way that the earth's rotation on its axis takes the place of the rotating person. The base is now in the plane of the equator, so that from the point of view of an actual person standing at the equator, the base and the initial position of the axis are vertical (the base is presumably anchored to a wall of some sort, which takes the place of our outstretched arms). The axle of the wheel will move in a horizontal plane, lining up with the axis of the earth's rotation, and thus point to the geometric north pole, or at any rate oscillate about this direction, so our device ought to serve as a "gyrocompass". We just have to see whether this will be practical, and to take into account what happens at other latitudes.

The gyrocompass. We will begin by considering our original gyroscope. The mathematical analysis will go right back to the Euler equations, in fact right back to our original equation

$$(\mathbf{E}_\tau) \qquad \tau = \mathbf{L}' + \boldsymbol{\omega} \times \mathbf{L}.$$

Our rotating coordinate system $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ will have $\mathbf{e}_1$ pointing along the axis



connecting the points where the gimbals intersect, which always lies in a horizontal plane, $\mathbf{e}_2$ a perpendicular vector lying along the wheel, and $\mathbf{e}_3$ pointing in the direction of the axis of our wheel. As usual, I_3 will be the moment of inertia around this axis, while $I_1 = I_2$ will be the moments of inertia about a diameter of the wheel. The only force on the wheel is the torsion caused by rotating the outer gimbal.

If a is the constant angular velocity of the outer gimbal, and θ is the angle from the vertical to $\mathbf{e}_2$, then $\boldsymbol{\omega}$, the angular velocity of our coordinate system in the body coordinates, is

$$\begin{aligned}\boldsymbol{\omega} &= \left\{(a\cos\theta)\mathbf{e}_3 + (-a\sin\theta)\mathbf{e}_1\right\} + \theta'\mathbf{e}_2 \\ &= (-a\sin\theta)\mathbf{e}_1 + \theta'\mathbf{e}_2 + (a\cos\theta)\mathbf{e}_3.\end{aligned}$$

If A is the angular velocity of the wheel about its axis—the large angular initial velocity we give it, plus the $a\cos\theta$ term, plus any small changes that the combined forces may produce—then the angular velocity of the wheel is

$$(-a\sin\theta)\mathbf{e}_1 + \theta'\mathbf{e}_2 + A\mathbf{e}_3,$$

and its angular momentum in body coordinates is

$$\mathbf{L} = (-I_1 a\sin\theta)\mathbf{e}_1 + I_1\theta'\mathbf{e}_2 + I_3 A\mathbf{e}_3.$$

We can write, for some X and Y,

$$\mathbf{L}' = X\,\mathbf{e}_1 + (I_1\theta'')\mathbf{e}_2 + (I_3A')\mathbf{e}_3$$
$$\boldsymbol{\omega} \times \mathbf{L} = Y\,\mathbf{e}_1 + (I_3 a A \sin\theta - I_1 a^2 \sin\theta\cos\theta)\,\mathbf{e}_2 + 0,$$

and $\boldsymbol{\tau}$ is a multiple of $\mathbf{e}_1$, so by looking at the coefficient of $\mathbf{e}_2$ in $(\mathbf{E}_{\boldsymbol{\tau}})$ we obtain

$$(*) \qquad I_1\theta'' + I_3 a A \sin\theta - I_1 a^2 \sin\theta\cos\theta = 0.$$

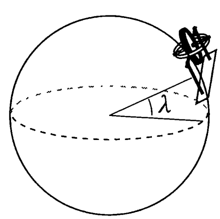
Looking at the coefficient of $\mathbf{e}_3$, we simply get $A' = 0$, so that A is constant. These conclusions can presumably also be derived from the equations (E_τ) on page 334.

This analysis can be immediately applied to a gyrocompass at the equator, where $a = 1$ revolution per 24 hours, or 1 revolution per 86,400 seconds, while the angular velocity A of the wheel in a typical gyroscope is about 20,000 revolutions per second, so $a \ll A \implies a^2 \ll aA$. If we therefore ignore the term with the factor a^2 in our equation, we simply get

$$\theta'' + k^2 \sin\theta = 0 \qquad\qquad k = \sqrt{\frac{I_3 a A}{I_1}}$$

with solutions a multiple of $\theta(t) = \sin kt$, oscillating about $\theta = 0$. For a thin rotating disk, I_1 is about $I_3/2$ (Problem 5-6), so that k is approximately $\sqrt{2aA}$, giving a period at the equator of approximately $2\pi/\sqrt{2aA} \approx 9.24$ seconds for $A = 20{,}000$. One can determine true north by bisecting the angle of swing, or wait for friction to cause the motion to cease.

When the gyrocompass is at latitude λ, we have



$$\boldsymbol{\omega} = (-a\sin\theta\cos\lambda)\,\mathbf{e}_1 + (\theta' + a\sin\lambda)\,\mathbf{e}_2 + (a\cos\theta\cos\lambda)\,\mathbf{e}_3$$
$$\mathbf{L} = (-I_1 a\sin\theta\cos\lambda)\,\mathbf{e}_1 + I_1(\theta' + a\sin\lambda)\,\mathbf{e}_2 + (I_3A)\,\mathbf{e}_3.$$

and we end up with the equation

$$I_1\theta'' + I_3 a A \sin\theta\cos\lambda - I_1 a^2 \sin\theta\cos\theta\cos^2\lambda = 0,$$

which we approximate by the same equation as before, but with

$$k = \sqrt{\frac{I_3 a A \cos\lambda}{I_1}},$$

so that the period $2\pi/\sqrt{2aA\cos\lambda}$ becomes longer, and the gyrocompass less useful, as the latitude increases; it is ≈ 11 seconds for latitude 45°, and infinite at the north pole.

Of course, all sorts of ingenious and complicated engineering mechanisms, and modifications, are required to make a practical gyrocompass, especially one that will work not only on land but on a ship, and respond quickly enough to allow automatic corrections to keep the ship on course.

Precession of the equinoxes. Finally, we should add a few words about the astronomical "precession of the equinoxes", a very slow precession of the axis of the earth not related to the Euler precession mentioned on page 339.

If the earth were a perfect sphere, homogeneous, or even radially symmetric, the only gravitational effect of the sun on the earth would be a force directed toward the sun; in particular, the sun could not produce any torque on the earth. Because the earth isn't exactly spherical, the sun does produce a torque on the earth, but this torque is not directly related to the earth's spinning on its axis, except for that fact that this spinning is what produces the bulging near the equator in the first place. In fact, this torque is due to the "tidal forces" of the sun's gravitational field (Problem 4-20), since a uniform gravitational force would produce no torque on the earth no matter what its shape.

This small torque makes the spinning earth act like a gyroscope with a slow precession having a period of about 26,000 years, with constellations appearing in the night sky in different seasons during this cycle (the additional orbital motion of the earth around the sun is so small compared to the distance to the constellations that it is usually completely ignored). At the "spring equinox", the time when the sun is directly overhead at noon, the sun will appear in different constellations, hence the name "precession of the equinoxes", known even to ancient astronomers/astrologers. In a mere 600 years or so the sun, now in Pisces, will be in Aquarius, as anticipated in the musical *Hair.*

Calculations of the precession become exceedingly involved (Newton presented a geometric one, with fudging, cf. Cohen-Whitman [1; pg. 265]), and are seldom even mentioned in text books. A description of one the first serious attempts, by d'Alembert, can be found in Hand and Finch [1; pp. 317 ff.]; an exercise in Goldstein [1; Chap. 5] also tackles the problem.

ADDENDUM 9A

THE EULER EQUATIONS FOR ROTATING PRINCIPAL VECTORS

THE ROLLING DISC

When $I_1 = I_2$, it is occasionally useful to choose perpendicular unit principal vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ that are not fixed in the body, but that are rotating in the body. For simplicity, we will actually consider the equations arising for an arbitrary rotating orthonormal basis $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ at the fixed point O, although in practice the cases of interest involve only two equal principal moments of inertia, with $\mathbf{u}_3$ being fixed.

The rotating orthonormal basis has its own angular velocity $\boldsymbol{\xi}$, and we write $\boldsymbol{\omega}$ and $\boldsymbol{\xi}$ in terms of this rotating coordinate system as

$$\begin{aligned}\boldsymbol{\omega} &= \omega_1 \cdot \mathbf{u}_1 + \omega_2 \cdot \mathbf{u}_2 + \omega_3 \cdot \mathbf{u}_3 \\ \boldsymbol{\xi} &= \xi_1 \cdot \mathbf{u}_1 + \xi_2 \cdot \mathbf{u}_2 + \xi_3 \cdot \mathbf{u}_3.\end{aligned}$$

Similarly, we write the angular momentum vector $\mathbf{L}$ of the body as

$$\mathbf{L} = \omega_1 I_1 \cdot \mathbf{u}_1 + \omega_2 I_2 \cdot \mathbf{u}_2 + \omega_3 I_3 \cdot \mathbf{u}_3.$$

Our rotating differentiation ’ now refers to the rotating coordinates $\mathbf{u}_i$ and the equation

$$\boldsymbol{\tau} = \mathbf{L}' = \mathbf{L}\text{’} + \boldsymbol{\xi} \times \mathbf{L}$$

then gives us

$$\text{(T)} \qquad \begin{aligned}\tau_1 &= I_1\omega_1{}' - I_2\omega_2\xi_3 + I_3\omega_3\xi_2 \\ \tau_2 &= I_2\omega_2{}' - I_3\omega_3\xi_1 + I_1\omega_1\xi_3 \\ \tau_3 &= I_3\omega_3{}' - I_1\omega_1\xi_2 + I_2\omega_2\xi_1.\end{aligned}$$

We will also consider the more general case where the center of mass $\mathbf{x}$ of our rigid body is moving, and the ω_i and ξ_i refer to the appropriate rotations about the center of mass, and we write the velocity $\mathbf{v} = \mathbf{x}$’, computed in the $\mathbf{u}_i$ coordinates, as

$$\mathbf{v} = v_1 \cdot \mathbf{u}_1 + v_2 \cdot \mathbf{u}_2 + v_3 \cdot \mathbf{u}_3.$$

If $\mathbf{F}$ is the total force on the rigid body, of mass m, then

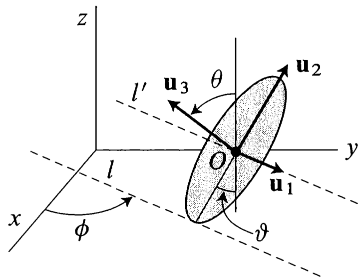
$$\mathbf{F} = m\mathbf{v}' = m(\mathbf{v}\text{’} + \boldsymbol{\omega} \times \mathbf{v}),$$

which gives us

(F)
$$\begin{aligned} F_1 &= m(v_1{}' - v_2\xi_3 + v_3\xi_2) \\ F_2 &= m(v_2{}' - v_3\xi_1 + v_1\xi_3) \\ F_3 &= m(v_3{}' - v_1\xi_2 + v_2\xi_1). \end{aligned}$$

We will apply these results to find a set of equations for the general case of a disc of radius a and mass m rolling on a plane;[1] a very different approach will be used in Addendum 12A.

In the figure below, the line l is the intersection of the plane of the disc with the (x, y)-plane, and l' is the line in the plane of the disc parallel to l through



the center O. We choose $\mathbf{u}_1$ to lie along l', with $\mathbf{u}_2$ lying along the line from O to the contact point. We let ϕ be the angle through which the contact point has rotated from the x-axis, as before, while θ is now the angle that $\mathbf{u}_3$ makes with the vertical. If one imagines the center O moved over to the origin, then our ϕ is just the Euler angle ϕ from page 342 and our θ is the Euler angle θ (and l' is the line of nodes). It will actually be more convenient for us to write things in terms of the angle $\vartheta = \pi/2 - \theta$, with the upright disc corresponding to $\vartheta = 0$.

The vectors $\mathbf{v}$ and $\boldsymbol{\omega}$ are related by the rolling condition (compare page 240)

$$\mathbf{v} + \boldsymbol{\omega} \times (-a\mathbf{u}_2) = 0,$$

which gives us

(a)
$$v_1 = -a\omega_3, \qquad v_2 = 0, \qquad v_3 = a\omega_1,$$

so the ω_i determine the v_i.

To find the components ξ_i of $\boldsymbol{\xi}$ with respect to the $\mathbf{u}_i$, we note that the coefficient of $\mathbf{u}_1$ is just $\theta' = -\vartheta'$, while there is an angular velocity of ϕ' about the vertical line through O, and decomposing this into its components with respect to $\mathbf{u}_2$ and $\mathbf{u}_3$ we have

(b)
$$\xi_1 = -\vartheta', \qquad \xi_2 = \phi' \cos\vartheta, \qquad \xi_3 = \phi' \sin\vartheta.$$

[1] From Synge and Griffith [1].

Since the angular velocities of the disk and the triple $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ differ only in the $\mathbf{u}_3$ component, we also have

$$\text{(c)} \qquad \omega_1 = -\vartheta', \qquad \omega_2 = \phi' \cos\vartheta.$$

Because of equations (a) – (c), we so far really have only three unknown functions, ϕ, ϑ, and ω_3 = the rate at which the disc is rolling.

The total force $\mathbf{F}$ on the disc is the downward gravitational force of magnitude gm and the reaction force $\mathbf{R}$ of the plane on the disc at the contact point, which we write in terms of three more unknowns,

$$\mathbf{R} = R_1 \cdot \mathbf{u}_1 + R_2 \cdot \mathbf{u}_2 + R_3 \cdot \mathbf{u}_3,$$

so that

$$\mathbf{F} = \mathbf{R} - mg(\cos\vartheta \cdot \mathbf{u}_2 + \sin\vartheta \cdot \mathbf{u}_3).$$

Using (a) – (c) in the equations (F) to express everything in terms of our six unknowns, we have

$$\text{(F}'\text{)} \qquad \begin{aligned} R_1 &= -ma(\omega_3{}' + \vartheta'\phi'\cos\vartheta) \\ R_2 - mg\cos\vartheta &= -ma(\vartheta'^2 + \phi'\sin\vartheta \cdot \omega_3) \\ R_3 - mg\sin\vartheta &= -ma(\vartheta'' - \phi'\cos\vartheta \cdot \omega_3). \end{aligned}$$

Similarly, equations (T) give (noting that the downwards gravitational force produces no torsion around O)

$$\text{(T}'\text{)} \qquad \begin{aligned} aR_3 &= I_1\vartheta'' + I_1\phi'^2\cos\vartheta\sin\vartheta - I_3\phi'\cos\vartheta \cdot \omega_3 \\ 0 &= I_1(\phi'\cos\vartheta)' + I_3\vartheta'\omega_3 - I_1\phi'\vartheta'\sin\vartheta \\ aR_1 &= I_3\omega_3{}'. \end{aligned}$$

Equations (F′) and (T′) are now 6 equations for 6 unknowns ϑ, ϕ, ω_3, R_1, R_2, R_3, and in terms of these unknowns we can find the ω_i by (c), and then the v_i by (a). Although we would obviously have to resort to numerical solutions in general, we can examine some special cases, and root out some additional information.

One obvious solution is

$$\begin{aligned} \vartheta &= 0 & R_1 &= 0 \\ \phi &= \text{constant} & R_2 &= mg \\ \omega_3 &= \text{constant} & R_3 &= 0, \end{aligned}$$

which is just the disc rolling vertically along a straight line. We can also look for a solution with the disc rolling along a circle, inclined at a fixed angle,

$$\vartheta = \text{constant}$$
$$\phi' = \text{constant}$$
$$\omega_3 = \text{constant}.$$

Equations (F′) then become

$$R_1 = 0$$
$$R_2 = m(-a\phi' \sin\vartheta \cdot \omega_3 + g\cos\vartheta)$$
$$R_3 = m(a\phi' \cos\vartheta \cdot \omega_3 + g\sin\vartheta),$$

and when we substitute into the first equation of (T′) we find the necessary condition

$$(I_3 + ma^2)\phi' \cos\vartheta \cdot \omega_3 + mga \sin\vartheta = I_1 \phi'^2 \sin\vartheta \cos\vartheta$$

(the angle at which the disc is inclined is related to the centripetal force that must be exerted in order for the disc to move in a circle).

We can also use the equations to investigate the stability of the straight line motion. When we roll a coin, or a hoop or thin tire like a bicycle tire along a surface, it stays close to straight line motion when it is spinning rapidly, but as it slows down it suddenly wobbles and falls over. In the straight line motion, the quantities

$$\vartheta, \quad \vartheta', \quad \vartheta'', \quad \phi', \quad \phi'', \quad \omega_3', \quad R_1, \quad R_2 - mg, \quad R_3$$

are all 0, so they will all be small for a small deviation from straight line motion and equations (F′) and (T′) give, up to first order,

(1) $R_1 = -ma\omega_3'$ (4) $aR_3 = I_1\vartheta'' - I_3\phi' \cdot \omega_3$

(2) $R_2 - mg = 0$ (5) $0 = I_1\phi'' + I_3\vartheta'\omega_3$

(3) $R_3 - mg\vartheta = -ma\vartheta'' + ma\phi' \cdot \omega_3$ (6) $aR_1 = I_3\omega_3'$.

Equations (1) and (6) give $\omega_3 = \text{constant}$, and (5) then gives

$$I_1\phi' + I_3\vartheta\omega_3 = \text{constant}.$$

Substituting this into the equation obtained by eliminating R_3 from (3) and (4) then gives us

$$\boldsymbol{A}\vartheta'' + \boldsymbol{B}\vartheta = \text{constant}$$

$$\boldsymbol{A} = I_1(I_1 + ma^2), \quad \boldsymbol{B} = I_3(I_3 + ma^2)\omega_3^2 - I_1 mga.$$

This gives small oscillations for $\boldsymbol{B} > 0$, but unbounded solutions for $\boldsymbol{B} \leq 0$, so for stability we need $\boldsymbol{B} > 0$, or

$$\omega_3^2 > \frac{I_1 mga}{I_3(I_3 + ma^2)}.$$

ADDENDUM 9B

SECRETS OF THE HERPOLHODE

Since the center O of the inertia ellipsoid is at a constant distance from the invariable plane as the ellipsoid rolls on it, the polhode can be described as the set of points on the ellipsoid whose tangent planes are at a fixed distance from the center. This definition can be made for any ellipsoid, even though not every ellipsoid is an inertia ellipsoid (Problem 5-10). If the equation of the ellipsoid is

$$ax^2 + by^2 + cz^2 = 1,$$

a computation shows that the distance d from the tangent plane at (x, y, z) to the origin is $d = 1/\sqrt{a^2x^2 + b^2y^2 + c^2z^2}$, so these general polhodes satisfy

$$\begin{aligned} ax^2 + by^2 + cz^2 &= 1 \\ a^2x^2 + b^2y^2 + c^2z^2 &= \frac{1}{d^2} = D, \quad \text{say,} \end{aligned}$$

for various constants D.

For the case of an inertia ellipsoid, we will switch to x_1, x_2, x_3 for the components of a point in the body coordinates, so that it has the equation

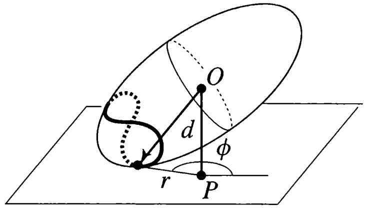
$$I_1{x_1}^2 + I_2{x_2}^2 + I_3{x_3}^2 = 1,$$

and we will let $x_1(t), x_2(t), x_3(t)$ be the body coordinates of the point of the polhode corresponding to some distance $d = 1/\sqrt{D}$. Then we have

$$(1) \qquad \begin{aligned} {x_1}^2 + {x_2}^2 + {x_3}^2 &= r^2 + \frac{1}{D} \\ I_1{x_1}^2 + I_2{x_2}^2 + I_3{x_3}^2 &= 1 \\ {I_1}^2{x_1}^2 + {I_2}^2{x_2}^2 + {I_3}^2{x_3}^2 &= D, \end{aligned}$$

where r is the distance from the point with body coordinates x_1, x_2, x_3 to the point P in the invariable plane directly below the fixed point O, which will be used as the origin for polar coordinates (r, ϕ) in that plane. Recall also (page 337) that for the inertia ellipsoid, the distance d is given by

$$(2) \qquad d = \frac{\sqrt{2T}}{L}.$$



Setting

$$\Delta = (I_1 - I_2)(I_2 - I_3)(I_3 - I_1)$$

and solving the three equations in (1) for the x_i in terms of the I_i, r and D gives

$$(3)\quad \begin{aligned} x_1{}^2 &= \frac{I_2 I_3 (I_3 - I_2)}{\Delta}(r^2 - a_1) \\ x_2{}^2 &= \frac{I_2 I_1 (I_1 - I_3)}{\Delta}(r^2 - a_2) \\ x_3{}^2 &= \frac{I_1 I_2 (I_2 - I_1)}{\Delta}(r^2 - a_3) \end{aligned} \qquad \text{for} \qquad \begin{cases} a_1 = -\dfrac{(I_2 - D)(I_3 - D)}{I_2 I_3 D} \\ a_2 = -\dfrac{(I_3 - D)(I_1 - D)}{I_3 I_1 D} \\ a_3 = -\dfrac{(I_1 - D)(I_2 - D)}{I_1 I_2 D}. \end{cases}$$

We will assume that $I_1 > I_2 > I_3$. If $1/d^2 = D = I_1$, so that $d = 1/\sqrt{I_1}$, the smallest semi-axis of the inertia ellipsoid, then the polhode and herpolhode are just points, and similarly if $D = I_3$. The special case $D = I_2$ will be disposed of at the end, and we will assume that D is between I_2 and I_3, the case where D is between I_1 and I_3 being similar, with various signs changed. In the current case, $\Delta < 0$ and $a_1, a_2 > 0$ while $a_3 < 0$.

Since the points on the polhode are of the form $\boldsymbol{\omega}/\sqrt{2T}$, their coordinates satisfy

$$(4)\qquad \omega_i = \sqrt{2T} x_i,$$

and the Euler equations yield

$$(5)\qquad \begin{aligned} I_1 x_1{}' + \sqrt{2T}(I_3 - I_2)x_2 x_3 &= 0 \\ I_2 x_2{}' + \sqrt{2T}(I_1 - I_3)x_3 x_1 &= 0 \\ I_3 x_3{}' + \sqrt{2T}(I_2 - I_1)x_1 x_2 &= 0. \end{aligned}$$

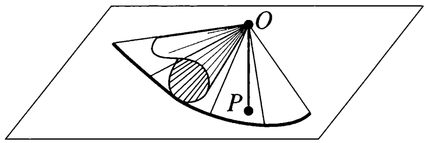
The first equation of (1) gives $rr' = xx' + yy' + zz'$, which together with (5) gives

$$r\frac{dr}{dt} = \sqrt{2T} x_1 x_2 x_3 \left(\frac{I_2 - I_3}{I_1} + \frac{I_3 - I_1}{I_2} + \frac{I_1 - I_2}{I_3} \right) = -\frac{\Delta \sqrt{2T} x_1 x_2 x_3}{I_1 I_2 I_3},$$

and then using (3),

$$(\mathrm{A})\qquad r\frac{dr}{dt} = \sqrt{2T}\sqrt{-(r^2 - a_1)(r^2 - a_2)(r^2 - a_3)}.$$

Finding a formula for $d\phi/dt$ will be quite a bit more interesting. In the figure below, we have drawn the polhode cone shown in (b) of the figure on



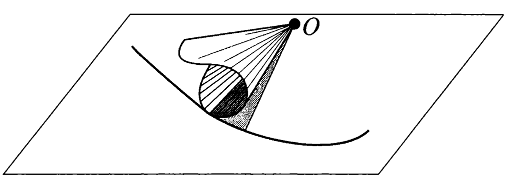
page 337, together with the cone from the same origin O to the herpolhode. As the polhode cone rotates about the point O, it moves, generator by generator, on the herpolhode cone, providing a mapping from the polhode cone to the herpolhode cone, and there is no stretching during this process, so this mapping must be an *isometry*.

One can show this formally by considering the parameterizations

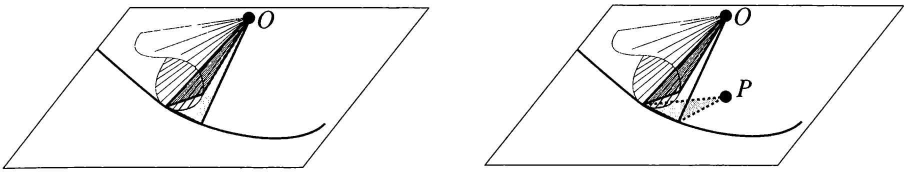
$$\begin{aligned} (s,t) &\mapsto s\cdot \mathbf{p}(t) \quad \text{of the polhode cone}\\ (s,t) &\mapsto s\cdot \mathbf{h}(t) \quad \text{of the herpolhode cone,} \end{aligned}$$

where $\mathbf{p}$ is the vector from O to a point on the polhode, and $\mathbf{h}$ the vector from O to the corresponding point on the herpolhode. Since each generator of the polhode goes to a generator of the same length on the herpolhode, $\partial/\partial s$ has the same length at corresponding points. Since rolling implies that equal lengths are marked off on the polhode and the herpolhode at all times, $\partial/\partial t$ has the same length at corresponding points. And since the mapping at each time is, up to first order, a rotation about a generator, the inner products of $\partial/\partial s$ and $\partial/\partial t$ have the same values at corresponding points.

This means that the region between two generators of the polhode cone has the same area as the region of the herpolhode cone between the corresponding



generators, and so the same is certainly true for the rate of change of these areas, keeping one generator fixed and varying the other. But for the rate of change, we can approximate these curved regions with the triangular regions bounded by the generators (a), so that the rate of change of the triangular regions must be the same. And for the triangular regions we can say that the rate of change



(a) (b)

of the projections of the areas on the invariable plane must also be the same. In the case of the herpolhode cone, this projection (b) is simply the triangular region between two radii from P to the herpolhode, and the rate of change is

$$\tfrac{1}{2}r^2\frac{d\phi}{dt}$$

since $\frac{1}{2}r^2\,d\phi$ is the integrand for area in polar coordinates. So we just have to determine the rate of change of the projection of the corresponding triangular region of the polhode cone, the triangle having $\mathbf{p}(t)$ and $\mathbf{p}(t+h)$ as its sides.

We do this in two steps, first finding the answer for the projection on the various coordinate planes in the body coordinates. If $\mathbf{p}(t) = (x_1(t), x_2(t), x_3(t))$, then for the projection in the (x_2, x_3) plane we are looking at the area of the triangle bounded by the lines from the origin to

$$(x_2(t), x_3(t)) \quad \text{and} \quad (x_2(t+h), x_3(t+h)),$$

which is just half the determinant

$$x_2(t)x_3(t+h) - x_2(t+h)x_3(t),$$

and the rate of change is thus $\frac{1}{2}(x_2x_3' - x_3x_2')(t)$, with similar formulas for the other planes.

Note that equations (5) give

$$x_2x_3' - x_3x_2' = \frac{x_1\sqrt{2T}}{I_2I_3}[I_2(I_1 - I_2)x_2{}^2 + I_3(I_1 - I_3)x_3{}^2]$$

and the quantity in brackets is simply $I_1 - D$, as one sees by eliminating $x_1{}^2$ from the last two equations of (1). Similarly for the other two expressions, so that we have

$$
\begin{aligned}
x_2x_3' - x_3x_2' &= \frac{x_1\sqrt{2T}(I_1 - D)}{I_2I_3} \\
x_3x_1' - x_1x_3' &= \frac{x_2\sqrt{2T}(I_2 - D)}{I_3I_1} \\
x_1x_2' - x_2x_1' &= \frac{x_3\sqrt{2T}(I_3 - D)}{I_3I_2}.
\end{aligned} \tag{6}
$$

Now we note that since the vector $\mathbf{L}$ is, in body coordinates,

$$\mathbf{L} = (\omega_1 I_1, \omega_2 I_2, \omega_3 I_3),$$

the cosines of the angles that $\mathbf{L}$ makes with the axes in the body coordinates are

$$\frac{\omega_i I_i}{L} = \frac{I_i\sqrt{2T}x_i}{L} \quad \text{by (4).}$$

In the standard coordinate system, where $\mathbf{L}$ is now the z-axis, these give the cosines of the angle between the z-axis and the coordinate planes in the body, so using (6) we find that the rate of change of the area of projection of the triangle in the polhode cone onto the invariable plane is

$$\frac{2T}{L}\left(\frac{I_1 - D}{I_2I_3}I_1{x_1}^2 + \frac{I_2 - D}{I_3I_1}I_2{x_2}^2 + \frac{I_3 - D}{I_1I_2}I_3{x_3}^2\right).$$

Replacing the ${x_i}^2$ be their values in (3) we find that

$$r^2\frac{d\phi}{dt} = \frac{2T}{L}(r^2 + E) \tag{B}$$

for

$$E = \frac{(I_1 - D)(I_2 - D)(I_3 - D)}{I_1I_2I_3D} = -\sqrt{-a_1a_2a_3D}.$$

From (A) and (B), we then get

$$\frac{dr}{d\phi} = \frac{dr}{dt}\Big/\frac{d\phi}{dt} = \frac{r\sqrt{D}\sqrt{-(r^2 - a_1)(r^2 - a_2)(r^2 - a_3)}}{r^2 + E}.$$

For those adventurous enough to compute the curvature κ, using Problem 6, it will appear that the formula involves positive terms together with the factors $(I_2 + I_3 - I_1)$, $(I_1 + I_3 - I_2)$, and $(I_1 + I_2 - I_3)$, so that $\kappa > 0$ for inertia ellipsoids, where each of this factors is positive.

In the case $D = I_2$, we have $E = 0$ and $a_1 = a_3 = 0$, and

$$\frac{dr}{d\phi} = r\sqrt{I_2}\sqrt{a_2 - r^2},$$

which we can integrate explicitly. Writing

$$d\phi = \frac{dr}{r\sqrt{I_2}\sqrt{a_2 - r^2}}$$

and using the substitution $r = \sqrt{a_2}/u$ we obtain

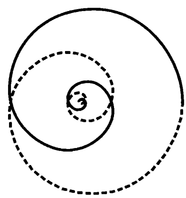
$$d\phi = -\frac{du}{\sqrt{a_2 I_2}\sqrt{u^2 - 1}}$$

so

$$\phi = -\frac{1}{\sqrt{a_2 I_2}}\cosh^{-1} u = \frac{1}{\lambda}\cosh^{-1} u, \quad \text{say,}$$

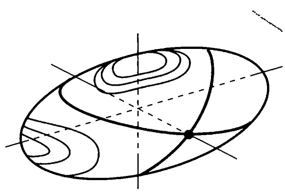
or

$$\frac{\sqrt{a_2}}{r} = \cosh \lambda\phi.$$



The dashed line in the figure comes from negative values of ϕ.

Geometrically, $D = I_2$ means that we are rolling the ellipsoid along one of the two ellipses that make up the polhode. If we start at an intersection of



the two ellipses, the solid spiral in the above picture of the herpolhode is the path obtained by rolling in one direction, the dashed spiral the path obtained by rolling in the other direction. Choosing the other ellipse simply gives the mirror image of this spiral.

This material is adapted from Appell [1; Vol. 2, sect. 393].

PROBLEMS

1. Equations (1) and (2) on page 334 can also be written as

$$I_1\omega_1{}^2 + I_2\omega_2{}^2 + I_3\omega_3{}^2 = 2T$$
$$I_1{}^2\omega_1{}^2 + I_2{}^2\omega_2{}^2 + I_3{}^2\omega_3{}^2 = L.$$

(The fact that the left sides of these equations are constants can be derived directly from the Euler equations: for the first, we multiply the i^{th} Euler equation by ω_i and add; for the second we multiply by $I_i\omega_i$.)

(a) Solve to obtain

$$\omega_1{}^2 = P - Q\omega_2{}^2$$
$$\omega_3{}^2 = R - S\omega_2{}^2$$

for positive P, Q, R, and S, and conclude that

$$\omega_2{}'^2 = \left(\frac{I_3 - I_1}{I_2}\right)^2 (P - Q\omega_2{}^2)(R - S\omega_2{}^2).$$

(b) Transform this to

$$\left(\frac{d\xi}{d\tau}\right)^2 = (1-\xi^2)(1-k^2\xi^2), \qquad \xi = \frac{\omega_2}{\beta}, \quad \tau = pt$$

for positive constants β, p and k. The solution to this equation is the elliptic function sn, so that

$$\omega_2 = \beta\,\mathrm{sn}[p(t - t_0)]$$

for a constant t_0, and ω_1 and ω_3 can then be found in terms of ω_2 (in fact, it turns out that they can be written in terms of the elliptic functions cn and dn). Further details can be found in several sources, including Synge and Griffith [1] and Landau and Lifschitz [1].

2. For the symmetric top, let P be the plane containing $\mathbf{L}$, $\boldsymbol{\omega}$, and $\mathbf{u}_3$.

(a) P rotates with frequency λ around $\mathbf{u}_3$, so the angular velocity of the body around P is $-\lambda\mathbf{u}_3$.
(b) The angular velocity of P in the inertial frame is $\mathbf{L}/I_1$.
(c) The angular velocity of the body in the inertial frame is $\mathbf{L}/I_1 - \lambda\mathbf{u}_3 = \boldsymbol{\omega}$.

3. (a) For a rotating body with $I_1 < I_2 < I_3$, suppose we have $\omega_1 = 0$, $\omega_2 = 0$, $\omega_3 =$ constant, and we consider small changes in ω_1 and ω_2, with ω_3 staying constant. Show that there are solutions $\omega_i(t) = a_i e^{pt}$ with

$$p^2 = \frac{(I_3 - I_2)(I_1 - I_3)}{I_1 I_2}\omega_3{}^2,$$

so that we have a stable oscillatory solution, and a similar result holds for $\omega_1 =$ constant, while in the case of ω_2 we will have an unstable exponentially increasing solution.

Problem 4-12 may now be used to show that in general, rotations with ω_1 or ω_3 constant are stable, while those with ω_2 constant are unstable.

4. Given two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, suppose that $A(t)$ is rotation about $\mathbf{a}$ by the angle $a(t)$ and $B(t)$ is rotation about $\mathbf{b}$ by the angle $b(t)$. Show that the angular momentum vector $\boldsymbol{\omega}$ of $C(t) = B(t)A(t)$ is given by

$$\boldsymbol{\omega}(t) = a'(t) \cdot \mathbf{a} + b'(t) \cdot A(t)(\mathbf{b}),$$

and generalize to multiple compositions.

5. (a) Recall that for a curve given in polar coordinates by $r = r(\phi)$, i.e., parameterized by

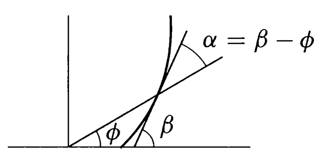
$$x(\phi) = r(\phi)\cos\phi, \quad y(\phi) = r(\phi)\sin\phi,$$

the slope of the tangent line is

$$\frac{r(\phi)\cos\phi + r'(\phi)\sin\phi}{-r(\phi)\sin\phi + r'(\phi)\cos\phi}.$$

If this tangent line makes an angle of β with the horizontal axis, so that $\alpha = \beta - \phi$ is the angle between the tangent line and the line from the origin to the point, then

$$\tan\alpha = \tan(\beta - \phi) = \frac{r(\phi)}{r'(\phi)}.$$



In Leibnizian notation we have

$$\tan\alpha = r/(dr/d\phi) = r\frac{d\phi}{dr},$$

which thus gives $\tan\alpha$ when we instead consider "ϕ as a function of r".

(b)[1] Consider the projection of the curve traced out by the axis of a top onto the (x, y)-plane, where for the functions $r(t)$, $\phi(t)$ we have

$$(r, \phi) = (\sin\phi, \phi) = (\sqrt{1-\cos^2\phi}, \phi) = (\sqrt{1-u^2}, \phi).$$

Opportunistically mixing $'$ notation for derivatives with respect to time with Leibnizian notation, show that

$$\begin{aligned}\tan\alpha &= r\frac{d\phi}{dr} = r\frac{d\phi}{dt}\frac{dt}{dr}\\ &= \frac{(u^2-1)\phi'}{uu'}.\end{aligned}$$

Recalling that $b = a\cos\phi = au_1$ (page 350), we see that the formula for $\tan\alpha$ has a factor of $(u - u_1)$ in the numerator, while the denominator has only the factor

$$uu' = u\sqrt{(u-u_1)(u-u_2)(u-u_3)},$$

so that when $u = u_1$ we have $\tan\alpha = 0$, i.e., the tangent line is pointing radially toward the origin.

6. From the equations $x(\phi) = r(\phi)\cos\phi$, $y(\phi) = r(\phi)\sin\phi$ in Problem 5, find both the first and second derivatives of x and y, and then use the formula

$$\kappa = \frac{x'y'' - y'x''}{(x'^2 + y'2)^{3/2}}$$

for the curvature of a curve $t \mapsto (x(t), y(t))$ to deduce that

$$\kappa = \frac{r^2 + 2r\left(\dfrac{dr}{d\phi}\right)^2 - r\dfrac{d^2r}{d\phi^2}}{\left(r^2 + \left(\dfrac{dr}{d\phi}\right)\right)^{3/2}}.$$

428
7. (a) If all principal moments are equal to I, so that $\boldsymbol{\tau} = I\boldsymbol{\omega}$ (page 190), use the vector form $(\mathbf{E}_{\boldsymbol{\tau}})$ of the Euler equations to show that $\boldsymbol{\tau} = \mathbf{L}'$ also holds *in body coordinates*.
(b) Also deduce this from the equations (E_τ).

---

[1] From Cabannes [1].

429 **8.** The situation considered in Addendum A, where $I_1 = I_2$, can also be approached somewhat differently. Let $\boldsymbol{\omega}$ be the angular velocity of the rotating orthonormal bases $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ and let $\mathbf{L} = I_1\omega_1 \cdot \mathbf{u}_1 + I_1\omega_2 \cdot \mathbf{u}_2 + I_3\omega_3 \cdot \mathbf{u}_3$ be the expression for the angular momentum with respect to these axes. If α is the angular velocity of fixed axes in the body with respect to these axes, then the angular momentum $\mathbb{L}$ with respect to fixed axes in the body is

$$\mathbb{L} = \mathbf{L} + \alpha I_3 \mathbf{e}_3.$$

Use $\boldsymbol{\tau} = \mathbb{L}' + \omega \times \mathbb{L}$ to show that the Euler equations become

$$\begin{aligned}
\tau_1 &= I_1\omega_1{}' + (I_3 - I_1)\omega_2\omega_3 + I_3\alpha\omega_2 \\
\tau_2 &= I_1\omega_2{}' + (I_1 - I_3)\omega_3\omega_1 - I_3\alpha\omega_1 \\
\tau_3 &= I_3(\omega_3{}' + \alpha'),
\end{aligned}$$

or deduce the results directly from the equation (T) on page 362.

CHAPTER 10

NON-INERTIAL SYSTEMS AND FICTITIOUS FORCES

In Chapter 9 we derived equations for the motion of a body in terms of a coordinate system located within the body itself. More generally, we now consider what the equations of motion for any particle become in an arbitrarily moving, usually non-inertial, coordinate system. One of the main reasons for this investigation is that the behavior of numerous common phenomenon depend on the fact that they are really being observed in such a moving coordinate system, namely one located on the rotating earth.

The basic equations. For the case of rotating coordinate systems, we derived in Chapter 9 the basic formula

$$\mathbf{r}' = \mathbf{r}\text{'} + \boldsymbol{\omega} \times \mathbf{r}$$

for any curve $\mathbf{r}$. Applying this to a particle $\mathbf{x}$, and then extending our computations, we have

$$\begin{aligned}
\mathbf{x}' &= \mathbf{x}\text{'} + \boldsymbol{\omega} \times \mathbf{x} \\
\mathbf{x}'' &= (\mathbf{x}\text{'})' + \boldsymbol{\omega}' \times \mathbf{x} + \boldsymbol{\omega} \times \mathbf{x}' \\
&= (\mathbf{x}\text{''} + \boldsymbol{\omega} \times \mathbf{x}\text{'}) + \boldsymbol{\omega}' \times \mathbf{x} + \boldsymbol{\omega} \times (\mathbf{x}\text{'} + \boldsymbol{\omega} \times \mathbf{x}) \\
&= \mathbf{x}\text{''} + 2 \cdot \boldsymbol{\omega} \times \mathbf{x}\text{'} + \boldsymbol{\omega}' \times \mathbf{x} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{x}).
\end{aligned}$$

We are going to be working almost entirely in the rotating coordinate system, so we will set $\mathbf{x}\text{'} = \mathbf{v}$, with the understanding that $\mathbf{v}$ denotes the velocity as computed in the rotating coordinate system, and similarly we will set $\mathbf{x}\text{''} = \mathbf{a}$, where the acceleration $\mathbf{a}$ is also computed in this rotating coordinate system. Multiplying our equation by the mass m of $\mathbf{x}$, and rearranging, we have

$$m\mathbf{a} = m\mathbf{x}'' - m \cdot \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{x}) - 2m \cdot \boldsymbol{\omega} \times \mathbf{v} - m \cdot \boldsymbol{\omega}' \times \mathbf{x},$$

and if $\mathbf{F}$ is the force acting on $\mathbf{x}$ we can write this as

$$\boxed{m\mathbf{a} = \mathbf{F} - m \cdot \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{x}) - 2m \cdot \boldsymbol{\omega} \times \mathbf{v} - m \cdot \boldsymbol{\omega}' \times \mathbf{x}}$$

This equation says that, using measurements in the rotating coordinate system, the particle behaves as if it were under the influence of the force **F** together with three other "fictitious forces":

$-m \cdot \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{x})$	the centrifugal force
$-2m \cdot \boldsymbol{\omega} \times \mathbf{v}$	the Coriolis force
$-m \cdot \boldsymbol{\omega}' \times \mathbf{x}$	the azimuthal or Euler force

If our moving coordinate system includes translations, with the origin being translated to a new origin $\mathbf{b}(t)$ at time t, then we will have one more term in the equation, giving another "fictitious force",

$-m \cdot \mathbf{b}''$	the translational or acceleration force.

These "fictitious forces", always having the factor m, are just the forces that would be needed to account for the not-at-all-fictitious correction terms that need to be made to the acceleration because our observations are made in the moving coordinate system. They are useful, though often confusing, theoretical constructs precisely because, as mentioned at the beginning of this chapter, we sometimes do make observations in a convenient coordinate system that is not an inertial coordinate system, the rotating earth being a prime example.

The translational or acceleration force. This force is easily envisioned by imagining that you are sitting on a totally frictionless flat cart at rest; if the cart accelerates, you remain at rest, so you will seem to accelerate backwards to the cart—in the cart's coordinate system, it *appears* (if the observers on the cart are insensitive to the fact that *they* are being accelerated) that there is a force acting on you, of which you are blissfully unaware. If the cart has a back to lean against, so that you don't fall off, then you will feel a force, which is not this fictitious translational force, but the opposite force that keeps you on the cart.

A more involved example involves what happens when we push a container of water along a surface, first accelerating it to get it moving, and then letting it decelerate to come to a stop; the water will tilt toward the rear of the container

as the container accelerates, and tilt in the other direction as it decelerates. The reason is that the water is subject to the force of gravity, straight downwards, plus a horizontal fictitious force, so that the resultant is at an angle, and the effect is the same as if gravity were actually acting at that angle.

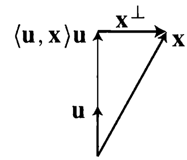
At first this example seems to be a typical case where the physicist's way of looking at things is confoundingly different from the mathematician's; we seem to have resorted to a physics trick instead of analyzing the problem in terms of the laws of motion. But we're actually in no position to make such an analysis, since we currently have no principles for dealing with a liquid. In fact, even in the case where there is no acceleration at all we haven't discussed any principles to demonstrate that the surface of the water will remain horizontal!

But even if we can't give an analysis, the physicist recognizes a principle, and the mathematician should recognize what might be called an "invariant" of the problem: the surface of the water at any point is always perpendicular to the total force at that point.

Of course, as good physicists and/or mathematicians, we should seek to generalize this example, and the next force will play a role in this.

The centrifugal force. The complicated looking formula for the centrifugal force can be expressed more simply geometrically. If $\mathbf{u}$ is a unit vector along $\boldsymbol{\omega}$, and $\omega = |\boldsymbol{\omega}|$, then the centrifugal force is (recall the formula on page 189)

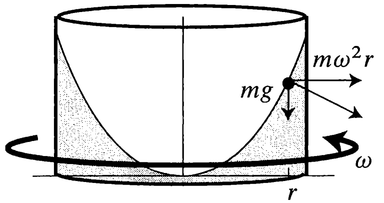
$$\begin{aligned} -m\cdot\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x}) &= -m\cdot\big(\langle\boldsymbol{\omega},\mathbf{x}\rangle\boldsymbol{\omega}-\omega^2\mathbf{x}\big)\\ &= m\omega^2\big(\mathbf{x}-\langle\mathbf{u},\mathbf{x}\rangle\mathbf{u}\big)\\ &= m\omega^2\mathbf{x}^{\perp}, \end{aligned}$$



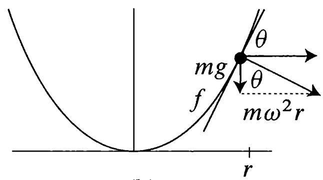
where $\mathbf{x}^{\perp}$ is the vector from the line along $\boldsymbol{\omega}$ perpendicular to $\mathbf{x}$, so the magnitude of the centrifugal force depends on the distance from the axis of rotation, and is directed away from it.

If you are revolving rapidly with a weight suspended by a string in front of you, and then release the string, the weight is now free to move in a straight line, and seems to fly away from you. In your (rotating) coordinate system it acts *as if* there is a force on it, the centrifugal force. Until the time of release, the weight has stayed in front of you because of a counteracting actual force that you exert on it, the "centripetal" force. By the third law, the weight is exerting an equal and opposite force on your hand. Unfortunately, the name "centrifugal force" is often applied to *this* force, leading people to think that they are "feeling the centrifugal force", and wondering why it should be called fictitious—yet another example of how easily the notion of fictitious forces can lead to confusion. It might help to speak of the translational and the centrifugal "acceleration corrections", remembering that we multiply them by mass to get forces, which we then of course divide by the mass to get accelerations.

To generalize our previous example involving liquids, suppose that we rotate a cylindrical container of water with angular speed ω. Then a particle of mass m



(a)



(b)

on the surface of the water at distance r from the center is subject to a downward force mg and a centrifugal force of $m\omega^2 r$, as in (a). Our invariant says that the resultant force should be perpendicular to the surface of the water, whose profile curve is the graph of some function f. From (b) we see that

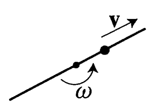
$$f'(r) = \tan\theta = \frac{\omega^2 r}{g},$$

and thus

$$f(r) = \frac{\omega^2}{2g} \cdot r^2,$$

and the profile curve is a parabola.

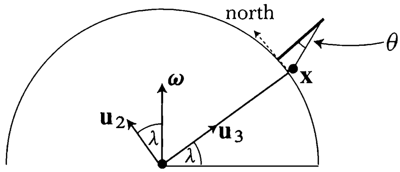
Another simple example of centrifugal force arises for the problem of a bead on a rotating rigid wire, which we considered on page 227, where the angular rotation is $\omega = |\boldsymbol{\omega}|$ for a vector $\boldsymbol{\omega}$ pointing out from the plane of the paper. In



the rotating coordinate system, the only force along the wire is the centrifugal force on the bead, which is $m\omega^2 x$ when the bead is at distance x from the center (the other two forces are perpendicular to the wire, since both $\mathbf{x}$ and $\mathbf{v}$ are along the wire). So the equation of motion is simply $x'' = \omega^2 x = \theta'^2 x$, as obtained previously.

The deflection of a hanging body. Centrifugal force also made an unheralded appearance in our discussion of the Eötvös experiment in Addendum 1B, the whole point of this experiment being that the proportionality of mass to weight amounts to saying that weight acts *just like a fictitious force*, so we can test the proportionality by comparing it to the centrifugal force, which is definitely proportional to mass, since it *is* a fictitious force.

To calculate the deflection θ of a hanging body, like that used in the Eötvös experiment, we consider the hanging particle $\mathbf{x}$ hovering just above the earth's



surface, at latitude λ, and a rotating coordinate system $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ at the center of the earth defined by letting $\mathbf{u}_3$ point toward $\mathbf{x}$ and choosing $\mathbf{u}_2$ perpendicular to it, parallel to the direction of north at $\mathbf{x}$; then $\mathbf{u}_1$ is parallel to the direction of east at $\mathbf{x}$, so that in the picture it points into the plane of the paper. The rotation vector of the earth, $\boldsymbol{\omega} = (0, 0, \omega)$, can be written as

$$\boldsymbol{\omega} = 0 \cdot \mathbf{u}_1 + \omega \cos\lambda \cdot \mathbf{u}_2 + \omega \sin\lambda \cdot \mathbf{u}_3.$$

If R is the radius of the earth, then $\mathbf{x} = R \cdot \mathbf{u}_3$, so

$$\begin{aligned} \boldsymbol{\omega} \times \mathbf{x} &= \boldsymbol{\omega} \times R \cdot \mathbf{u}_3 = \omega R \cos\lambda \cdot \mathbf{u}_1 \\ -m \cdot \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{x}) &= -m\omega^2 R(\cos\lambda \sin\lambda \cdot \mathbf{u}_2 - \cos^2\lambda \cdot \mathbf{u}_3). \end{aligned}$$

The gravitational force on $\mathbf{x}$ is $-mg \cdot \mathbf{u}_3$, so the total force on $\mathbf{x}$ is

$$\begin{aligned} (1) \qquad \mathbf{F} &= -mg \cdot \mathbf{u}_3 - m\omega^2 R(\cos\lambda \sin\lambda \cdot \mathbf{u}_2 - \cos^2\lambda \cdot \mathbf{u}_3) \\ &= -m\omega^2 R \cos\lambda \sin\lambda \cdot \mathbf{u}_2 - m(g - \omega^2 R \cos^2\lambda) \cdot \mathbf{u}_3. \end{aligned}$$

So the downwards acceleration is decreased from g to $g - \omega^2 R \cos^2\lambda$, resulting in a teensy decrease in weight as we approach the equator, and a small southward force is added, resulting in a southward deflection θ with

$$\tan\theta = \frac{\omega^2 R \cos\lambda \sin\lambda}{g - \omega^2 R \cos^2\lambda}.$$

Even though R is large, the angular velocity $\omega = 2\pi/(24 \cdot 3600)$ radians per second is so small that $\omega^2 R/g$ has the small value $\omega^2 R/g \approx .00344$, so that we can write

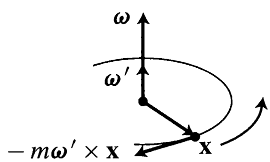
$$\tan\theta \approx \frac{\omega^2 R}{g} \cos\lambda \sin\lambda.$$

Computations show that for a plumb bob hung from a height of 50 meters at latitude 45° (the leaning tower of Pisa is a good approximation), the bob should end up about 8.5 cm south, which one might assume is easily measurable on a windless day. But there's actually no "origin" from which to measure this distance, since "straight down" is always *determined by a plumb bob*! What this comes down to is that for actual measurements, the direction we choose for $\mathbf{u}_3$ won't really be exactly in the radial direction, and for later use we should really write (1) as

$$(1') \qquad \mathbf{F} = m\mathbf{g} = mg \cdot \mathbf{u}_3,$$

where $g = |\mathbf{g}|$ is the acceleration of gravity that we measure at a particular spot on earth, and where $\mathbf{g}$ points in the direction of a plumb bob, which is the direction we will really have ended up choosing for $\mathbf{u}_3$. Actually, the bulging equator of the rotating earth causes an additional deflection of the plumb bob, and this too is taken into account in equation (1′).

The azimuthal or Euler force. A simple example of the azimuthal force arises when you are in a car $\mathbf{x}$ accelerating as it goes around a curve, so that in addition to $\boldsymbol{\omega}$, giving the centrifugal force that seems to push you towards the outside

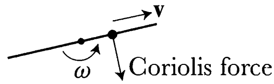


of the car, there is a non-zero $\boldsymbol{\omega}'$. This gives the additional force $-m\boldsymbol{\omega}' \times \mathbf{x}$, which seems to push you towards the back of the car, a rotational analogue of the translational or acceleration force.

This force doesn't play an important role for observations on the rotating earth, because the earth's rotation $\boldsymbol{\omega}$ is so nearly constant, though as we've seen in the discussion of the Euler precession (page 339), $\boldsymbol{\omega}$ moves along a small cone around the line to the North star, so that $\boldsymbol{\omega}'$ is a tiny vector pointing inward. This might explain the name "Euler force", or it might have been given simply because the whole equation on page 376 had actually been deduced by Euler very early on (cf. Persson [1; pg. 15]). One hardly ever sees the Euler force mentioned in mechanics problems, but it makes a surprise appearance in the final chapter.

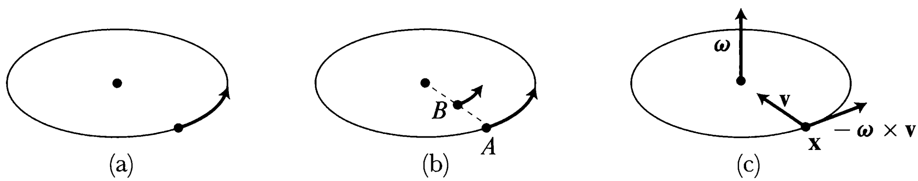
The Coriolis force. Unlike the previous fictitious forces, the Coriolis force on a particle depends not on its position but on its *velocity*. It was introduced by Coriolis in a paper of 1835, expanding on one from 1832, both to be found in Coriolis [1]. This work originated with the study of water-wheels according to Dugas [1], where a description of Coriolis' work shows how amazingly complicated the origins of a simple idea can be (cf. also page 420).

As a simple example of the Coriolis force, consider the case of the rotating bead. Here the Coriolis force is perpendicular to the wire, with magnitude $2m\omega|\mathbf{v}|$. The wire will have to exert the negative of the Coriolis force on the



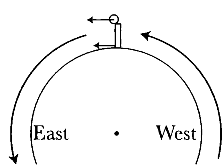
bead in order to keep it moving around, so as the bead moves both further away and faster, the wire must be stiffer to resist bending.

A more involved situation arises when you are on a rotating platform, like a carousel, as in (a) of the figure below. If you are facing the center, you will have to dig in your heels to prevent yourself from being thrown off by the centrifugal force. An additional, Coriolis, force will arise if you try to walk



towards the center, going from A to B in (b) of the figure, basically because you are already moving in a tangential direction, and the point B that you are trying to reach is moving less quickly in this direction. You will have to exert pressure on the right side of your feet to tack left, giving the impression that you are moving against a force to the right, which is the direction of $-\boldsymbol{\omega} \times \mathbf{v}$, as shown in (c); you can feel the effect at one blow by jumping from a wooden horse on the carousel to one nearer the center. Also see the amusing movie at `ww2010.atmos.uiuc.edu/(Gh)/guides/mtr/fw/gifs/coriolis.mov`.

The deflection of a *falling* body. One significant consequence of the differing tangential speeds at different distances from the center of rotation was pointed out by Newton in 1679 in a brief reply to a letter from Hooke (see Newton [1; Vol. 2, pg. 301]). Because of the controversy over the Copernican theory, there had been a history of experiments to prove the earth's rotation by dropping objects from a height to observe a supposed westward deflection during the fall, the futility of which had already been explained by Galileo (cf. Chapter 7). Newton pointed out that not only does an object dropped from a tall building already have a horizontal motion east but also, since this horizontal motion is slightly greater than the horizontal motion of the earth below it (as



Galileo had also noted), the object should actually end up very slightly to the east of that building when it hit the earth, "quite contrary to the opinion of the vulgar who think that if the earth moved, heavy bodies in falling would be outrun by its parts & fall on the west side of the perpendicular." Newton proposed a means of detecting such a tiny deflection to the east, by comparing the distribution of many dropped objects, though Hooke's trials proved quite inconclusive.

A naive calculation based on Newton's observation might be the following. For a building of height h, the time of descent is very close to

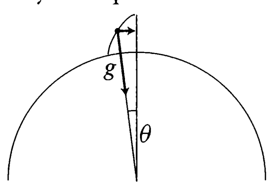
$$T = \sqrt{2h/g},$$

the usual answer when we ignore the earth's turning. If the building is at latitude λ, and ω is the angular velocity of the earth, then the body's horizontal motion is greater than that of the earth's surface by a factor of $h\omega(\cos\lambda)$. So it seems it should fall to the east by the amount $h\omega(\cos\lambda)T = h\omega(\cos\lambda)\sqrt{2h/g}$ or

$$\text{(E)} \qquad \omega(\cos\lambda)\sqrt{2h^3/g}.$$

This computation actually assumes that the body and the base of the tower are each moving in straight lines, rather than circular arcs, but for such short distances the difference is presumably negligible.

On the other hand, there is another effect that we might not notice, and might assume was also negligible even if we did notice it. When the falling body has reached a position making a (very small) angle θ with the perpendicular at the center of the earth, gravity will produce an acceleration of magnitude g



along the line to the center of the earth, and this acceleration will have a tiny horizontal westward component of

$$\mathrm{acc}_{\mathrm{west}} = (\cos\lambda)g\sin\theta \approx (\cos\lambda)g\theta = (\cos\lambda)g\omega t.$$

So the total westward motion during the descent, from time 0 to time T, will be

$$\text{(W)} \qquad (\cos\lambda)g\omega\cdot\frac{1}{6}T^3 = \frac{1}{3}(\cos\lambda)\omega\sqrt{2h^3/g},$$

and the total eastward deflection should actually be the difference (E) – (W),

$$\frac{2}{3}(\cos\lambda)\omega\sqrt{2h^3/g}.$$

This calculation naturally leaves us a little queasy—who knows what else we might have ignored!—and a very easy alternate solution is given in Problem 1.

It isn't clear whether any such calculations were even made before the 19th century.[1] The first calculations usually mentioned were made in 1803, when the question of the path of a falling object was analyzed by Laplace [1; Vol. 14, 267–277] and independently by the younger Gauss [1; Vol. 5, 498–503], for the benefit of the experimenter Benzenberg, who eventually measured deflections in a deep mine shaft.

Gauss elegantly analyzed the problem by translating the equations of motion in an inertial system into equations for a system on the rotating earth, essentially a derivation of our basic equation on page 376 for the case of uniform rotation. We will continue to use the coordinate system on page 380, and write vectors in terms of these coordinates, so that (a, b, c) will stand for $a \cdot \mathbf{u}_1 + b \cdot \mathbf{u}_2 + c \cdot \mathbf{u}_3$. Note that, as explained on page 381, the third axis simply points along the direction of a plumb bob, whose lowest position is what we use as the point from which to measure the deflection. Since we are now dealing with a moving particle, we must amend equation (1′) on page 381 to read

$$\mathbf{F} = (0, 0, -mg) + \mathbf{F}_{\text{Coriolis}}.$$

The Coriolis force $\mathbf{F}_{\text{Coriolis}}$ for the particle $(x(t), y(t), z(t))$, with velocity vector $\mathbf{v}(t) = (x'(t), y'(t), z'(t))$, is given by

$$\begin{aligned} \mathbf{F}_{\text{Coriolis}} &= -2m \cdot \boldsymbol{\omega} \times \mathbf{v} \\ &= -2m\omega \cdot (0, \cos\lambda, \sin\lambda) \times (x', y', z') \\ &= -2m\omega(z'\cos\lambda - y'\sin\lambda,\ \ x'\sin\lambda,\ \ -x'\cos\lambda), \end{aligned}$$

so that the particle (x, y, z) satisfies

$$\begin{cases} x'' = -2\omega(z'\cos\lambda - y'\sin\lambda) \\ y'' = -2\omega x'\sin\lambda \\ z'' = 2\omega x'\cos\lambda - g. \end{cases}$$

Although we could find useful approximate solutions directly (Problem 2), these particular equations can actually be solved exactly. For a dropped object, which has the initial conditions $0 = x(0) = y(0)$ and $0 = x'(0) = y'(0) = z'(0)$, the last two equations give

$$\begin{aligned} y'(t) &= -2\omega(\sin\lambda)x(t) \\ z'(t) &= 2\omega(\cos\lambda)x(t) - gt, \end{aligned}$$

and substituting these into the first equation then gives

$$x''(t) + 4\omega^2 x(t) = 2\omega g(\cos\lambda)t,$$

[1] Newton's letter doesn't include one, though Arnold [3; pp. 19–20] seems to assume that Newton's observation amounts to the "naive calculation" on the previous page.

with the particular solution $x(t) = g(\cos\lambda)t/(2\omega)$, and the general solution

$$x(t) = \frac{g\cos\lambda}{2\omega}\cdot t + A\sin 2\omega t + B\cos 2\omega t.$$

The initial condition $x(0) = 0$ gives $B = 0$ and then the initial condition $x'(0) = 0$ gives $A = -g\cos\lambda/(4\omega^2)$, so that

$$x(t) = \frac{g\cos\lambda}{2\omega}\left(t - \frac{\sin 2\omega t}{2\omega}\right).$$

Having obtained the correct exact formula, it is now easy to make a useful approximation. We have $2\omega t \ll 1$ for a drop from any reasonable height, so we can use the approximation

$$\sin 2\omega t \approx 2\omega t - \frac{(2\omega t)^3}{6},$$

giving the approximate equation

$$x(t) = \frac{1}{3}g\omega(\cos\lambda)\cdot t^3.$$

Letting T be the total time of descent, $T \approx \sqrt{2h/g}$, we find that the total displacement $x(T)$ in the $\mathbf{u}_1$ direction, east, is approximately

$$\frac{2}{3}\omega(\cos\lambda)\sqrt{2h^3/g}.$$

It might be of interest to note that Gauss, in a letter to Benzenberg, expressed great surprise at having found "a deflection to the east only $\frac{2}{3}$ of that which Dr. Olbers has found." The physician Olbers, a good friend of Gauss, made several important astronomical discoveries, though he is probably most famous for "Olbers' paradox" that the night sky should be infinitely bright. Perhaps his calculation of the deflection was the naive calculation on page 383, without the added correction—I have the sneaking suspicion that this correction wasn't noticed until after the calculations of Laplace and Gauss had been made.[1]

[1] Just to make everything a bit more complicated, Gauss' equations had extra terms to account for the change of direction and magnitude of $\mathbf{g}$, but he excluded them as being insignificant for an almost vertical fall; note, however, that this refers to the very slight variation of direction of $\mathbf{g}$ in the *rotating* coordinate system, due to the fact that the object passes over different positions above the not exactly spherical earth.

The southward deflection. Since there is a factor of ω in the formula for x, and thus for x', our original equation for y'' will have a factor of ω^2, and thus the southward deflection will be negligible. But see Addendum B!

Using the formula for $x(t)$ to get an approximation for $x'(t)$ gives $y'' \approx -g\cos\lambda\sin\lambda(1-\cos 2\omega t) \approx -2g\omega^2\cos\lambda\sin\lambda\cdot t^2$; with initial conditions $y(0) = y'(0) = 0$, we find[1] that $y(T)$ is $(2/3)gh^2\omega^2\cos\lambda\sin\lambda$.

Stupid experimenter tricks. A related problem, which also vexed the "vulgar", giving rise to tales of experiments that one hopes are apocryphal, concerns the behavior of a cannonball shot directly vertically upwards. If the projectile has an initial speed of v_0, the only difference in the above calculations is that the equation for z' on page 384 becomes

$$z'(t) = 2\omega(\cos\lambda)x(t) - g\cdot t + v_0$$

leading to

$$x''(t) + 4\omega^2 x'(t) = 2\omega g\cos\lambda\,(t - v_0/g),$$

with the general solution

$$x(t) = \frac{g\cos\lambda}{2\omega}\left(t - \frac{v_0}{g}\right) + A\sin 2\omega t + B\cos 2\omega t.$$

The initial condition $x(0) = 0$ gives $B = v_0\cos\lambda/(2\omega)$, and $x'(0) = 0$ gives the same result as before for A, so that

$$x(t) = \frac{g\cos\lambda}{2\omega}\left(t - \frac{v_0}{g}\right) - \frac{g\cos\lambda}{4\omega^2}\sin 2\omega t + \frac{v_0\cos\lambda}{2\omega}\cos 2\omega t.$$

The Taylor series for $\sin 2\omega t$ and $\cos 2\omega t$ then give the approximate equation

(a) $$x(t) = \frac{1}{3}g\omega\cos\lambda\cdot t^3 - \omega v_0\cos\lambda\cdot t^2.$$

The total time from the firing of the projectile to its landing back on earth is $T = 2v_0/g$, or $T = 2\sqrt{2h/g}$ if h is its maximum height, and we find that

$$x(T) = -\frac{4}{3}\omega\cos\lambda\sqrt{8h^3/g}.$$

Thus, the deflection of the cannonball is to the *west*—by an extremely tiny amount that won't change the denouement of the experiment.

Note that the approximation (a) might lead us to the equation

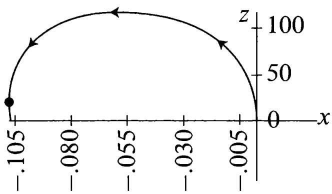
$$x'(t) = g\omega\cos\lambda\cdot t(t - 2v_0/g),$$

so that $x'(t) < 0$ for $t < 2v_0/g$, which is the entire time of the trip except for the last moment, implying that the projectile is heading westward throughout

---

[1] A more careful calculation (with refinements involving terms of order ω^2 that are insignificant for the value of the eastward deflection) yields the answer $4gh^2\omega^2\cos\lambda\sin\lambda$; see Belorizky and Sivardière [1] for details.

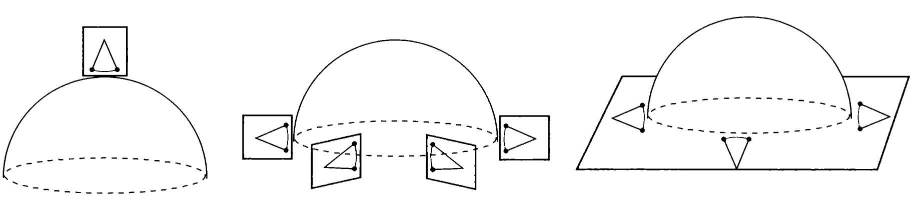
the trip, except that $x'(2v_0/g) = 0$, with the projectile landing exactly vertically. Thus, the eastward Coriolis force as the projectile falls from the highest point seems to exactly cancel out the westward velocity that has been acquired at the highest point. Actually, the computer generated graph of the path shown below indicates that the maximum westward position occurs a very short time



before the end of the trip, which ends with a tiny eastward velocity. An analytic verification from the exact equations might be rather unpleasant.

Foucault's pendulum. Coriolis had been a student of Poisson, who soon applied the results of Coriolis' paper to calculate the deflection of artillery shells, which turned out to be less than those due to wind and other effects for the artillery of that time (though not for modern artillery), and Poisson likewise decided that the effect on a pendulum was too small to be observable. But the enterprising experimenter Foucault realized that with a spherical pendulum, allowing the direction of swing to change with time, one could demonstrate the small Coriolis force caused by the rotation of the earth in the way pendulums have always been exploited, by observing the cumulative effect of many cycles. He eventually created a 67 meter pendulum, with a bob weighing 28 kilograms, in the Panthéon in Paris, which created a sensation at the Paris exhibition of 1851, where the plane in which the pendulum was swinging could clearly be seen to rotate over time, allowing people to "see the earth go round".

Of course, it wasn't so clear just what people were actually seeing, since the popular explanation for Foucault's pendulum was (and still is) that the plane of the pendulum's swing stays fixed, while the earth turns beneath it. This would be true for a pendulum at the north pole (a), where the clockwise rotation would have the same magnitude ω as the earth's counterclockwise rotation, but at the equator, the plane of the pendulum's swing appears stationary to an observer, and if the pendulum is swinging in the north-south direction (b), this

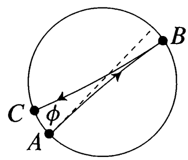


(a) (b) (c)

plane is rotating in space right along with the earth (the case of an east-west swing (c) is usually used instead, to deviously bolster the erroneous idea that the plane of the pendulum's swing is always fixed).

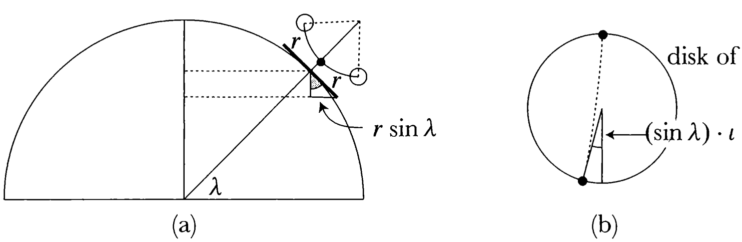
Foucault recognized that the situation was more complicated, and intuited that at latitude λ the rate of change of the pendulum's swing should be $\omega \sin \lambda$, certainly a good guess for a quantity that varies from ω to 0 as the latitude changes from 90° to 0! But intuitive, basically geometric, arguments for the $\sin \lambda$ formula have never been entirely convincing—for good reason as we shall see.

First of all, we need a general idea of what the path of the pendulum bob will look like to an observer on the earth. In the figure below, our pendulum starts its swing at A. Instead of moving along the diameter (dashed line), it will be deflected slightly to the east because of the Coriolis force, arriving at the



point B, where its velocity is 0 and it begins to swing back in the other direction. This makes the Coriolis force change direction also, so the bob is now deflected in the other direction, ending at C. Thus the bob will seem to have rotated around the circumference of the circle by an angle ϕ as it goes from A to C. Of course, the angle ϕ has been grossly exaggerated in this figure. Nevertheless, in Foucault's Panthéon demonstration, the distance from A to C was almost a centimeter, and the rotation was observable in a short time.

For the special case where the pendulum's swing is exactly north-south over the disk of radius r, as shown in (a) of the figure below, we can determine the amount of rotation geometrically.[1] The shaded angle in (a) is also λ, so

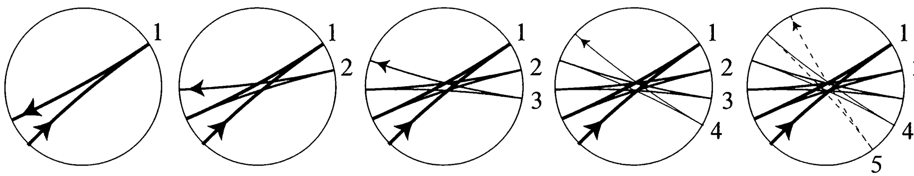


the southernmost point of the disk is further from the axis of the earth's rotation than the center of the disk by the amount $r \sin \lambda$. If the center of the disk rotates by the very small angle ι during one swing of the pendulum, then the southernmost point rotates an extra $(\sin \lambda) \cdot \iota$ counterclockwise along

[1] From Kittel *et al.* [1].

the circumference of the disk, so the position of the pendulum bob, as observed by some one standing on the earth, looking down at the disk (b), has moved clockwise along the circumference by this amount.

Now one might hope that the rotation is the same for any direction of the pendulum's swing. In fact, textbooks usually show the full path of the pendulum bob, seen on the earth, as made up of repetitions of the basic piece of the picture on the previous page, just as for the orbits discussed on page 128. But this



picture can't be *exactly* correct, because the Coriolis force, involving the cross-product with $\boldsymbol{\omega} = \omega(0, \cos\lambda, \sin\lambda)$, does in fact depend on the direction of the pendulum's swing.

The details will emerge as a by-product of our analytic computation, which actually involves approximations right from the start—not all that surprising, since approximations are needed even for an ordinary pendulum. We will use the same coordinate system $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ as before, and consider only small oscillations of the pendulum, of length l, say. If we ignore the Coriolis force, then, as on page 291, the coordinates $x(t), y(t)$ of the pendulum bob satisfy

$$\begin{aligned} x'' &= -\alpha^2 x \\ y'' &= -\alpha^2 y \end{aligned} \qquad \text{for } \alpha = \sqrt{g/l}.$$

When adding in the Coriolis force $\mathbf{F}_{\text{Coriolis}}$ on page 384, we note (aha!) that in comparison to x' and y', the quantity z' is small,[1] so we gleefully hasten to discard it, obtaining approximate equations conveniently free of the $\cos\lambda$ term,

$$(*) \qquad \begin{aligned} x'' &= -\alpha^2 x + (2\omega \sin\lambda) y' \\ y'' &= -\alpha^2 y - (2\omega \sin\lambda) x'. \end{aligned}$$

Thus, $\cos\lambda$, the y-component of the Coriolis force, has a negligible role because its effect in the (x, y)-plane depends on the z component of the velocity, which is negligible.

[1] At any time t_0 we can rotate our x and y axes far from the direction of the pendulum swing, eliminating the problem of x' or y' being very small, or even zero.

Now we use the trick in the footnote on page 339, custom-made for a coupled pair of equations. If we set $\zeta = x + iy$, we obtain

$$\zeta'' + i(2\omega \sin \lambda)\zeta' + \alpha^2 \zeta = 0,$$

which is just the equation for damped oscillations, in this case for a function that is complex-valued to begin with. Setting $\zeta(t) = e^{\rho t}$, this becomes

$$\rho^2 + i(2\omega \sin \lambda)\rho + \alpha^2 = 0.$$

Since $\omega^2 \ll \alpha^2$, we have

$$\rho \approx -i(\omega \sin \lambda) \pm i\alpha,$$

and thus

$$\zeta(t) \approx e^{-i(\omega \sin \lambda)t} \cdot (ae^{i\alpha t} + be^{-i\alpha t}).$$

The second factor describes the motion of the spherical pendulum of Chapter 8, which will simply be an ordinary pendulum motion if we release it without any sidewise motion (Foucault obtained this condition by attaching the bob to a stationary point with a thread, and then burning the thread). So we are considering only a and b that reduce the second factor to $c \cos \alpha t$ for a constant c. The solution for ζ then yields approximate equations

$$\begin{aligned} x(t) &= \cos\big(-(\omega \sin \lambda)t\big) \cdot c \cos(\alpha t) \\ y(t) &= \sin\big(-(\omega \sin \lambda)t\big) \cdot c \cos(\alpha t), \end{aligned}$$

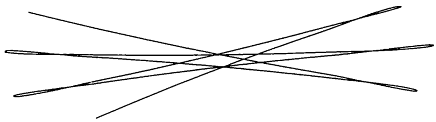
which we might write as

$$(**) \qquad \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos\big(-(\omega \sin \lambda)t\big) \\ \sin\big(-(\omega \sin \lambda)t\big) \end{pmatrix} \cdot c \cos(\alpha t).$$

Since $\omega \sin \lambda \ll \alpha$, we have something like the phenomenon of "beats": the pendulum swings with angular velocity α in a plane that is rotating with the slow angular velocity $-\omega \sin \lambda$. This rotation is clockwise because of the minus sign (counterclockwise in the southern hemisphere, where ω is replaced by $-\omega$).

Because $(*)$ is not exact, the angle by which the pendulum advances actually varies ever so slightly from swing to swing; thus, $\sin \lambda$ is merely an extremely good approximation for the factor to be applied to the average of many nearly equal advances.

Although we pictured the path of the Foucault pendulum as having cusps, in practice it is extremely hard to guarantee that the second factor of our solutions $(**)$ will be exactly equal to $\cos \alpha t$; even without the Coriolis force acting, the motion of the bob, projected on the horizontal plane, will usually be an extremely narrow ellipse rather than degenerating precisely into a straight line segment, so when the Coriolis force is added in, the cusps would actually have very tiny loops.

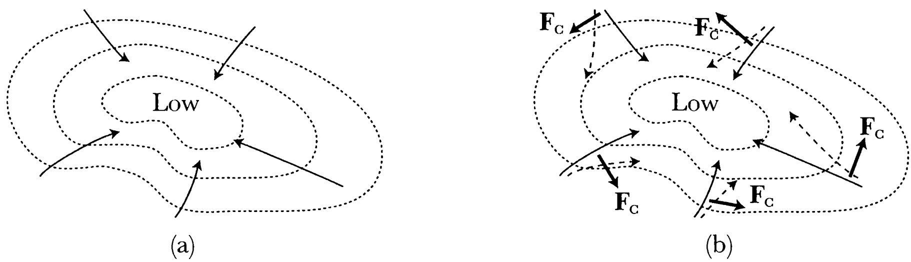


In this regard, it might be mentioned that although Foucault's pendulum experiment of 1851 was generally regarded as a more convincing, and certainly a more dramatic, demonstration of the earth's rotation than the experiments nearly a half-century earlier measuring the eastward deflection of falling bodies, there are many possible sources of error in constructing the pendulum, which actually requires great care. It was not until 1879 that these were thoroughly investigated, allowing the construction of a Foucault pendulum that did not move in an increasingly elliptical fashion. This was the subject of a physics doctoral thesis at Groningen, "Nieuwe bewijzen voor aswenteling der aarde" (New proofs of the rotation of the earth), by Heike Kamerlingh Onnes, later renowned as the discoverer of superconductivity. An exposition of this impressive work, which will also be mentioned in Chapter 22, can be found in Schulz-Dubois [1].

Because of the difficulties encountered trying to picture what is really happening with his pendulum, Foucault realized, perhaps in discussions with Poinsot, who shared his preference for geometric over analytic arguments, that the movement of the earth could be demonstrated much more directly by a gyroscope, whose axis of rotation would stay fixed. The instruments that had been constructed up to that time were inadequate for the purpose, because the friction and imprecision of the gimbal bearings created distortions totally masking the desired effect. Though Foucault managed to overcome these obstacles, a microscope was needed to see the deflection, because he could only get the gyroscope to keep spinning unaided for about ten minutes. Wonderful pictures of Foucault's gyroscope can be found in Tobin [1; pp. 163, 288].

Foucault also invented the gyrocompass, as described at the end of Chapter 9, although its eventual usefulness on ships required the contributions of many others, especially Hermann Anschütz-Kaempfe, culminating in the Anschütz-gyroscope. In addition, Foucault carried out many other important investigations, and even before his famous pendulum demonstration he had performed one of the first experiments showing that the speed of light was less in water than in air, which was extremely important in the debate about the nature of light, as we shall briefly discuss in Chapter 15. Much more about Foucault, and the scientific milieu of his time, can be found in the recent estimable biography Tobin [1] mentioned on the previous page; a discussion of Onnes' work may also be found there.

Hurricanes and bath-tubs. The Coriolis force is of interest to meteorologists because, among other things, it is the reason for the rotational winds of hurricanes, which form around centers of low air pressure as outer air flows toward the center, as in (a) of the figure below, where the dotted lines are "isobars", indicating areas of constant pressure, and the Coriolis force causes the inward paths to be deflected to the right, as in (b), causing the distinctive counterclockwise swirling



associated with hurricanes, with, as usual, many additional complications to the general picture.

The actual calculations, which we leave to the meteorologists, show that the very small Coriolis force can really have such an appreciable effect because the winds are moving with such high velocities, and for such long distances.

In theory, and also in fact, hurricanes in the southern hemisphere rotate clockwise rather than counterclockwise. Likewise, theoretically water drains out of a bath-tub in a counterclockwise direction in the northern hemisphere, but clockwise in the southern hemisphere. This can be verified with very careful experiments (Shapiro [1] for the northern hemisphere and Trefethen *et al.* [1] for the southern[1]), but in practice the direction depends on small effects, such as the slight motions that the water already has, which are of much greater magnitude than the Coriolis force could produce on water moving so slowly for such a short distance. Tourist scams demonstrating the reversal of the draining direction as ships pass the equator merely attest to the ease with which the direction can be subtly influenced.

On a larger scale, it might be mentioned (cf. Persson [1; pg. 24]) that at one time plans to build a space station rotating rapidly to create an artificial gravity had to be abandoned when it was realized that the rotation necessary would create Coriolis forces thousands of times stronger than on earth, with all sorts of dire consequences.

[1] And a multitude of others, cf. *Am. J. Physics*, **62** (1994), pg. 1063.

On a still larger scale, Addendum A discusses a much more complicated example where we get to consider the role of the Coriolis force in a feature of the solar system.

Mach's Principle. It is, alas, virtually impossible to conclude a discussion of rotating coordinate systems without mentioning Mach's Principle, which, like most philosophical principles, seems to be simultaneously both extremely significant for, and totally irrelevant to, everything physicists do.

Near the beginning of Chapter 7 (cf. page 275), we mentioned the philosophical problem of determining what one means by velocity and acceleration if one hasn't already decided on a coordinate system, leading Newton to resort to the idea of "absolute space"; as we suggested in that chapter, we hope to have avoided an appeal to "absolute space" by rephrasing the first law in terms implying the *existence* of an inertial system, without the expectation of a way to distinguish between all the different inertial systems.

Although Newton did not suggest any way to distinguish one inertial system from another, he pointed out that one could definitely distinguish rotating systems from non-rotating ones. He cited in particular the fact that water in a rotating bucket will assume a concave shape "(as experience has shown me)". Though not specifically mentioning a parabolic shape, he added that "The rise of the water reveals its endeavor to recede from the axis of motion, and from such an endeavor one can find out and measure the true and absolute circular motion of the water." A little later he also gave a more direct theoretical example: "if two balls, at a given distance from each other with a cord connecting them, were revolving about a common center of gravity, the endeavor of the balls to recede from the axis of motion could be known from the tension of the cord, and thus the quantity of circular motion could be computed. ... ", similar in some respects to the discussion on page 18. And of course, the bulging of the earth at the equator, the deflection of falling bodies, Foucault's experiment, etc., are all further examples, and the whole *raison d'être* for this chapter.

The notion that it is impossible to distinguish between uniform rectilinear motion and absolute rest has only been strengthened by later developments in physics, especially special relativity. On the other hand, rotation seemed to be quite a different story, and philosophers' rejection of the notion that one could determine absolute rotation seemed to be, not so much arguments that it wasn't true, as complaints that it simply couldn't, or shouldn't, be true.

This situation was changed considerably by the physicist Ernst Mach, who became interested in questions of perception and thence to a philosophy of science rather inimicable to reliance on theoretical concepts. Mach's analysis of the rotating bucket experiment was that it can only reveal that the water is rotating with respect to something else, and this something else must be the "fixed stars", or at any rate the mean position of these stars. As his ideas on this question developed, he seemed indeed to be putting forth the proposal, rather startling even to himself, that it was somehow the effect of all these stars in concert that caused the water to "endeavor to recede from the axis of motion", and that this wouldn't happen if they weren't there.

This notion, though embraced by some physicists, was dismissed with both horror and disdain by many others, and attempts have even been made to cite experimental results that might make it untenable. For example, the earth is located far out on a limb of our galaxy, so that the distribution of the nearby mass of the universe around the earth is quite unsymmetric, while numerous very delicate experiments have revealed no dependence of inertial mass on the direction of acceleration.

But Machians dismiss such observations as simply showing that it is not the influence of nearby matter, but of all the matter in the universe, that is responsible for these phenomena. Of course, we can't eliminate all the other matter in the universe to test this idea, which has thus been formulated so as to be unfalsifiable (which one would have thought ought to perturb the philosophers just a bit). And, in a modern rejoinder, Feynman [1; pg. 16-2] has pointed out mischievously, if not a bit petulantly, that if rotation is really being measured with respect to the fixed stars, then why can't rectilinear motion also be measured absolutely with respect to them?

The whole contretemps might have remained permanently in the shadowy demimonde of the philosophy of science, had Mach's views not attracted the attention of a bright young physicist, one Albert Einstein, as he was developing his theory of general relativity, very briefly mentioned at the end of Chapter 1.

Einstein initially felt that general relativity demonstrated the correctness of Mach's ideas, and there is a famous letter from Einstein to Mach (reproduced in Misner, Thorne, and Wheeler [1; §21.12]), written as Einstein was developing the theory of general relativity, in which he says that if the expected verification during an eclipse is made, then Mach's ideas "will receive brilliant confirmation" (Mach did not reply to this letter, which is not surprising, since he didn't even accept special relativity, or even the reality of atoms and molecules!). Einstein discussed these ideas in several papers, and in Einstein [2] he stated what he called *Mach's principle* "because this principle has the significance of a generalization of Mach's requirement that inertia should be derived from an interaction of bodies." Nevertheless, it is generally acknowledged that general relativity and Machian ideas have never actually been unified.

The discussion in §21.12 of Misner, Thorne and Wheeler [1], written shortly after Feynman's book, practically declares Mach's ideas to be correct if properly understood. But contentious discussions of Mach's ideas have continued, and may well continue forever. For those who just can't get enough of this sort of thing, a good place to begin further reading might be

> *Mach's Principle: From Newton's Bucket to Quantum Gravity*, Barbour, J. and Pfister, H. eds., Birkhauser, 1995.

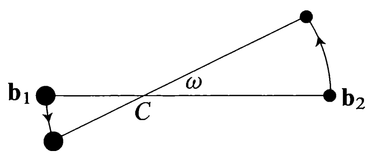
This volume is based on a conference held in 1993, perhaps the first ever devoted exclusively to Mach's Principle. Being fairly recent, it may be close to the latest word on Mach's Principle (though, given the philosophical nature of the question, undoubtedly not the *last!*).

ADDENDUM 10A

THE TROJAN ASTEROIDS

In Chapter 4 we easily progressed from the one-body problem to the two-body problem, and it might be thought that the three-body problem would require just a bit more effort. As the old adage says, however, two's company, three's a crowd. The three-body problem exhibits almost all of the intractability of the more general N-body problem. This circumstance led to the investigation of many special cases of the three-body problem where one could at least get some results, even if these results seemed unlikely to have any application.

The restricted three-body problem. In particular, consider the case where one of the bodies has a negligible mass m compared to the other two, $\mathbf{b}_1$ and $\mathbf{b}_2$, with masses $m_1 > m_2$; it is usually assumed also that $\mathbf{b}_1$ and $\mathbf{b}_2$ move in a plane, and are actually in circular orbits about their center of mass C.



Problem 4-10 shows that $\mathbf{b}_1$ and $\mathbf{b}_2$ rotate about C with an angular velocity ω given by

$$\omega^2 a^3 = G(m_1 + m_2),$$

where a is the distance between $\mathbf{b}_1$ and $\mathbf{b}_2$. One could also deduce this from the equations (page 121) for circular motion: if $d_i = |\mathbf{b}_i - C|$, then

$$\frac{Gm_1m_2}{a^2} = m_2 d_2 \omega^2 = m_1 d_1 \omega^2.$$

Since the force on the third body depends on its distances from $\mathbf{b}_1$ and $\mathbf{b}_2$, the easiest coordinate system to work with is the rotating one in which they are fixed. We choose the line from $\mathbf{b}_1(t)$ to $\mathbf{b}_2(t)$ as the rotating x-axis, with C at $(0,0,0)$, so that $\mathbf{b}_1(t)$ has coordinates $(-d_1, 0, 0)$ and $\mathbf{b}_2(t)$ has coordinates $(d_2, 0, 0)$, where

$$d_1 = a \cdot \frac{m_2}{m_1 + m_2}, \qquad d_2 = a \cdot \frac{m_1}{m_1 + m_2}.$$

Then $\boldsymbol{\omega} = (0, 0, \omega)$ and for the third particle $c(t) = (c_1(t), c_2(t), c_3(t))$, we have

$$\text{centrifugal force} = +m \cdot \omega^2 (c_1(t), c_2(t), 0)$$

$$\text{Coriolis force} = +2m \cdot \omega (c_2{}'(t), -c_1{}'(t), 0),$$

so if we let $r_i(t) = |c(t) - \mathbf{b}_i|$, we have

$$(*) \quad (c_1''(t), c_2''(t), c_3''(t)) = -G\left[m_1 \frac{c(t)-\mathbf{b}_1}{r_1{}^3} + m_2 \frac{c(t)-\mathbf{b}_2}{r_2{}^3}\right]$$
$$+\, \omega^2 \cdot (c_1(t), c_2(t), 0) + 2\omega \cdot (c_2'(t), -c_1'(t), 0).$$

We are not going to try to solve even this simplified equation in general. We will merely look for the special case of a stationary orbit, in other words we seek a solution with $c_i(t) = \mathring{c}_i$ some constants $\mathring{c}_i$, which means that the orbit is stationary *in our rotating coordinate system.* Since we then have $0 = c'(t) = c''(t)$, the components of equation $(*)$ give

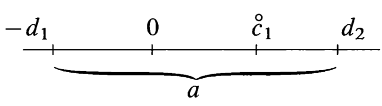
$$\text{(a)} \qquad 0 = -G\left[\frac{m_1}{r_1{}^3}(\mathring{c}_1 + d_1) + \frac{m_2}{r_2{}^3}(\mathring{c}_1 - d_2)\right] + \omega^2 \mathring{c}_1$$

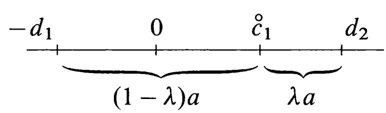
$$\text{(b)} \qquad 0 = \left(\omega^2 - G\left[\frac{m_1}{r_1{}^3} + \frac{m_2}{r_2{}^3}\right]\right) \cdot \mathring{c}_2$$

$$\text{(c)} \qquad 0 = -G\left[\frac{m_1}{r_1{}^3} + \frac{m_2}{r_2{}^3}\right] \cdot \mathring{c}_3.$$

Equation (c) immediately gives $\mathring{c}_3 = 0$, so the stationary orbits must be in the plane in which $\mathbf{b}_1$ and $\mathbf{b}_2$ revolve.

Equation (b) will automatically be satisfied if $\mathring{c}_2 = 0$, i.e., for a stationary orbit that lies along the line between $\mathbf{b}_1$ and $\mathbf{b}_2$; for such "collinear" solutions, the only condition is that given by (a). Suppose first that $\mathring{c}_1$ is between $-d_1$ and d_2, where $d_1 + d_2 = a$, and choose λ with $r_2 = d_2 - \mathring{c}_1 = \lambda a$, and thus $r_1 = (1-\lambda)a$.





Since

$$\mathring{c}_1 = d_2 - \lambda a = \left(\frac{m_1}{m_1 + m_2} - \lambda\right) a,$$

equation (a) becomes

$$0 = -G\left[\frac{m_1}{(1-\lambda)^3 a^3}(1-\lambda)a + \frac{m_2}{\lambda^3 a^3}\lambda a\right] + \omega^2\left(\frac{m_1}{m_1+m_2} - \lambda\right) a,$$

and since

$$\frac{G}{\omega^2} = \frac{a^3}{m_1 + m_2},$$

this reduces to

$$0 = f(\lambda) = m_1 - (m_1 + m_2)\lambda - \frac{m_1}{(1-\lambda)^2} + \frac{m_2}{\lambda^2}.$$

We have $f'(\lambda) < 0$ for $0 < \lambda < 1$ and $f(\lambda) \to \infty$ as $\lambda \to 0$ and $f(\lambda) \to -\infty$ as $\lambda \to 1$, so there is a unique λ in $(0, 1)$ satisfying this equation, and thus a unique $\mathring{c}_1$ between $-d_1$ and d_2. Similar arguments show that there is also a unique $\mathring{c}_1 < -d_1$ and a unique $\mathring{c}_1 > d_2$, so there are exactly 3 stationary orbits on the line between c_1 and c_2, discovered by Euler in 1767.

Any other stationary orbits, "triangular" orbits, must satisfy

$$\text{(b}')\qquad \omega^2 - G\left[\frac{m_1}{r_1{}^3} + \frac{m_2}{r_2{}^3}\right] = 0,$$

while (a) can be rewritten in the form

$$\text{(a}')\qquad \left(\omega^2 - G\left[\frac{m_1}{r_1{}^3} + \frac{m_2}{r_2{}^3}\right]\right)\mathring{c}_1 - G\frac{m_1 m_2}{m_1 + m_2}a\left[\frac{1}{r_1{}^3} - \frac{1}{r_2{}^3}\right] = 0,$$

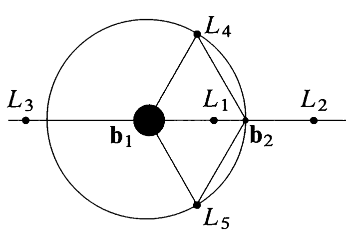
and (b$'$) then implies that

$$r_1 = r_2.$$

Thus (b$'$) simply becomes

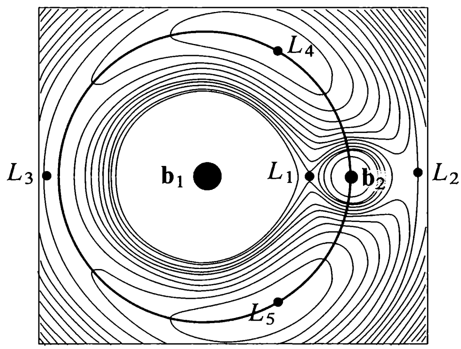
$$\omega^2 = \frac{G(m_1 + m_2)}{r_1{}^3},$$

and using the relation at the top of the page, this shows that $r_1 = r_2 = a$. Thus $\mathbf{b}_1, \mathbf{b}_2, c$ form an equilateral triangle. These triangular orbits were discovered by Lagrange in 1733. When $m_1 \gg m_2$ and the orbit of $\mathbf{b}_2$ is essentially a circle around $\mathbf{b}_1$, the orbit of c is on that same circular orbit. It is customary to label the one preceding $\mathbf{b}_2$ in its orbit as L_4 and the one following as L_5, with Euler's collinear stationary orbits numbered L_1, L_2, and L_3 (though there is no fixed convention for the numbering of these three). $L_1, \dots, L_5$ are called "libration" points by astronomers, although all five are often called Lagrange points.



Stability. Even allowing for the various idealizations in our analysis, like the assumption that c has negligible mass, the foregoing results require further consideration. Although we have found five mathematically stationary orbits, only stable ones are of interest; the tiniest asymmetry of the body at the orbit, or even a passing bit of cosmic dust, will change the orbit by a tiny amount, and we need to know that this will not cause the body to start wandering far away.

To get some idea about the stability of the stationary orbits, consider the contour plot below for the potential function V associated with the vector field

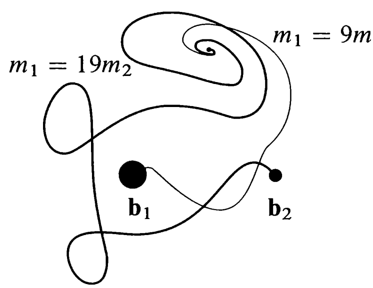


indicated by the first two terms on the right of $(*)$; we have to omit the Coriolis force, which is a function of velocity, not position, but at least we know that this is 0 for stationary orbits. The potential V approaches $-\infty$ around $\mathbf{b}_1$ and $\mathbf{b}_2$, while the Lagrange points, where the combined gravitational forces of $\mathbf{b}_1$ and $\mathbf{b}_2$ just cancel the centrifugal force, are critical points for V. If a body at a collinear Lagrange point is displaced horizontally, it will move into a region where V has a smaller value, and thus continue to move away from its position, so we would expect these Lagrange points to be unstable. Of course, we've ignored the Coriolis force, which comes into play as soon as the particle near the Lagrange point starts to move, so this argument needs to be verified by calculations. As we will mention later, the calculations show that these points really are unstable, implying that the Coriolis force doesn't materially change the situation.

These calculations will also show that in the case of the triangular Lagrange points L_4 and L_5, the Coriolis force does make a decisive difference: as c moves away from L_4 or L_5, "rolling down the potential hill", the Coriolis force deflects it, causing it to spiral, and when $m_1 \gg m_2$ the spiraling is sufficiently fast to insure that it stays in the vicinity of L_4 or L_5. Of course, one need not invoke the Coriolis force for this explanation; it can be rephrased directly in terms of the rotation of the Lagrange point around the center of mass of $\mathbf{b}_1$ and $\mathbf{b}_2$.

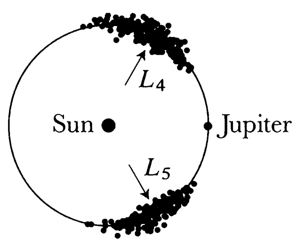
Though this may seem like a lot of analysis for a special instance of the three-body problem that was concocted simply to produce a solvable solution, that all changed in 1906, when the intrepid asteroid hunter Max Wolf found his 588[th] asteroid, 588 Achilles, at L_4 for the Sun-Jupiter system. The asteroid Patroclus was then found at L_5, and Hektor at L_4. After that, the next asteroids Nestor (L_4), Priamus (L_5), Agamemnon (L_4), Odysseus (L_4), Aneas (L_5), ... , were always given the names of Greeks in the Trojan war if they were near L_4 and the names of Trojans if near L_5, so that only Patroclus and Hektor are in the wrong camps, and the combined groups were called the Trojan asteroids.

The web site `www.kw.igs.net/~jackord/bp/f8.html` shows plots of the path of an asteroid starting very close to L_4 for various ratios m_1/m_2. The figure below shows the asteroid falling into the heavier body $\mathbf{b}_1$ when the ratio



is 9, and into $\mathbf{b}_2$ when the ratio is 19. By the time we get to a ratio of 99, which includes the Sun-Jupiter system, we have stability, with the path staying very close to L_4.

Periodic orbits near the Lagrange points have also been studied (see page 404 for a reference). In the Sun-Jupiter case, there are "stable orbits", closed curves that rotate about the sun in sync with Jupiter, one of which is also shown in the above web site. In fact there are actually many asteroids of this sort near L_4 and L_5, which continue to be discovered at an alarming rate, with over 1,000 at each point, and Greek and Trojan names have been exhausted long ago. The rough picture below is based on a diagram that can be found in the web site `www.dtm.ciw.edu/sheppard/satellites/trojan.html`, which has a continually updated count of the known Trojans.



It also turns out that there are "Mars Trojans" at both the L_4 and L_5 points of the Sun-Mars system, as well as Neptune Trojans. In addition, although there are no known Earth Trojans, there are some very strange asteroid companions. The web site `math.ucr.edu/home/baez/lagrange.html`, which covers many topics about Trojans, give references to other sites where these strange asteroid companions of earth are discussed, together with animated pictures of their orbits.

Stability calculations. In Addendum 6C we described stability criteria for first order equations. A second order equation

$$(c_1, c_2, c_3)''(t) = F(c_1(t), c_2(t), c_3(t), c_1{}'(t), c_2{}'(t), c_3{}'(t))$$

is handled in the usual way, by introducing new variables v_1, v_2, v_3 and writing the equation as

$$(c_1, c_2, c_3, v_1, v_2, v_3)' = (v_1, v_2, v_3, F(c_1, c_2, c_3, v_2, v_2, v_3)).$$

Applying this to our second order equation

$$(*) \quad (c_1{}''(t), c_2{}''(t), c_3{}''(t)) = -G\left[m_1\frac{c(t)-\mathbf{b}_1}{r_1{}^3} + m_2\frac{c(t)-\mathbf{b}_2}{r_2{}^3}\right]$$
$$+\,\omega^2\cdot(c_1(t), c_2(t), 0) + 2\omega\cdot(c_2{}'(t), -c_1{}'(t), 0),$$

we will now be considering a 6×6 Jacobian matrix made up of a 3×3 matrix with all entries 0, a 3×3 identity matrix, and two other matrices J_1 and J_2, the second of which, coming from the term $2\omega\cdot(c_2{}'(t), -c_1{}'(t), 0)$, is simply

$$J_2 = -2\omega\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

As for J_1, since

$$\frac{\partial r_1}{\partial x} = \frac{\partial\sqrt{(x+d_1)^2+y^2+z^2}}{\partial x} = \frac{x+d_1}{r_1},$$

the derivatives with respect to x of the components of

$$\frac{(x,y,z)-\mathbf{b}_1}{r_1{}^3} = \frac{(x+d_1,y,z)}{r_1{}^3}$$

are

$$\frac{1}{r_1{}^3} - \frac{3(x+d_1)^2}{r_1{}^5}, \qquad \frac{-3(x+d_1)y}{r_1{}^5}, \qquad \frac{-3(x+d_1)z}{r_1{}^5}$$

and similarly for the other partials. Remembering the $\omega^2 \cdot (c_1(t), c_2(t), 0)$ term, the complete expression for $J_1(x, y, z)$ will be

$$\omega^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} - G\frac{m_1}{r_1{}^3}\left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{3}{r_1{}^2}(x+d_1, y, z)^{\mathbf{t}} \cdot (x+d_1, y, z)\right]$$

$$- G\frac{m_2}{r_2{}^3}\left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{3}{r_2{}^2}(x-d_2, y, z)^{\mathbf{t}} \cdot (x-d_2, y, z)\right],$$

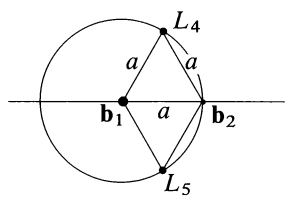
here written using terms like the product of $(x + d_1, y, z)$ with its transpose on the left (also called the outer product of $(x + d_1, y, z)$ with itself): explicitly, the whole first bracketed term is

$$\begin{pmatrix} 1 - \dfrac{3(x+d_1)^2}{r_1{}^2} & -\dfrac{3y^2}{r_1{}^2} & -\dfrac{3z^2}{r_1{}^2} \\ -\dfrac{3(x+d_1)y}{r_1{}^2} & 1 - \dfrac{3y^2}{r_1{}^2} & -\dfrac{3zy}{r_1{}^2} \\ -\dfrac{3(x+d_1)z}{r_1{}^2} & -\dfrac{3yz}{r_1{}^2} & 1 - \dfrac{3z^2}{r_1{}^2} \end{pmatrix},$$

with a similar expression for the other term.

When our point (x, y, z) is L_4 or L_5 we have

$$(x + d_1, y, z) = \frac{a}{2}(1, \pm\sqrt{3}, 0)$$
$$(x - d_2, y, z) = \frac{a}{2}(-1, \pm\sqrt{3}, 0),$$



where $+\sqrt{3}$ applies for L_4 and $-\sqrt{3}$ for L_5. Remembering that $r_1 = r_2 = a$ and that $\omega^3 a^3 = G(m_1 + m_2)$, our first matrix comes down to

$$J_1(x, y, z) = \frac{3\omega^2}{4}\begin{pmatrix} 1 & \pm\sqrt{3}\,\mu & 0 \\ \pm\sqrt{3}\,\mu & 3 & 0 \\ 0 & 0 & -\frac{4}{3} \end{pmatrix}, \qquad \mu = \frac{m_1 - m_2}{m_1 + m_2}.$$

Taking the corresponding approximating first order equations in 6 variables, and putting them back into the form of a second order approximating equation for $\mathbf{c} = (c_1, c_2, c_3)^{\mathbf{t}}$, we get

$$\mathbf{c}'' - \frac{3\omega^2}{4}\begin{pmatrix} 1 & \pm\sqrt{3}\,\mu & 0 \\ \pm\sqrt{3}\,\mu & 3 & 0 \\ 0 & 0 & -\frac{4}{3} \end{pmatrix}\mathbf{c} + 2\omega\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\mathbf{c}' = 0,$$

which separates into a system for the first two components and the single equation

$$c_3{}'' + \omega^2 c_3 = 0.$$

The latter is simply the harmonic oscillator, with solutions that are periodic functions, indicating that small changes in the z direction, off of the plane containing $\mathbf{b}_1$ and $\mathbf{b}_2$, will not lead to instability.

Ignoring this third component, and considering $\mathbf{c} = (c_1, c_2)^{\mathbf{t}}$, we then have the system

$$\mathbf{c}'' + 2\omega\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\mathbf{c}' - \frac{3\omega^2}{4}\begin{pmatrix} 1 & \pm\sqrt{3}\,\mu \\ \pm\sqrt{3}\,\mu & 3 \end{pmatrix}\mathbf{c} = 0,$$

for which we seek a solution of the form

$$\mathbf{c}(t) = e^{\alpha t}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

This leads to the equations

$$\begin{pmatrix} \alpha^2 - \frac{3}{4}\omega^2 & -\omega\left(2\alpha \pm \frac{3\sqrt{3}}{4}\omega\mu\right) \\ \omega\left(2\alpha \mp \frac{3\sqrt{3}}{4}\omega\mu\right) & \alpha^2 - \frac{9}{4}\omega^2 \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0,$$

which have non-trivial solutions only if the determinant of the matrix is zero. Hence we need

(a) $$\alpha^4 + \omega^2\alpha^2 + \tfrac{27}{16}(1 - \mu^2) = 0,$$

or

$$\alpha^2 = -\frac{\omega^2}{2}\left(1 \pm \sqrt{1 - \tfrac{27}{4}(1-\mu)^2}\right).$$

Since any solution α will also give the solution $-\alpha$, if we have any solution that is not purely imaginary, then we will definitely have one with a real positive part, so our stationary orbit will definitely not be stable. So we obviously need

$$0 \leq 1 - \tfrac{27}{4}(1 - \mu^2) < 1.$$

Since $m_1 \geq m_2$ implies that $0 \leq \mu \leq 1$, to have only purely imaginary solutions we need

$$\mu^2 \geq \tfrac{23}{27} \implies \frac{m_1 - m_2}{m_1 + m_2} \geq .9229582.$$

This holds for the Sun and a planet. For example, the mass of the Sun is 1047.35 times the mass of Jupiter, so for the Sun and Jupiter we have

$$\mu = \frac{1 - \frac{1}{1047.35}}{1 + \frac{1}{1047.35}} = .9980922.$$

Of course, we have simply found the condition that the characteristic equation of our matrix has purely imaginary roots, and this is the critical case in which we can't determine stability for sure. Proving stability for this particular case is an extremely difficult mathematical problem, which was only solved (nearly completely) in the 1960's, though physicists presumably weren't particularly concerned with these mathematical niceties. For references, see Abraham and Marsden [1; end of §10.2]; also see §10.3 of that book for a discussion of closed orbits around the Lagrange points.

The collinear Lagrange points. Although we will not carry out the calculations, it turns out that the collinear Lagrange points are all unstable (Abraham and Marsden [1], Beutler [1], and Boccaletti and Pucacco [1]).

In the case of the Sun-Earth system, although the collinear Lagrange points are unstable, L_1 and L_2 are promising positions for artificial satellites, which can be kept in stable orbit by rockets, which they have to have anyway. In fact, SOHO, the Solar and Heliospheric Observatory, is at L_1, with an unobstructed view of the Sun, and WMAP, the Wilkinson Microwave Anisotropy Probe is at L_2.

As for L_3, always on the other side of the Sun from Earth, it has starred in various science fiction stories and movies as the hidden "Planet-X", ideally situated to attack Earth. With an instability on a time scale of 150 years, it isn't a very good candidate for a planet, though it might make a good temporary launching field for an attack on Earth.

ADDENDUM 10B

THE SOUTHWARD DEFLECTION

In the words of an accomplished experimental physicist, Hall [1], the question of a southward deflection of a falling body "has been answered in the negative, on theoretical grounds, by Gauss and by Laplace, and in the positive, on experimental grounds, by nearly every one of the investigators who have from time to time through more than two centuries made the actual trial."

In fact, the calculations by Laplace and Gauss were made after Benzenberg's first experiments in 1802, dropping balls from the tower of St. Michael's in Hamburg, and in these experiments Benzenberg had found both an eastward deflection and a southward deflection. For the later experiments of 1804, in a mine shaft in Schlebusch, Benzenberg dropped 40 balls, and then selected the drops that he considered to have been made under the most favorable conditions. The results then gave an eastward deflection close to the 8.8 mm predicted by Laplace's and Gauss's analyses, while the southward deflection now seemed to have become negligible.

Hall concludes, "The honesty of Benzenberg is not to be questioned. But we may well ask how far he may have been influenced, in 'selecting' his evidence from the whole body of data, by the knowledge that the authority of Gauss and of Laplace was dead against the southerly direction." Indeed, Hall allows that such biases might be what a careful investigation of this strange discrepancy between theory and experiment would ultimately reveal: "If the whole mystery is the consequence of mental bias in the experimenters, the proof and explanation of this bias would have, at least, the merit of psychological interest."

The second paper by Hall in the same journal is devoted to Hall's own experiments. They too detected a southward deflection, but it was within the limits of experimental error for his particular setup, which had the advantage of being carried out in a specially constructed enclosed tower, but the disadvantage of involving a much shorter fall.

Among the various possible explanations given for an actual southward deflection, the influence of the bulging of the earth near the equator has seemed the most promising, but there are apparently still disputes about the matter. See French [1] for references.

Hall's paper mentions the many experimental difficulties involved in experiments of this sort and considering how many there are, it's something of a wonder that reasonable results have ever been obtained. Aside from obvious problems like air resistance, one of the most delicate points is assuring that the objects are being dropped without having any sidewards motion imparted. Near the end of his paper, Hall mentions one other factor—the reliability of plumb bobs—that I have not seen considered in any other investigations of this question. In particular, our calculation of the deflection of the hanging pendulum bob on page 380 essentially assumes that it hangs on a weightless filament, but in practice there will be considerable weight distributed along the supporting wire, so the vector **g** in equation (1′), which plays a crucial role in our analysis of the equations for a falling body, won't really be indicated accurately by an actual plumb line.

I do not know of any calculations trying to determine the difference that this makes, but it could conceivably explain the small southern deflections obtained, as arising from making measurements from an incorrect base point.

PROBLEMS

1.[1] Let R be the radius of the earth (thought of as a perfect sphere) and h the height from which an object falls at the equator (at the end, all equations can be multiplied by $\cos\lambda$ for the result at latitude λ). For convenience we also assume the object has mass $m = 1$. Let ω be the angular velocity of the earth.

(a) At the time of release, the speed of the object is $v = \omega(R+h)$, and the angular momentum has magnitude

$$v(R+h) = \omega(R+h)^2.$$

Letting $\omega(x)$ be the angular velocity of the object after it has fallen a vertical distance $x = \frac{1}{2}gt^2$ [so that $\omega = \omega(0)$], show that

$$\omega(R+h)^2 = \omega(x)(R+h-x)^2,$$

so that

$$\omega(x) = \frac{\omega(R+h)^2}{(R+h-x)^2}.$$

(b) For $h \ll R$ we have approximately

$$\omega(x) = \omega(1 + 2x/R) = \omega(1 + gt^2/R).$$

(c) If the time of descent is T, then the total angular displacement is

$$\omega T + \frac{\omega g T^3}{3R},$$

and the displacement measured along the earth's surface is R times this.
(d) Noting that the foot of the tower has an angular displacement of ωT, conclude that the displacement from the foot of the tower is

$$\frac{1}{3}\omega g T^3.$$

2. (a) The easiest way to handle the equations on page 384 is simply to ignore the quantities x' and y', since these are surely going to be very small, obtaining

$$x'' = -2\omega z' \cos\lambda$$
$$z'' = -g.$$

Using the initial conditions, obtain $z' = -gt$, substitute into the first equation, and obtain immediately

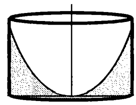
$$x(t) = \frac{1}{3}\omega g t^3 \cos\lambda.$$

(b) Treat the cannon problem in the same way.

---

[1] From Reddingius [1].

3. Show that the volume of a segment of a paraboloid of revolution is half the volume of the circumscribing cylinder, of radius a, say. So if a liquid in a



rotating cylinder rises to a height h above its initial height, the depth below the initial height is also h. Conclude that the angular velocity ω of rotation satisfies

$$\omega^2 = \frac{4gh}{a^2}.$$

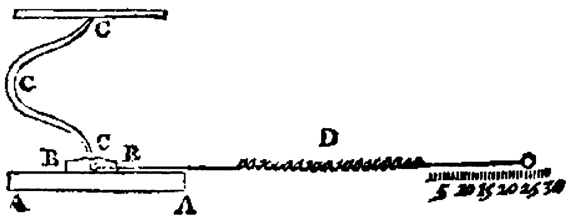
If the wheels of a vehicle are attached by gears, at a known ratio, to an upright cylinder filled with liquid inside the vehicle, this allows us to use the cylinder as a speedometer. Such a mechanism, with oil as the liquid, was actually once used for trains.

CHAPTER 11

FRICTION, FRIEND AND FOE

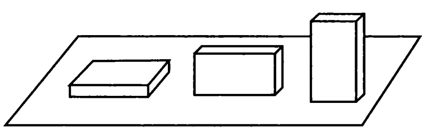
Friction is generally regarded as an unwelcome intruder in theoretical mechanics, where it is usually banished by fiat in elementary problems, or, as in the discussion of rolling in Chapter 6, cleverly given a figurehead role without any actual influence. To be sure, the very first problem in this book acknowledged its importance, and a footnote in Chapter 9 implied that it might behoove us to consider it in more detail.

The laws of friction. Tribology, the study of friction (from Greek $\tau\rho\iota\beta\omega$ = rub), nowadays a specialty, was originally just a side-line for researchers. The earliest such investigations, aside from those of Leonardo da Vinci, unearthed only later, were made by Guillaume Amontons (1663–1705), mainly noted for his improvements to thermometers, barometers, and hygrometers. Amontons' first



law of friction states that the frictional force of one object sliding on another is proportional to the applied load. Though da Vinci measured the friction of a block sliding on an inclined plane, the design of Amontons' mechanism suggests that he did not. In any case, nowadays we are careful to say that friction is proportional to the normal (component of the) force, as mentioned in Chapter 6.

The usual, somewhat mystifying, statement of Amontons' second law, that the force depends only on the load, not on the amount of surface area where



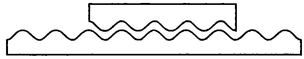
the objects are in contact, refers to a single body; for example, the block in the above figure will have the same friction with the surface on which it slides in all three positions. It appears that this conclusion, which had also been made by da Vinci, was always considered very suspect, since friction seemed to be a surface phenomenon.

In this regard, mention is sometimes made of J. T. Desaguliers (1683-1744), an experimenter in many fields, who in 1725 demonstrated a force of adhesion between two balls of lead (what we would now call molecular adhesion), and suggested, rather far-sightedly, that this might have something to do with friction, though this proposal had to contend with the fact that the force of adhesion is not independent of the contact area, but proportional to it.

The prolific Euler (1707–1783) performed his own analysis and experiments on friction, and noted that for a block of steel on an inclined steel plane, when the inclination reaches a critical angle the block starts moving, but only very slowly, while in the case of wood on wood the block starts moving relatively fast. This lead him to distinguish between "kinetic friction", the frictional force when the object is moving, and "static friction", the frictional force opposing motion when an applied force doesn't cause motion, with kinetic friction less than or equal to static friction, and he is generally credited with being the first to make this distinction.

Coulomb (1736–1806) is the person always sure to be mentioned in connection with friction, though he is much more famous for his law of electrostatic forces. He became interested in friction through Amontons' work, and carefully verified that friction was independent of the contact area. Coulomb not only distinguished between static friction and kinetic friction, but also noted that static friction often increased when the block and the surface remained in contact for a long time, even providing an empirical formula for the rate of increase. He stated several rather detailed rules (cf. Meyer [1; pg. 9]) and nowadays special attention is given to "sliding friction" satisfying Coulomb's law that it is independent of speed, though Coulomb's rule was not quite so definitive. Coulomb's experiments were later extended in scope by Morin (1795–1880), who sometimes appears as the third musketeer of French friction researchers.

Coulomb attributed friction to the roughness of the surfaces, and the effort needed to slide the protruding humps over each other, which would help explain



Amontons' second law, since both the surface area and the pressure would be involved. Nowadays, this explanation is cavalierly dismissed on the grounds that the work expended moving up the humps is retrieved as the block slides back down under the normal force. Indeed, this criticism would seem reasonable for a regular array of humps as in the figure, but might not seem so applicable to a more random distribution of humps of various sizes. In any case, the explanation seemed to conflict with some of Morin's experiments, and was later seen to conflict with other evidence. For example, it turned out that friction between surfaces was often lower when one was significantly *rougher* than the other, not to mention the fact that highly polished surfaces might exhibit *increased* friction, as mentioned on page 216, which fit in nicely with the ideas of Desaguliers.

A resolution seemed imminent around 1950, when the contact area was examined more carefully. Because of microscopic irregularities, the actual contact area of two surfaces is much smaller than the apparent macroscopic area, and, most importantly, an increase in the normal load pushes these irregularities closer together, so that they overlap more, and even flattens some, thus increasing the contact area. This lent support to the idea that friction does result primarily from molecular adhesion, and one can actually observe tiny fragments of the surfaces being worn away because of this adhesion force.

The picture was not destined to remain so simple, however, since it was later shown that in some cases, as with the ultra-smooth surfaces of a cleaved piece of mica, there is friction even though there is no wear whatsoever, and the modern picture of friction invokes waves in the atomic lattice generated by the protrusions being deformed, described in some detail in Krim [1], and briefly in Feynman [1; pg. 12-4].

This entire discussion involves "sliding friction", and there are whole other areas that we haven't even begun to consider, like rolling friction, which occurs when real, rather than idealized, bodies roll; viscous friction, which is not independent of velocity, but proportional to it (cf. page 295); the use of lubricants; etc., and we will studiously continue to ignore them.

So where does this all leave us? Basically we will be considering only the simplified laws of Coulomb friction, sometimes called Amontons-Coulomb friction, and even sometimes Coulomb-Morin friction:

Consider a fixed planar surface, and a body in contact with it, with the total force on this body decomposed as $\mathbf{N} + \mathbf{F}$, where $\mathbf{N}$ is normal to the surface, with magnitude N, and $\mathbf{F}$ parallel to it, with magnitude F.

Then there is a critical value μ_S, the *static friction*, and a number μ with $0 \leq \mu \leq \mu_S$, the *coefficient of (kinetic) friction*, so that

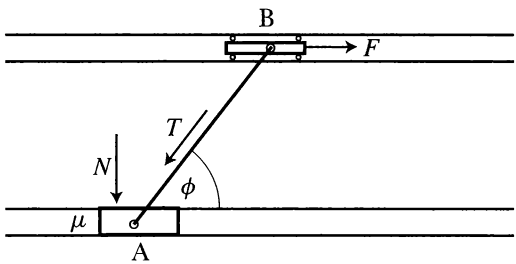
> for $F \leq \mu_S \cdot N$, there is no motion,
>
> for $F > \mu_S \cdot N$ there is motion in the direction of $\mathbf{F}$, and the body acts as if it is under the influence of a force of magnitude $F - \mu N$, i.e., there is now an additional "frictional force" of magnitude μN.

Note that the stipulation $\mu \geq 0$ implies that the frictional force is always in the opposite direction from $\mathbf{F}$. And since μ is simply a number, it is implied that this frictional force is independent of the speed of motion.

Of course, this is a purely empirical law, and like almost every empirical law, it is hedged in on all sides with restrictions and exceptions. Nevertheless, the very basic idea, that friction is proportional to the normal load, or sometimes merely the fact that friction exists, and, importantly, is always in the opposite direction of motion, helps explain, or is even crucial to explain, certain phenomena. Before discussing some complex examples (a couple of simple examples are given in the Problems), we ought to examine the problems that the Amontons-Coulomb laws of friction present on purely logical grounds.

The Painlevé paradoxes. In machinery with parts moving against each other, it often happens, even when the designers have taken friction into account to estimate the amount of force that needs to be applied, that the machinery may stop sliding smoothly, moving in starts and stops, or vibrating, or even locking up in certain positions. In a more mundane example, those of us of a certain age, who once wrote with chalk on blackboards, know that when a hard piece of chalk is pushed on the blackboard at an angle nearly perpendicular to it, the chalk will sometimes start skipping erratically instead of sliding, resulting in a dashed or dotted line. One might imagine various reasons for these effects, but the work of Painlevé [2], best known for his mathematical work, first brought attention to the fact that they are to be expected because of some inconsistencies between the basic laws of mechanics and the laws of Amontons-Coulomb friction.[1]

Modern presentations of the Painlevé paradoxes usually begin with a simple example called the Painlevé-Klein problem. Consider two carts moving along



parallel guides, connected by a rigid rod, with ϕ the angle determined by the length of the rod and the distance between the guides. Cart A is dragged along at a constant distance behind cart B, which is supposed to slide frictionlessly, as indicated by the little wheels or ball bearings, while cart A slides with coefficient of friction μ. For the simplest case we assume that the two carts each have mass

[1] This work was somewhat anticipated by Jellett [1], who will reappear in a later section.

$m = 1$, while the connecting rod has negligible mass, and that a force in the direction of the guides, with magnitude F, is applied at B. The normal force of magnitude N on A now arises from the tension force of magnitude T along the rod, with

$$N = T \sin\phi.$$

Note that N and T may be positive or negative. As usual, we let sign a be $+1$ for $a > 0$, and -1 for $a < 0$, and 0 for $a = 0$.

The total horizontal force acting on B is $F - T\cos\phi$, so if x is the coordinate of B we have

$$\text{(a)} \qquad x'' = F - T\cos\phi = F - \frac{N}{\tan\phi}.$$

On the other hand, the whole system, consisting of the two connected carts, satisfies

$$\text{(b)} \qquad \begin{aligned} 2x'' &= F - \mu|N| \operatorname{sign} x' \\ &= F - \mu N \operatorname{sign} N \operatorname{sign} x', \end{aligned}$$

and equations (a) and (b) yield

$$N = \frac{F\tan\phi}{2 - \mu\tan\phi \operatorname{sign} N \operatorname{sign} x'}.$$

Now suppose that $\boldsymbol{\mu \tan \phi > 2}$.

If we seek a solution with $x' > 0$, we immediately obtain a contradiction for either value of sign N.

For $x' < 0$ the solution is not unique, for we have both

$$N = \frac{F\tan\phi}{2 + \mu\tan\phi} > 0 \quad \text{and} \quad N = \frac{F\tan\phi}{2 - \mu\tan\phi} < 0.$$

If instead we have $\boldsymbol{\mu \tan \phi < 2}$, then we obtain the unique solution

$$N = \frac{F\tan\phi}{2 - \mu\tan\phi \operatorname{sign} x'} > 0, \qquad x'' = \frac{F(1 - \mu\tan\phi \operatorname{sign} x')}{2 - \mu\tan\phi \operatorname{sign} x'}.$$

As long as F is bounded, N will stay bounded also, so if the system at least starts in a reasonable state it can never reach a singular position where $\mu\tan\phi = 2$. Nevertheless, not everything is honky-dory. The most interesting property of the system involves its motion starting at rest, $x' = 0$. The frictional force on A has magnitude

$$\mu N \operatorname{sign} x' = \frac{\mu F \tan\phi \operatorname{sign} x'}{2 - \mu \tan\phi \operatorname{sign} x'},$$

while the horizontal force on A due to the force applied at B is

$$T \cos\phi = \frac{N}{\tan\phi} = \frac{F}{2 - \mu \tan\phi \operatorname{sign} x'}.$$

In order for A to start moving, the second must be greater than the first,

$$\frac{F}{2 - \mu \tan\phi \operatorname{sign} x'} > \frac{\mu F \tan\phi \operatorname{sign} x'}{2 - \mu \tan\phi \operatorname{sign} x'}.$$

When $F > 0$, so that B is being pushed to the right, this implies

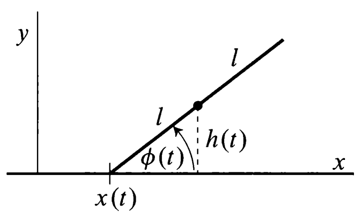
$$\mu \tan\phi \operatorname{sign} x' < 1.$$

In order for B to move right, with $\operatorname{sign} x' = 1$, we thus must have

$$\mu \tan\phi < 1.$$

Hence, if there is any friction at all, $\mu > 0$, then for a large enough initial angle ϕ it will be impossible to start, and the system is "wedged", or self-braking.

The standard Painlevé example, which can serve as a model of the chalk phenomenon, involves a uniform rod of mass 1 and length $2l$, which is sliding on a horizontal surface, with a coefficient of friction μ. We will let $x(t)$ be the



x-coordinate at time t of the bottom end of the rod, $\phi(t)$ the angle that the rod makes with the horizontal, and $h(t)$ the height of the center of the rod. Finally, let $y(t)$ be the y-coordinate of the bottom end of the rod—since we do not rule out the possibility that the rod will eventually rise above the surface—with

(a) $$h = y + l \sin\phi.$$

The magnitude $N(t)$ of the normal force of the rod on the surface at time t is also the magnitude of the force exerted by the surface upward on the rod at time t, so we have

$$\text{(b)} \qquad x'' = -\mu N$$
$$\text{(c)} \qquad h'' = -g + N.$$

Finally, the moment of inertia of the rod about the axis passing through the center and perpendicular to the plane of the figure is $\frac{1}{3}l^2$ (cf. Problem 5-6). So equation (b) gives

$$\tfrac{1}{3}l^2\phi'' = l(\mu \sin\phi - \cos\phi)N,$$

or

$$\text{(d)} \qquad \phi'' = 3N(\mu \sin\phi - \cos\phi)/l.$$

Now (a) gives

$$h'' = y'' + l(\phi'' \cos\phi - \phi'^2 \sin\phi),$$

and when we substitute in from (c) and (d) we get

$$\begin{aligned} y'' &= [1 + 3\cos\phi(\cos\phi - \mu\sin\phi] \cdot N \;+\; [l\phi'^2 \sin\phi - g] \\ &= A \cdot N \;+\; b, \quad \text{say.} \end{aligned}$$

Since the force N arises from the contact of the surface with the rod, which cannot go below the surface, we have two additional conditions: $y''(t) > 0 \implies N(t) = 0$ and $N(t) > 0 \implies y''(t) = 0$.

When $\boldsymbol{A > 0}$, it is easy to find the solution for $b \neq 0$:

For $b > 0$, we have $y'' > 0$, so $N = 0$ and thus $y'' = b$.

For $b < 0$, we can't have $y'' > 0$, for then we would have $N = 0$, and thus $y'' = b < 0$. So $y'' = 0$, and thus $N = -b/A$.

When $\boldsymbol{A < 0}$, on the other hand, things are quite different:

If $b > 0$, the solution is not unique; in fact, we have the two solutions given previously,

$$y'' = b \text{ and } N = 0,$$
$$y'' = 0 \text{ and } N = -b/A.$$

If $b < 0$ then there is no solution, since we would have $y'' < 0$.

But the condition $A < 0$ can certainly occur, simply by making μ large enough. More specifically, we have $A = 0$ for

$$\mu = \frac{1 + 3\cos^2\phi}{3\sin\phi\cos\phi} = \frac{1}{3}(\sin\phi\cos\phi)^{-1} + \cot\phi,$$

and the minimum value of this for all ϕ is for $\phi = \arctan 2$, with μ having the value[1] $\mu = \frac{4}{3}$. Moreover, given $\mu \geq \frac{4}{3}$, we have $A = 0$ for

$$3\mu \sin\phi \cos\phi = 1 + 3\cos^2\phi,$$

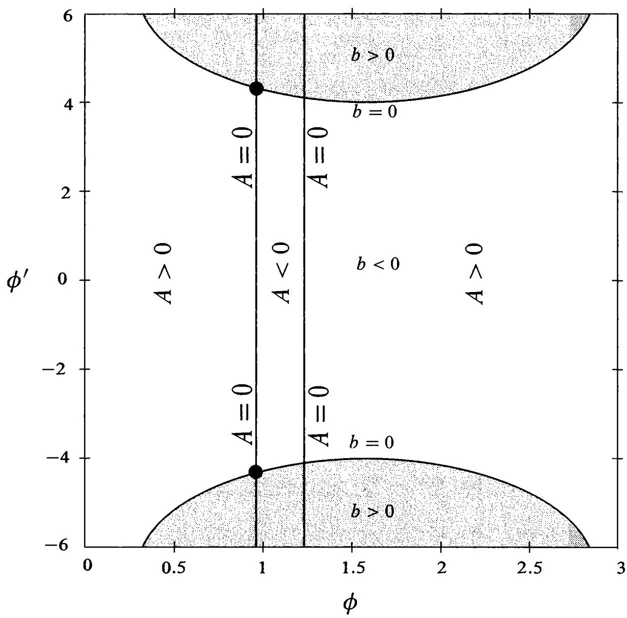
or

$$3\mu \tan\phi = \frac{1}{\cos^2\phi} + 3 \implies \tan^2\phi - 3\mu\tan\phi + 4 = 0,$$

so we have $A < 0$ for any ϕ between

$$\arctan\left(\frac{3\mu - \sqrt{9\mu^2 - 16}}{2}\right) \quad \text{and} \quad \arctan\left(\frac{3\mu + \sqrt{9\mu^2 - 16}}{2}\right).$$

This is summed up in the following figure, based on Génot and Brogliato [1],



where a detailed analysis is given to show that there are solutions that reach the two singular points indicated by dots, at which existence and uniqueness of solutions no longer holds. At these two points we have $A = 0$ and $b = 0$, which would lead us to expect that $y'' = 0$. But in fact there are solutions where we have $y'' \neq 0$, which turns out to be possible because as we approach these points, the normal load N becomes infinite.

[1] This is quite large, coefficients of friction usually being considerably less than 1, but more realistic models of a piece of chalk, as a thin cylinder with rounded edges, give more realistic values for μ.

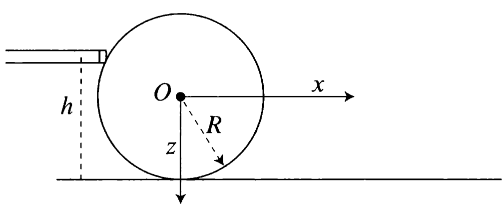
Of course, none of these paradoxical phenomena pose any threat to our basic understanding of theoretical mechanics, depending as they do on empirical laws that we know to be only approximate, treating situations that would become hopelessly complicated to describe accurately in detail. But they present great challenges for the actual design of machinery, and form an important area of research in applied, or engineering, mechanics, where the goal is to obtain reasonable models that can handle these situations. These models, which may involve considerably different approaches and assumptions, often depend on rather complex and sophisticated mathematics. Overviews, and extensive bibliographies, may be found in Brogliato [1] and Anh [1].

Extracts from Painlevé's work, as well as discussion and references to the criticisms that followed—basically early attempts to create models for the problem—can also be found in the imposing work Hamel [1; pp. 543 ff., 629 ff.].

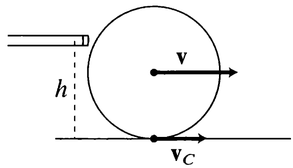
The noble game of billiards. In Chapter 6 we considered the theoretical case of a perfect sphere rolling on a flat rigid surface, with friction playing the irrelevant, yet essential, role of insuring rolling. Of course, that picture would not be entirely correct even for a perfectly spherical billiard ball, touching the table at a theoretical single point, because the weight of the ball causes the cloth of the table to compress slightly, and this produces friction, which would eventually causes a rolling ball to stop. However, this effect is very small, and we will ignore it.

Friction does play an extremely important role in the game of billiards, but for quite a different reason: a billiard ball is often *not* rolling, but instead moving with a combination of spinning and sliding, with the speed of the center of the ball and the rate of rotation not being in the relation required for rolling, even if the ball does happen to be rotating about an axis.

The figure below shows the simplest situation where at time $t = 0$ a billiard ball of radius R has been hit straight on by the cue, with the contact point at



height h above the billiard table. For convenience we set the mass of the ball to be $m = 1$, and choose the origin O of our coordinate system to be the center point of the ball at $t = 0$, with the x-axis in the direction of the cue, the y-axis also parallel to the table (perpendicular to the plane of the paper in this figure), and the z-axis vertical, with the positive part pointing downward, so that it passes through the point of contact. The strike of the cue, which is usually quite abrupt, can be thought of as imparting an impulsive force **P** that causes the



center of the ball to move along the x-axis with velocity function $\mathbf{v}$ satisfying

$$v(0) = |\mathbf{v}(0)| = |\mathbf{P}|,$$

and, as in Chapter 6, we let $\mathbf{v}_C$ be the velocity function of the contact point.

At $t = 0$ the torque of the force $\mathbf{P}$ about the center of the ball is

$$\boldsymbol{\tau}(0) = (h - R)v(0) \cdot (\text{unit vector along the } z\text{-axis}),$$

so for the angle θ through which the ball has rotated around the y-axis, measured clockwise, we have

$$I\theta'(0) = (h - R)v(0),$$

where the moment of inertia I of the ball is $\frac{2}{5}R^2$ (Problem 5-6). It follows that

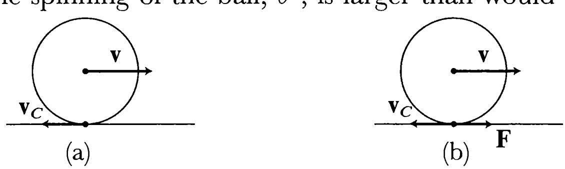
$$\theta'(0) = \frac{5}{2}\left(\frac{h-R}{R^2}\right)v(0),$$

which implies that at $t = 0$ the vector $\mathbf{v}_C$ has magnitude

$$v_C(0) = v(0) - R\theta'(0) = \left(\frac{7R - 5h}{2R}\right)v(0).$$

For the ball to start out with a rolling motion we must have (compare page 240), $v_C(0) = 0$, which means that the ball must be hit at exactly the height $h = \frac{7}{5}R$, which just so happens to be the height of the cushions—presumably found from experience by the makers of billiard tables.

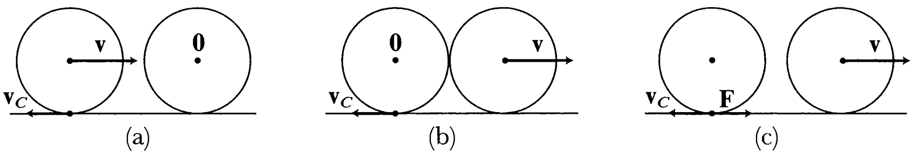
For a "high shot", with $h > \frac{7}{5}R$, the direction of $\mathbf{v}_C$ will, as in (a), be opposite to that of $\mathbf{v}$—the spinning of the ball, θ', is larger than would be expected for



rolling, given the speed v of the center. But now the non-zero velocity $\mathbf{v}_C$ gives rise to a force of friction **F** in the opposite direction (b), and this force acts on the rigid ball as a whole: it causes v to increase, and at the same time it causes θ' to decrease until rolling begins, and the ball then simply continues to roll.

For a "low shot", with $h < \frac{7}{5}R$, the situation is exactly the opposite: the direction of $\mathbf{v}_C$ will be the same as that of $\mathbf{v}$, with θ' less than expected, possibly even negative (with the ball spinning backwards). So the force of friction will be in the opposite direction as $\mathbf{v}$, causing v to decrease, and θ' to increase until rolling begins.

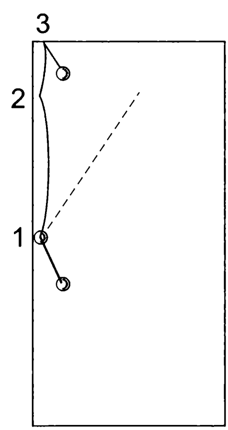
If a player wants the ball to roll, it is only necessary to strike it with h close to $\frac{7}{5}R$, and rolling will soon ensue. But sometimes high or low shots are specifically used to control the behavior of the cue ball after it collides with an object ball at rest. In the case of a high shot (a), since the balls have the same mass,



and the collision is almost perfectly elastic, at the moment of collision (b) the velocities are interchanged, so that the cue ball now has velocity $\mathbf{0}$, and the object ball acquires the velocity $\mathbf{v}$ in the same direction. If the cue ball hasn't reached the rolling stage, so that $\mathbf{v}_C \neq 0$, then (c) the friction force $\mathbf{F}$ causes the cue ball to move in the same direction, so that we have a *follow shot.*

In the case of a low shot, the situation is again exactly the opposite: now $\mathbf{v}_C$ points in the same direction as $\mathbf{v}$, so $\mathbf{F}$ points in the opposite direction, causing the cue ball to move backward, a *draw shot.*

This analysis of the simple case of billiard balls rolling on a straight line can only whet the appetite of mathematically inclined billiard players, for it hardly touches on the intricacies encountered in the game. In fact, the 19^{th} century saw the appearance of a thorough analysis of the dynamics of billiards, including the impressive shot shown below. As with all the studies previously mentioned, it was written by some one who is nowadays much better known for other work.



Point scored in a game of three-cushion billiards

Cueball drove first object ball along dotted path and followed solid path, contacting the long rail at point 1, curving back to the same rail at point 2, and continuing to point 3 on the end rail to score on the second object ball

Biographical accounts of Coriolis (1792–1843) [basically from *Dictionary of Scientific Biography*] note that he was a student at the École Polytechnique and worked for several years in the engineering corps, until his poor health led him in 1816 to accept a position as tutor at the École Polytechnique; his life was from then on was dedicated to the teaching of science, and he eventually became a greatly admired Director of Studies. He felt that the results of *mechanique rationelle* ("rational mechanics") should be used to give general principles applicable to the operation of machinery, and in his first book, *Du calcul def l'effet des machines* ("On the Calculation of Mechanical Action"), he introduced the modern meaning of the term "work" in physics, as well as the proper factor $\frac{1}{2}$ into the definition of kinetic energy for conservation of energy to hold. (His investigations of 1835 into the "Coriolis force" were made to account for conservation of energy, and thus followed a much more complicated path than our modern purely kinematic approach.)

This all sounds rather dutiful, if not a bit dreary, but Coriolis presumably had his diversions and times of relaxation—in the year 1835 he also published his second book: *Theéorie mathématique des effets du jeu du billiard* ("Mathematical Theory of the Game of Billiards") [though in his preface he dutifully says "I think that persons acquainted with rational mechanics, such as the students at the École Polytechnique, will be interested in the explanation of the surprising effects observed in the motion of billiard balls."]. By the way, in case you were wondering, the Coriolis force is not relevant to the game of billiards—unless, of course, you are playing on a rotating billiard table, an extra delight that Gilbert and Sullivan's Mikado forgot to include when making the punishment fit the crime of being a billiard shark.

Until its reprinting as Coriolis [1], this book was extremely hard to find, even if one was up to reading the French. So I was delighted to learn from a friend of old that he had made an English translation, Coriolis [2], which presents Coriolis' theory, together with accounts of the careful experiments he made to determine coefficients of friction.

Though Coriolis' exposition is quite straightforward, modern terminology helps the exposition. The material presented here, covering just a bit of Coriolis' work, has appeared in various classical works on mechanics, and can be found in notes available on the web.[1] Having these notes at hand may help smooth the study of the rest of Coriolis' book—for those readers in pursuit of the lost time of their misspent youth.

[1] They may be found at the web site `http://billiards.colostate.edu`; click on "Technical Proof (TP) analyses", and scroll down to the TP A section; here we are covering material from TP A.4.

We first study the path of the cue ball when it is not necessarily hit head on, but possibly to the left or right of center (with *side english*). Coriolis considers both the sliding friction and the very small rolling friction for the initial part of his analysis, and then ignores the rolling friction later on, but we will simplify things by ignoring it right from the start, which doesn't actually affect the outcome.

We will use the same coordinate system as before, letting $\mathbf{e}_1$ be a unit vector pointing along the x-axis, with $\mathbf{e}_2$ pointing along the y-axis, and $\mathbf{e}_3$ pointing (downward) along the z-axis. We now use the more general equation on page 190,

$$\boldsymbol{\tau} = I\boldsymbol{\omega}' = \frac{2}{5}R^2\boldsymbol{\omega}', \tag{1}$$

with the velocity $\mathbf{v}_C$ of the contact point given by

$$\mathbf{v}_C = \mathbf{v} + \boldsymbol{\omega} \times R\mathbf{e}_3. \tag{2}$$

It will be convenient to let $\mathbf{u}(t)$ denote the unit vector in the direction of $\mathbf{v}_C(t)$, and write the frictional force at the contact point as

$$\mathbf{F} = -\mu\mathbf{u}$$

(where μ really denotes the product of the coefficient of friction, the mass m, which we've taken to be 1, and the force of gravity g, to get the weight). This means that the acceleration $\mathbf{v}'$ of the ball is

$$\mathbf{v}' = -\mu\mathbf{u}. \tag{3}$$

The torque $\boldsymbol{\tau}$ of the force $\mathbf{F}$ about the center of the ball is

$$\boldsymbol{\tau} = R\mathbf{e}_3 \times \mathbf{F} = -\mu R(\mathbf{e}_3 \times \mathbf{u}),$$

so equation (1) gives

$$\boldsymbol{\omega}' = -\frac{5\mu}{2R}\mathbf{e}_3 \times \mathbf{u}. \tag{4}$$

Differentiating (2), and substituting from (3) and (4), we have

$$\mathbf{v}_C{}' = -\mu\mathbf{u} - \frac{5\mu}{2R}(\mathbf{e}_3 \times \mathbf{u}) \times R\mathbf{e}_3$$

or simply

$$\mathbf{v}_C{}' = -\frac{7\mu}{2}\mathbf{u}. \tag{5}$$

Notice that $\mathbf{v}_C{}'$ is always in the direction of $\mathbf{v}_C$, since $\mathbf{u}$ is, by definition. But this implies that $\mathbf{v}_C$ *does not change direction*, and thus that $\mathbf{u}$ is a constant vector.

So from (3) we can write

$$\mathbf{v}(t) = \mathbf{v}(0) - \mu t \mathbf{u}, \tag{6}$$

and thus the (center of the) ball follows a path c with

$$c(t) = c(0) + t\mathbf{v}(0) - \tfrac{1}{2}\mu \mathbf{u} t^2,$$

which is a parabola. From (2) we have the explicit formula

$$\mathbf{u} = \frac{\mathbf{v}(0) + \boldsymbol{\omega}(0) \times R\mathbf{e}_3}{v_C(0)}, \qquad v_C = |\mathbf{v}_C|. \tag{7}$$

This all holds, of course, only while the ball is *not* rolling—as we found in Chapter 6, as soon as it starts rolling, it will continue on a straight path, along the tangent line to the parabola. Since (5) gives

$$\mathbf{v}_C(t) = \mathbf{v}_C(0) - \frac{7\mu t}{2}\mathbf{u},$$

and rolling starts at the time t_* where we first have $\mathbf{v}_C(t_*) = 0$, we see that

$$t_* = \frac{2v_C(0)}{7\mu}.$$

So the velocity $\mathbf{v}(t_*)$ when the ball starts rolling is, from (6) and (7),

$$\mathbf{v}(t_*) = \mathbf{v}(0) - \mu\left(\frac{2v_C(0)}{7\mu}\right)\mathbf{u} = \mathbf{v}(0) - \frac{2v_C(0)}{7}\left[\frac{\mathbf{v}(0) + \boldsymbol{\omega}(0) \times R\mathbf{e}_3}{v_C(0)}\right]$$

or

$$\mathbf{v}(t_*) = \frac{5}{7}\mathbf{v}(0) - \frac{2}{7}\boldsymbol{\omega}(0) \times R\mathbf{e}_3. \tag{8}$$

We can write this result explicitly as

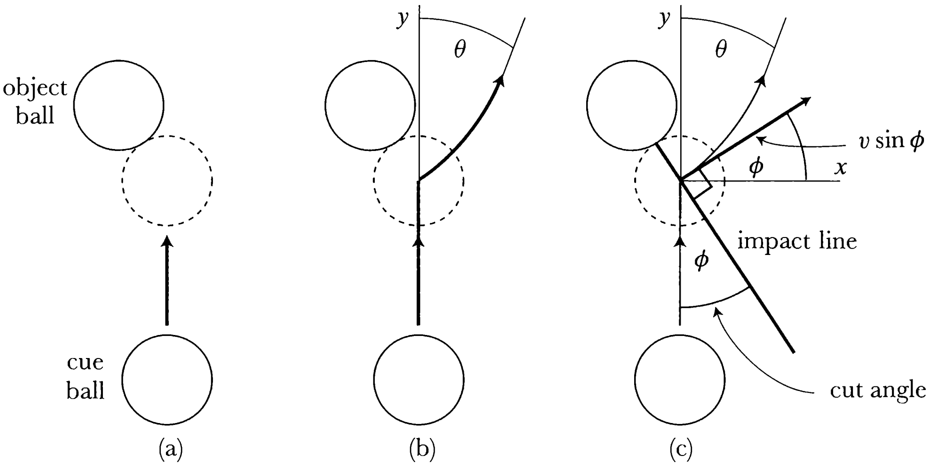
$$\mathbf{v}(t_*) = \frac{1}{7}\left[\big(5\mathbf{v}_1(0) + 2R\omega_2(0)\big)\mathbf{e}_1 + \big(5\mathbf{v}_2(0) - 2R\omega_1(0)\big)\mathbf{e}_2\right].$$

For the final deflected angle θ, which it will be convenient to measure from the y-axis rather than from the x-axis, we then have

$$\theta = \arctan\left(\frac{5\mathbf{v}_1(0) + 2R\omega_2(0)}{5\mathbf{v}_2(0) - 2R\omega_1(0)}\right). \tag{9}$$

Note that the quantity μ has disappeared from these equations, so the coefficient of friction of the ball on the cloth has no effect on this final result, even though the path itself would change.

Now suppose instead that the cue ball is hit straight on, without side english, along the y-axis to collide (a) at an angle with a stationary object ball. The collision will give the cue ball a spin and we want to consider the path of the cue ball after collision (b), and the deflected angle θ. Problem 6-4 shows that



the velocity of the cue ball after the (perfectly elastic) collision (c) will be perpendicular to the "impact line" (the line perpendicular to the two balls at the point of contact), and that if the velocity of the cue ball is v at impact, then the velocity after impact will have magnitude $v \sin \phi$, where ϕ is the "cut angle" between the original direction of the cue ball and the impact line. From the diagram, we see that this initial velocity, after the impact, is

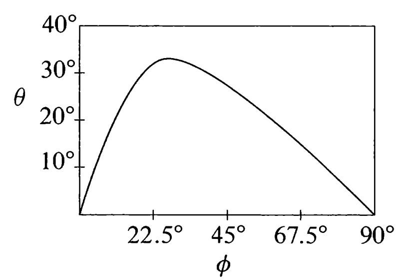
$$\mathbf{v}(0) = (v \cos \phi \sin \phi,\ v \sin^2 \phi,\ 0).$$

We are assuming that at impact, which we will consider as time $t = 0$, the cue ball has $\boldsymbol{\omega}_2(0) = 0$. Setting $\omega = \boldsymbol{\omega}_1(0)$, equation (9) becomes

$$\theta = \arctan\left(\frac{5\mathbf{v}_1(0)}{5\mathbf{v}_2(0) - 2R\omega}\right) = \arctan\left(\frac{5v \sin \phi \cos \phi}{5v \sin^2 \phi - 2R\omega}\right).$$

In particular, suppose that the cue ball is rolling at the time of impact. Then $\omega = -v/R$, so we get

$$\theta = \arctan\left(\frac{\sin \phi \cos \phi}{\sin^2 \phi + \frac{2}{5}}\right).$$

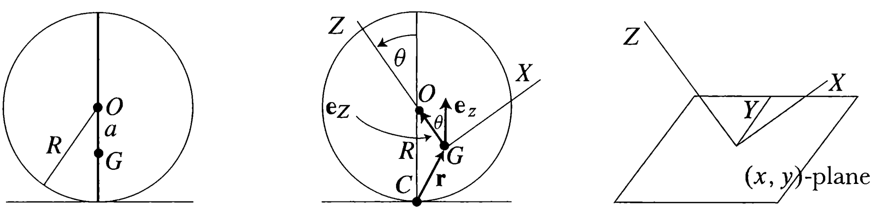


The graph of θ as a function of ϕ shows that for collisions that are neither too close to head on, nor too glancing, θ is about 30°, giving the "30 degree law" for estimating the direction in which the cue ball will bounce. You can find videos on the web explaining this rule, with discussion and computations at `http://billiards.colostate.edu`.

Apparently, billiard and pool players actually make use of rules like these, though I don't know if such crutches are ever needed by those who can execute shots like that shown on page 419. This shot is analyzed in Coriolis' book, whose preface contains the acknowledgment: "Monsieur de Tholozé, governor of the École Polytechnique, graciously showed me several complicated shots, which the theory afterwards explained: it is he whom I saw make the shot diagrammed ... which [is] explained in the course of the work."

Have fun, and good luck, reading the book!

The Jellett invariant. In Chapter 9 we analyzed the "heavy symmetrical top", corresponding to a typical toy top with a pointed end, which we assumed was fixed. By contrast, the last chapter of Jellett [1], one of the pioneering works in the theory of friction, analyzes a top with a non-pointed end. Jellett basically restricted his considerations to the case where the bottom was shaped like a part of a ball, and it is somewhat more convenient to consider a body that simply is in the shape of a ball, as in (a), with center O and radius R, with a density that is not necessarily uniform, though it is symmetric about an axis, so that the



(a) (b) (c)

center of mass G lies on this axis, at a distance a from O. The ball is thus a symmetric top with principal moments of inertia I_1, I_1, I_3.

In (b), showing the body after it has moved along the supporting surface, now at contact with the surface at C, we have introduced X, Y, and Z-axes along the principal directions (with the Y-axis pointing into the plane of the paper in this figure), as well as the Euler angle θ and the unit vectors $\mathbf{e}_Z$ and $\mathbf{e}_z$ along the Z-axis and the unrotating z-axis. Here we are allowing our (x, y, z)-plane to move so that G is always at the origin, as mentioned on page 344.

It will also be convenient to allow the X and Y axes to rotate within the body, as anticipated in Problem 9-8, and we choose X to lie in the plane through the

Z-axis that contains the contact point C, as in the figure, so ϕ measures the angle by which this plane rotates. The line of nodes, the intersection of the (x, y)-plane and the (X, Y)-plane, has the Y-axis lying in it (c), perpendicular to the (X, Z)-plane, so the angle ψ needed to bring the line of nodes to X is just $\psi = -\pi/2$. Thus our equations (ω) on page 344 simplify to

$$(\omega_s) \qquad \omega_1 = \phi' \sin\theta, \qquad \omega_2 = -\theta', \qquad \omega_3 = \phi' \cos\theta,$$

which one could also easily see directly.

The figure also shows the vector $\mathbf{r}$ from the contact point C to G. Note that since the distance from G to O is a, we have $\mathbf{r} + a\mathbf{e}_Z = R\mathbf{e}_z$ or

$$\text{(i)} \qquad \mathbf{r} = R\mathbf{e}_z - a\mathbf{e}_Z,$$

while the angular momentum $\mathbf{L}$ about G can be written as

$$\text{(ii)} \qquad \mathbf{L} = I_1\omega_1\mathbf{e}_X + I_1\omega_2\mathbf{e}_Y + I_3\omega_3\mathbf{e}_Z = I_1\boldsymbol{\omega} + (I_3 - I_1)\omega_3\mathbf{e}_Z.$$

If $\mathbf{F}$ is the total force applied to the ball at C, including the upward force of the surface on which the sphere is moving that balances out the downward gravitational force on the ball, as well as the force of friction, about which we will make no assumptions at all, then the torque $\boldsymbol{\tau}$ about the center of mass is simply $\boldsymbol{\tau} = -\mathbf{r} \times \mathbf{F}$, so we have

$$\text{(iii)} \qquad \mathbf{L}' = \boldsymbol{\tau} = -\mathbf{r} \times \mathbf{F}.$$

Now (i) gives $\mathbf{r}' = -a \cdot \boldsymbol{\omega} \times e_Z$, so (ii) gives

$$\begin{aligned} \langle \mathbf{L}, \mathbf{r}' \rangle &= \langle \mathbf{I}_1\boldsymbol{\omega} + (I_3 - I_1)\omega_3\mathbf{e}_Z,\ -a \cdot \boldsymbol{\omega} \times \mathbf{e}_Z \rangle \\ &= 0, \end{aligned}$$

since $\boldsymbol{\omega} \times \mathbf{e}_Z$ is perpendicular to both $\boldsymbol{\omega}$ and $\mathbf{e}_Z$. On the other hand, from (iii) we get

$$\langle \mathbf{L}', \mathbf{r} \rangle = 0,$$

and from these last two equations we conclude that $\langle \mathbf{L}, \mathbf{r} \rangle' = 0$, and thus that

$$\text{(1)} \qquad \langle \mathbf{L}, \mathbf{r} \rangle = \text{constant}.$$

Resubstituting $\mathbf{r} = R\mathbf{e}_z - a\mathbf{e}_Z$ into (1), and dividing by R, we get

$$\text{(2)} \qquad \langle \mathbf{L}, \mathbf{e}_z \rangle - \varepsilon \langle \mathbf{L}, \mathbf{e}_Z \rangle = J, \qquad \text{where } \varepsilon = a/R$$

for a constant J (*Jellett's constant*).

In either form, this may seem like a pretty elementary conclusion, but it is noteworthy because it holds no matter what is assumed about the frictional forces or motion of the top, so long as it remains in contact with the plane.

We can also make the result look a lot less elementary by writing it in terms of the Euler angles! Substituting

$$\langle \mathbf{L}, \mathbf{e}_z \rangle = L_1 \sin\theta + L_3 \cos\theta = I_1\omega_1 \sin\theta + I_3\omega_3 \cos\theta$$

into (2), we obtain

$$I_1 \sin\theta \cdot \omega_1 + I_3(\cos\theta - \varepsilon)\omega_3 = J$$

and then by (ω_s)

$$\text{(J)} \qquad I_1 \sin^2\theta \cdot \phi' + I_3(\cos\theta - \varepsilon)\omega_3 = J,$$

which is basically the form in Jellett [1; pg. 185], who used the same special sort of coordinates as used here.

This result is called *Jellett's integral*, where "integral" is here being used in the same sense that conservation of momentum is called an integral of the equations of motion. This 19th century discovery was long ignored and almost completely forgotten, until rescued from oblivion by a 20th century toy.

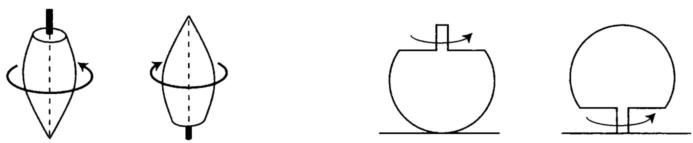
Tippe Tops and hard boiled eggs. The Tippe Top was invented, or perhaps merely reinvented, in 1950 by a Danish engineer Werner Østberg, who reported that on a visit to South America he saw people playing with a small fruit, which when spun by its stem would turn over and end up spinning on the stem. It appears that a typical college ring exhibits similar behavior—when spun with the heavy stone at the bottom, it turns over and spins with the stone on top. If you don't have a college ring, the toys can still be bought, quite inexpensively, and are usually in the shape of a sphere with a small section capped off, with a

(a) (b)

stem attached. In (a), at the beginning of the spin, the center of mass G of the Tippe Top lies slightly below the center O of the original sphere, while in (b), at the end of the spin, it ends up above it.

The Tippe Top seems to have fascinated physicists as well as kids. An oft-mentioned photograph from 1951 shows Wolfgang Pauli and Niels Bohr bent down peering at a Tippe Top in action,[1] and in 1952 and 1953 the basic principles were explained in several papers, followed by numerous later papers, some quite recent.

The first thing one immediately notices about the Tippe Top is that its potential energy *increases* when it turns over, since the center of mass ends up higher. A second fact is not so immediately apparent. If one takes a toy top of the sort mentioned on page 349, spinning around a protruding axis that can be held in the hand, and then turns the top over (a), it is naturally now spinning in the other direction. The angular momentum has reversed sign, thereby changing

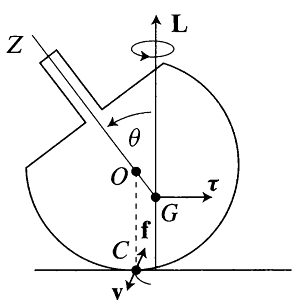


(a) (b)

by a fairly large amount; this requires a torque, which is why it's harder to invert the top when it is spinning. On the other hand, as observation confirms, when the Tippe Top turns itself over (b), it ends up spinning in the *same* direction as when it started, so the spin reverses direction within the body coordinates.

Nevertheless, since the angular momentum determines the kinetic energy, and since the potential energy increases slightly when the top turns over, the angular momentum must decrease slightly when it turns over. So a small torque is still required to produce this change of angular momentum. And just where does this torque come from? The answer is, it comes from friction. Here is a rough, approximate, description of the process.

To begin with, when the top is at an angle, and rotating about the vertical axis through its center of mass G, the contact point C, lying below O, is moving



[1] It can be found by googling `pauli picture`.

in a small curve, whose velocity vector is a horizontal vector $\mathbf{v}$ nearly perpendicular to the plane of the diagram. Thus the contact point is sliding, giving rise to a frictional force $\mathbf{f}$ in the opposite direction (pointing into the plane of the diagram), and the torque $\boldsymbol{\tau}$ of $\mathbf{f}$ around G is nearly horizontal, since O and G are very close. This torque rotates around with the same angular velocity ω as the top, and averages out to 0, so $\mathbf{L}' \approx 0$ on the average, and angular momentum is nearly conserved, as we basically already noted when we mentioned that the top ends up spinning in the same direction. Thus we have approximately $\mathbf{L} = \omega \cdot \mathbf{e}_Z$.

The principal moments of inertia I_1 and I_3 for a Tippe Top are usually close in value, so for a very rough seat-of-the-pants approximation[1] we will assume that they are actually equal. As shown in Problem 9-7, this implies that we then also have $\boldsymbol{\tau} = \mathbf{L}'$ *in body coordinates*. For the unit vector $\mathbf{e}_Z$ along the Z-axis, which is constant in body coordinates, we thus have

$$\langle \mathbf{e}_Z, \boldsymbol{\tau} \rangle = \langle \mathbf{e}_Z, \mathbf{L}' \rangle = \langle \mathbf{e}_Z, \mathbf{L} \rangle'.$$

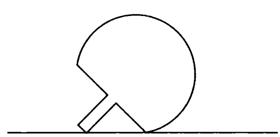
But

$$\begin{aligned} \langle \mathbf{e}_Z, \boldsymbol{\tau} \rangle &= -|\boldsymbol{\tau}| \sin\theta, \\ \langle \mathbf{e}_Z, \mathbf{L} \rangle' &= \big(|\mathbf{L}| \cos\theta\big)' = -|\mathbf{L}| \sin\theta \cdot \theta', \end{aligned}$$

giving

$$\theta' = |\boldsymbol{\tau}|/|\mathbf{L}| = \mu W R / I\omega,$$

where μ is the coefficient of friction, W the weight of the top, and R the radius of the original sphere. Thus θ keeps increasing, until the stem touches the plane (at some time $< \pi/\theta'$). From then on a similar argument about friction on the stem shows that the top will start to rise.



The early Tippe Top papers, which are referenced in some later papers that we will mention, gave considerably more detailed analyses, although they all used approximations of some sort, depending on the particular aspect of the top's behavior that they were focused on. One detail that we want to investigate is the process by which the top ends up revolving in the same direction.[2]

[1] Barger and Olsson [1].
[2] Pliskin [1].

Until the stem touches the plane, the center of mass G of the top is practically stationary, and we will consider it as the fixed point for the Euler equations, remembering, however, that we chose X and Y in a special way, so that they are not fixed in the body. If α is the angular velocity of the contact point C around the Z-axis, then the second Euler equation in Problem 9-8 becomes

$$\tau_2 = I_1\omega_2{}' + (I_1 - I_3)\omega_1\omega_3 - I_3\alpha\omega_1.$$

Pretending, as before, that the force of friction is exactly perpendicular to the plane of the diagram on page 428, the only torque about G along the Y-axis will be due to the upward force of magnitude W acting at C, at a perpendicular distance $a \sin\theta$ from G, and we find that $\tau_2 = aW\sin\theta$. So we have

$$aW\sin\theta = I_1\theta'' + (I_1 - I_3)\omega_1\omega_3 - I_3\alpha\omega_1.$$

Once again we approximate $\mathbf{L} = \omega\cdot\mathbf{e}_Z$, which implies that

$$\omega_1 = \omega\sin\theta, \qquad \omega_3 = \omega\cos\theta,$$

so that our equation becomes

$$aW\sin\theta = I_1\theta'' + (I_1 - I_3)\omega^2\sin\theta\cos\theta - I_3\alpha\omega\sin\theta.$$

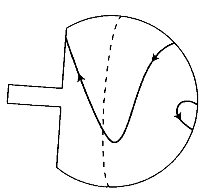
Now the Tippe Top's behavior indicates that θ' is small compared to ω (the top is set spinning at many revolutions per second, while it takes several seconds for the top to turn over). It seems reasonable to conclude that θ'' is also small (at least on average). If we therefore simply eliminate the $I_1\theta''$ term, we end up with

$$\alpha = \frac{aW}{I_3\omega} + \frac{(I_1 - I_3)}{I_3}\omega\cos\theta.$$

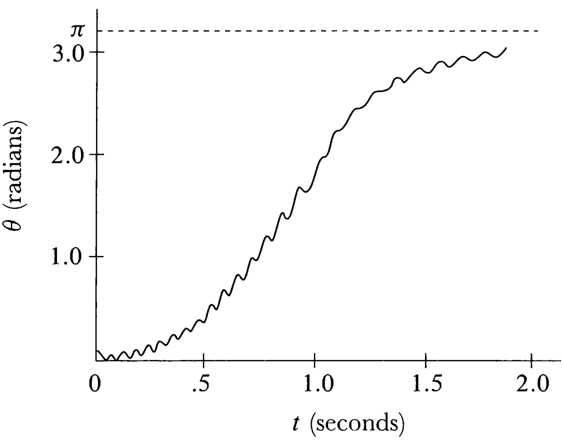
Thus $\alpha = 0$ for θ_0 satisfying

$$\cos\theta_0 = \frac{aW}{\omega^2(I_1 - I_3)},$$

and α changes sign at θ_0. When ω is very large, θ_0 is close to $\pi/2$, with the stem nearly horizontal at the time that the spin reverses. If a Tippe Top is set spinning on carbon paper, it gives a carbon trace of the path of C along the Tippe Top that illustrates these facts.



These rough approximations illustrate the general tenor of the discussion of the early Tippe Top papers. Then in 1974 Cohen [1], inspired after idly spinning his college ring, simply wrote down the complete set of equations and found computer-generated solutions for "these horribly nonlinear differential equations" without approximations, though he did work with a sphere, as in the analysis of Jellett's integral, the results then being valid for a Tippe Top up to the point where the stem touches the surface. They attest both to the general correctness of the simplifying assumptions, and the extent to which they oversimplify things. For example, the graph of θ over time looked something like the figure below.



The next stage of the Tippe Top investigations involved the rediscovery of Jellett's integral. The classical form of Jellett's integral was first rediscovered in an early Tippe Top paper by Hugenholtz [1], where it was used to discuss certain aspects of the Tippe Top's motion.

But Jellett's integral was given a decisive role in a paper by Leutwyler [1] in 1993, which was "written to provide entertainment", with the author being unencumbered by familiarity with the historical literature. Leutwyler once again rediscovered Jellett's integral, now in the form (2) on page 425, remarking that this integral "seems to have escaped attention", but serendipitously denoting the constant by J.

Leutwyler actually discovered only a special case of Jellett's integral, and his demonstration was rather strange, but the importance of his paper was the way in which he used Jellett's constant. He determined when the total energy E has the lowest possible value for a fixed value of J, the idea being that sliding friction will cause the top to loose energy, so it should end up in a position with E a minimum.

Before presenting Leutwyler's argument, we might point out that Jellett himself actually proved his integral only as an approximation, making simplifying assumptions about friction. The general result was then proved in the classic (distressingly difficult to read) book Routh [1; Art. 241c], where it was used to analyze the completely solvable case of a top with a spherical bottom rolling without slipping on a plane, an investigation that was completely ignored, as pointed out in 1999 by Gray and Nickel [1], which is basically a rewriting of Routh's results in modern terminology, together with an extensive bibliography of older papers. Our proof of Jellett's integral comes from this paper, which also reintroduces the much more complicated, much less well known, Routh integral.

To derive Leutwyler's result, leaving some of the details to his paper, we want to work with an ordinary set of principal directions fixed in the top, just like those used in Chapter 9. We find that now

$$J = \phi'(I_1 \sin^2\theta + I_3 \cos\theta(\cos\theta - \varepsilon)) + \psi' I_3(\cos\theta - \varepsilon).$$

Similarly, using the equation for the rotational energy T on page 346, we write the total energy, for a top of mass M, as

$$\begin{aligned} E = {} & \frac{M}{2}(x'^2 + y'^2 + z'^2) \\ & + \frac{I_1}{2}(\theta'^2 + \phi'^2 \sin^2\theta) + \frac{I_3}{2}(\phi' \cos\theta + \psi')^2 + Mgz, \end{aligned}$$

where we have $z = R(1 - \cos\theta)$.

Since J doesn't involve x', y', or θ', the minimum of E for a fixed J obviously occurs when $0 = x' = y' = \theta'$, so x, y, and θ are constant, which also implies that $z = R(1 - \cos\theta)$ is constant; thus the center of mass and the angle of rotation are fixed at a minimum for E.

Now for each fixed θ we look for the minimum of E for a fixed value of J. Computations show (cf. Problem 4) that this occurs when $\psi' + \varepsilon\phi' = 0$, with the energy being

$$E(\theta) = \frac{J^2}{2\mathcal{I}(\theta)} + MgR(1 - \varepsilon\cos\theta),$$

where

$$\mathcal{I}(\theta) = I_3(\cos\theta - \varepsilon)^2 + I_1 \sin^2\theta.$$

When the top spins rapidly, so that ϕ' and hence J is large, the minima of the $E(\theta)$ are the maxima of the $\mathcal{I}(\theta)$. We have $\mathcal{I}(0) < \mathcal{I}(\pi)$, so that, as expected, the top has less energy spinning upside down then in its initial condition. In fact, it turns out that for

$$(1-\varepsilon)I_3 < I_1 < (1+\varepsilon)I_3$$

the function $\mathcal{J}(\theta)$ is increasing, so that the top moves into the inverted position. If instead

$$(1+\varepsilon)I_3 < I_1,$$

then $\mathcal{J}(\theta)$ has a minimum close to

$$\cos\theta = -\frac{\varepsilon I_3}{I_1 - I_3},$$

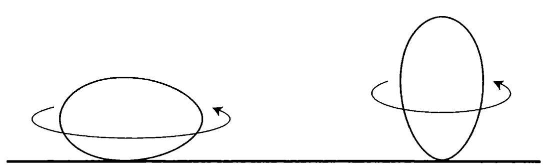
and the top ends up rolling around a vertical axis through the center of mass with the contact point moving in a circle of radius $a\sin\theta$. Finally, if

$$I_1 < (1-\varepsilon)I_3,$$

then both $\theta = \pi$ and $\theta = 0$ are local minima, so the initial position is stable.

Leutwyler's rediscovery of Jellet's integral was followed by several other papers with particular emphasis on questions of stability, using the integral to reduce the number of equations required for a complete set of equations for the Tippe Top. The latest I know of is Rauch-Wojciechowski, Sköldstam, and Glad [1], which bills itself as "a rigorous, and possibly complete analysis of ... the tippe top ... ", and has references to the important earlier investigations.

We'll end this long, though incomplete, discussion with a more mundane, and some might even say, frivolous, phenomenon. If the college ring is the well-to-do person's Tippe Top, then the poor person's Tippe Top might very well be the hard boiled egg. When a hard boiled egg is placed on a surface and set spinning it will (as experience has shown me) quickly stand up on its long axis and



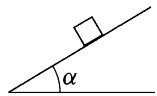
continue spinning—quite a bit more rapidly, because the moment of inertia around the long axis is less than that around the short axis. A spinning football is supposed to behave similarly.

One of the first explanations of the spinning egg behavior was given in Moffatt and Shimomura [1], and a very thorough investigation of the phenomenon was subsequently made by Bou-Rabee, Marsden, and Romero [1]. Both papers rely on the fact that although the Jellett invariant is defined only for a top with a spherical bottom, it is an "adiabatic invariant" for more general shapes. We will talk about adiabatic invariants in Chapter 22, but we won't be returning to the consideration of the spinning egg; I fear that if I tried to explain any details of these papers, I'd probably just end up with egg all over my face.

And yes, there does seem to be something about eggs that encourages frivolity. The paper by Bou-Rabee, Marsden, and Romero is entitled "A geometric treatment of Jellett's egg", while the article by Moffit and Shimomura, which appeared in the March 28, 2002 issue of *Nature*, is announced with the banner: An explanation for an odd egg performance is rolled out in time for Easter.

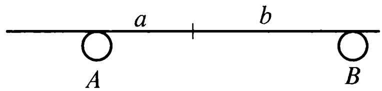
PROBLEMS

1. Consider a plane that can be slanted at various angles α to the floor, and an object placed on it. Show that the condition that the object doesn't move is



$\tan\alpha < \mu_S$. The "angle of friction" ϕ at which the object just starts to move thus gives $\mu_S = \tan\phi$ (while μ itself can be calculated by observing the speed of descent for $\alpha > \phi$).

2. Suppose a smooth cane or walking stick, or even a wooden ruler, of mass m, is balanced on two fingers, at different distances a and b from the midpoint.



We let μ be the coefficient of kinetic friction for wood on fingers, and μ_S the coefficient of static friction. Although we only stated that $\mu \le \mu_S$, in most cases we have $\mu < \mu_S$; this is certainly the case for wood on fingers.

(a) The forces on A and B are

$$\mathbf{F}_A = \frac{b}{a+b}mg, \qquad \mathbf{F}_B = \frac{a}{a+b}mg.$$

Assuming $b > a$, as in the figure, so that $\mathbf{F}_A > \mathbf{F}_B$, conclude that as the fingers are moved toward each other, A stays at the same position on the ruler while B moves closer until it reaches a distance $b_1 < a$ where the sliding friction of B equals the static friction of A, and thus

$$\mu a = \mu_S b_1, \qquad \frac{a}{b_1} = \frac{\mu_S}{\mu} > 1.$$

At this point, A starts moving until it reaches a_1 with

$$\frac{b_1}{a_1} = \frac{\mu_S}{\mu},$$

so A and B approach each other in geometric progression, ending (theoretically after an infinite number of steps, though in practice after only a few) at the midpoint.

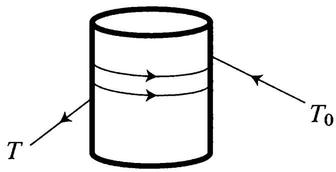
3. Consider the situation in Problem 6-18, except that there is a non-zero coefficient μ of friction between the filament and the fixed object, and suppose that the filament starts slipping in the direction from $c(s)$ to $c(s+h)$. Then there is an additional force at $c(s)$ in the opposite direction with a magnitude close to $\mu|\mathbf{N}|h$. Conclude that we now have

$$T'(s) = \mu|\mathbf{N}|, \qquad |\mathbf{N}| = T\theta'$$

and thus

$$\frac{dT}{d\theta} = \mu T \implies T(\theta) = T_0 e^{\mu\theta}.$$

This exponential increase explains the effectiveness of a capstan, where a rope



wrapped around a post several times allows a very small force to keep a very large one at bay. For example, if the coefficient of friction between the rope and the post is $\frac{1}{2}$, and the rope is wrapped twice around the post, then

$$T = T_0 e^{4\mu\pi} = T_0 e^{2\pi}$$

for a rope beginning to slip in the direction of T. Since

$$\frac{T_0}{T} = e^{-2\pi} = .0019,$$

a load of 2000 pounds can be kept from slipping by a force of about 3.8 pounds.

4. Find the minimum of

$$E(\Phi, \Psi) = \frac{I_1}{2}\sin^2\theta \cdot \Phi^2 + \frac{I_3}{2}(\Phi\cos\theta + \Psi)^2 + Mgz \qquad (\Phi = \phi',\ \Psi = \psi')$$

for a fixed value of

$$J(\Phi, \Psi) = (I_1 \sin^2\theta + I_3(\cos\theta(\cos\theta - \varepsilon)) \cdot \Phi + I_3(\cos\theta - \varepsilon)\Psi$$

by expressing E in terms of Φ (and the constant J) alone. Or, since we've already given the answer, verify it by using the fact (Problem 5-2) that at the desired (Φ, Ψ) there will be a λ for which the partial derivatives satisfy

$$\begin{aligned} E_\Phi &= \lambda J_\Phi \\ E_\Psi &= \lambda J_\Psi, \end{aligned}$$

noting that for $\Psi = -\varepsilon\Phi$, both hold, with $\lambda = \Phi$.

PART III

LAGRANGIAN MECHANICS

CHAPTER 12
ANALYTICAL MECHANICS

> We now come to a piece of work which united and crowned all the efforts which were made in the XVIIIth Century to develop a rationally organised mechanics. *Mécanique analytique* appeared in 1788. ... [Lagrange] became preoccupied with the organisation of mechanics ... the perfection of its mathematical language and the isolation of a general analytical method for solving its problems.
>
> — Dugas, *A History of Mechanics*

Parts I and II have covered material that would generally be regarded as "elementary" mechanics. Although we have allowed ourselves the luxury of using whatever mathematical techniques we needed, we have basically always reduced everything to derivations directly from Newton's laws.

Lagrangian mechanics, the first encounter with "advanced" mechanics, is also called analytical mechanics, as it sedulously shuns geometric arguments, and notably deals with scalar functions in preference to vectors, which so often require diagrams to understand and facilitate computations. The preface to Lagrange's famous *Mécanique Analytique*, Lagrange [1], notoriously declared "No figures will be found in this work."

Mécanique Analytique is now available in an English translation (the fourth edition of 1811), but despite the encomium offered by Dugas to the "perfection of its mathematical language", it is a mathematical language now quite foreign to us. So we will take advantage of modern notation on differentiable manifolds, expressly designed to elucidate the implicit conventions of this classical language, which mathematicians of the time grasped intuitively (or just mimicked).

The mathematical arena for analytical mechanics. We are going to be working with a differentiable manifold M together with the tangent bundle TM, with $M_{\mathscr{p}}$ the tangent space at $\mathscr{p} \in M$ (we use something like $\mathscr{p}$ for points in M, reserving p for later use, cf. page 445, and especially in Part IV.) We will use the notation and conventions of DG, especially Volume 1. In keeping with the conventions of mechanics books, we will use q for a typical coordinate system on M, though writing it, as usual in differential geometry, with superscripts $(q^1, \dots, q^n)$. Recall (DG, Chap. 3) that if $\pi : TM \to M$ is the projection, every tangent vector $\mathbf{v} \in M_{\mathscr{p}}$ at a point $\mathscr{p} \in M$ can be written uniquely as $\mathbf{v} = \displaystyle\sum_{i=1}^{n} a^i \left.\frac{\partial}{\partial q^i}\right|_{\mathscr{p}}$, and if we define $\dot{q}^i(\mathbf{v}) = a^i$, then on TM we have the coordinate system $(q^1 \circ \pi, \dots, q^n \circ \pi, \dot{q}^1, \dots, \dot{q}^n)$, often written simply as $(q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n)$ for convenience. And in general, for a function f on M we often use f as a shorthand for the map $f \circ \pi$ on TM.

Also recall that for a smooth curve $c\colon \mathbb{R} \to M$, we have the tangent vector $c'(t) \in M_{c(t)}$ for each t, so that c' is a curve $c'\colon \mathbb{R} \to TM$.

Since $(q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n)$ is a coordinate system on TM, for any differentiable function L on TM the expressions $\dfrac{\partial L}{\partial q^i}$ and $\dfrac{\partial L}{\partial \dot{q}^i}$ make sense; formally, we just write down a formula for L in terms of the q^i and $\dot{q}^i$, and then simply regard them as new "variables".

Specialized considerations for analytical mechanics. Suppose M is a submanifold of an N-dimensional manifold with the coordinate system $(x^1, \dots, x^N)$. The N-dimensional manifold might be M itself, so that we are comparing two coordinate systems on M, but the case of most interest will actually be $M \subset \mathbb{R}^N$.

For points p in the domain of a coordinate system q on M, we have

$$x^\alpha(p) = X^\alpha\big(q^1(p), \dots, q^n(p)\big), \qquad \alpha = 1, \dots, N$$

for certain smooth functions $X^\alpha\colon \mathbb{R}^n \to \mathbb{R}$, or simply

$$x^\alpha = X^\alpha \circ (q^1, \dots, q^n).$$

On TM we also have the functions $\dot{x}^\alpha$ and $\dot{q}^1, \dots, \dot{q}^n$, related by a similar equation, involving other functions, say

$$\dot{x}^\alpha = \dot{X}^\alpha \circ (q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n).$$

To figure out exactly what $\dot{X}^\alpha$ is, we consider a curve c in M with $c(0) = p$ and let $\mathbf{v} = c'(0) \in M_p$. Differentiating the equation

$$x^\alpha(c(t)) = X^\alpha\big(q^1(c(t)), \dots, q^n(c(t))\big)$$

we obtain, using D_i as the partial with respect to the i^{th} argument,

$$(x^\alpha \circ c)'(t) = \sum_{i=1}^{n} D_i X^\alpha\big(q^1(c(t)), \dots, q^n(c(t))\big) \cdot (q^i \circ c)'(t),$$

and evaluating at $t = 0$ gives

$$\dot{x}^\alpha(\mathbf{v}) = \sum_{i=1}^{n} D_i X^\alpha\big(q^1(p), \dots, q^n(p)\big) \cdot \dot{q}^i(\mathbf{v}),$$

which means that, shorn of extraneous symbols, we have

$$\dot{X}^\alpha(a^1, \dots, a^n, b^1, \dots, b^n) = \sum_{i=1}^{n} D_i X^\alpha(a^1, \dots, a^n) \cdot b^i.$$

In particular, we have

$$D_{n+i} \dot{X}^\alpha(a^1, \dots, a^n, b^1, \dots, b^n) = D_i X^\alpha(a^1, \dots, a^n).$$

As indicated previously, we will allow q^i to stand for $q^i \circ \pi$ on TM, and similarly for x^α. Remembering just how the symbols $\dfrac{\partial}{\partial q^i}$ and $\dfrac{\partial}{\partial \dot{q}^i}$ are defined, we see from the final equation on the previous page that on TM we now have the meaningful, and true, equations[1]

(a)
$$\frac{\partial \dot{x}^\alpha}{\partial \dot{q}^i} = \frac{\partial x^\alpha}{\partial q^i},$$

and we also easily derive

(b)
$$\frac{\partial \dot{x}^\alpha}{\partial q^i} = \sum_{j=1}^{n} \frac{\partial^2 x^\alpha}{\partial q^i \partial q^j} \dot{q}^j.$$

Lagrange's equations. Now let us recall, from page 210, d'Alembert's Principle for Constraints:

> If the constraints on a system confine the system to a configuration space M, and are perpendicular to M, then for all t the motions of the system under the external forces $\mathbf{F}$ satisfy
>
> (∗) $$\langle \mathbf{F}(c(t)) - m\mathbf{c}''(t),\ \mathbf{v}\rangle = 0 \qquad \text{for all } \mathbf{v} \in M_{c(t)},$$
>
> where $\langle\ ,\ \rangle$ denotes the usual inner product on $\mathbb{R}^N$.

In Chapter 6, we struggled to use this principle by expressing all vectors in terms of a set of coordinates for M. Now we will take a totally different approach, writing (∗) directly in terms of real-valued functions in an arbitrary coordinate system q on M.

Let $x^1, \dots x^N$ be the standard coordinates in $\mathbb{R}^N$, where $N = 3K$ for a system of K particles, and the x^α are naturally grouped in triplets, though it is often easiest to think of a curve c in $\mathbb{R}^N$ rather than a collection of K curves; for $m\mathbf{c}$, which denotes $(m_1c_1, \dots, m_Nc_N)$, the m_α are naturally equal in triplets.

Then d'Alembert's principle can be written as

$$\begin{aligned} 0 &= \sum_{\alpha=1}^{N} \Big(\mathbf{F}_\alpha(c(t)) - m_\alpha c_\alpha{}''(t)\Big) \cdot dx^\alpha(c(t)) && \text{on } M_{c(t)} \\ &= \sum_{\substack{\alpha=1,\dots,N \\ i=1,\dots,n}} \Big(\mathbf{F}_\alpha(c(t)) - m_\alpha c_\alpha{}''(t)\Big) \cdot \frac{\partial x^\alpha}{\partial q^i}\, dq^i(c(t)) && \text{on } M_{c(t)}, \end{aligned}$$

[1] Classically obtained from the "equation" $x^\alpha = x^\alpha(q^1, \dots, q^n)$ by first computing that $\dot{x}^\alpha = \sum_{i=1}^n (\partial x^\alpha/\partial q^i)\dot{q}^i$, and then, equally nonchalantly, that $\partial \dot{x}^\alpha/\partial \dot{q}^i = \partial x^\alpha/\partial q^i$.

and since the dq^i are linearly independent on each $M_{c(t)}$, this is equivalent to the set of n individual equations

$$\sum_{\alpha=1}^{N} \Big(\mathbf{F}_\alpha(c(t)) - m_\alpha c_\alpha{}''(t)\Big) \cdot \frac{\partial x^\alpha}{\partial q^i}(c(t)) = 0, \qquad i = 1, \dots, n.$$

For the case of conservative forces, where $\mathbf{F}_\alpha = -\partial V/\partial x^\alpha$, so that

$$\sum_{\alpha=1}^{N} \mathbf{F}_\alpha \frac{\partial x^\alpha}{\partial q^i} = -\sum_{\alpha=1}^{N} \frac{\partial V}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial q^i} = -\frac{\partial V}{\partial q^i},$$

we can thus write

$$(**) \qquad \sum_{\alpha=1}^{N} m_\alpha c_\alpha{}''(t)\frac{\partial x^\alpha}{\partial q^i}(c(t)) = -\frac{\partial V}{\partial q^i}(c(t)).$$

Now comes the fun part.

$$\begin{aligned}
c_\alpha{}''(t)&\frac{\partial x^\alpha}{\partial q^i}(c(t)) \\
&= \frac{d}{dt}\left(c_\alpha{}'(t)\frac{\partial x^\alpha}{\partial q^i}(c(t))\right) - c_\alpha{}'(t)\frac{d}{dt}\left(\frac{\partial x^\alpha}{\partial q^i}(c(t))\right)
\end{aligned}$$

which, using (a) on the first term,

$$\begin{aligned}
&= \frac{d}{dt}\left(c_\alpha{}'(t)\frac{\partial \dot{x}^\alpha}{\partial \dot{q}^i}(c'(t))\right) - c_\alpha{}'(t)\sum_{j=1}^{n}\frac{\partial^2 x^\alpha}{\partial q^i \partial q^j}(c(t))(q^j \circ c)'(t) \\
&= \frac{d}{dt}\left(\dot{x}^\alpha(c'(t))\frac{\partial \dot{x}^\alpha}{\partial \dot{q}^i}(c'(t))\right) - \dot{x}^\alpha(c'(t))\sum_{j=1}^{n}\frac{\partial^2 x^\alpha}{\partial q^i \partial q^j}(c(t))\dot{q}^j(c'(t)) \\
&= \frac{d}{dt}\left(\dot{x}^\alpha(c'(t))\frac{\partial \dot{x}^\alpha}{\partial \dot{q}^i}(c'(t))\right) - \dot{x}^\alpha(c'(t))\frac{\partial \dot{x}^\alpha}{\partial q^i}(c'(t)) \qquad \text{by (b)} \\
&= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}^i}\left(\tfrac{1}{2}[\dot{x}^\alpha(c'(t))]^2\right)\right) - \frac{\partial}{\partial q^i}\left(\tfrac{1}{2}[\dot{x}^\alpha(c'(t))]^2\right).
\end{aligned}$$

When we add up these equations for all α, equation $(**)$ becomes

$$(***) \qquad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^i}(c'(t))\right) - \frac{\partial T}{\partial q^i}(c'(t)) = -\frac{\partial V}{\partial q^i}(c(t)),$$

where T is the kinetic energy,

$$T(c'(t)) = \frac{1}{2}\sum_{\alpha=1}^{N} m_\alpha[\dot{x}^\alpha(c'(t))]^2, \qquad \text{or simply } T(\mathbf{v}) = \frac{1}{2}\sum_{\alpha=1}^{N} m_\alpha[\dot{x}^\alpha(\mathbf{v})]^2.$$

In the standard "condensed notation" of physics, where the curve c is suppressed, we can write (∗∗∗) as

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^i}\right) - \frac{\partial(T-V)}{\partial q^i} = 0.$$

Finally, since V on TM is really $V \circ \pi$ for V on M, we have $\partial V/\partial \dot{q}^i = 0$, so we can write

$$\frac{d}{dt}\left(\frac{\partial(T-V)}{\partial \dot{q}^i}\right) - \frac{\partial(T-V)}{\partial q^i} = 0.$$

Introducing the *Lagrangian*

$$L = T - V,$$

we thus have[1]

$$\boxed{\text{Lagrange's equations:} \qquad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0}$$

Using Lagrange's equations. To get an idea of what these convoluted calculations come down to, we first consider a simple case that has nothing to do with constraints, instead illustrating the usefulness of the fact that our equations can be written in terms of any convenient coordinate system. We consider motion in a plane under a central force, with the obvious choice of polar coordinates (r, θ). As usual, we also allow r and θ to be used as abbreviations for $r \circ c$ and $\theta \circ c$ [as a consequence, $\dot{r}$ also denotes $(r \circ c)'$, which under the same conventions we would denote by r']. As noted on page 120, we have

$$T(r, \theta, \dot{r}, \dot{\theta}) = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)$$

[in our current notation, we could deduce this by computing

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \quad \Longrightarrow \quad \begin{aligned} \dot{x} &= \dot{r}\cos\theta - r\sin\theta\cdot\dot{\theta} \\ \dot{y} &= \dot{r}\sin\theta + r\cos\theta\cdot\dot{\theta}, \end{aligned}$$

and substituting into the formula $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$ for standard coordinates], while we wrote the potential V as

$$V(r, \theta) = F(r),$$

for $F' = f$, where $-f(r)$ is the force at distance r from the origin. Thus the Lagrangian L is

$$L = T - V = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - F(r).$$

[1] For non-conservative forces, we have $\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}^i}\right) - \frac{\partial T}{\partial q^i} = Q_i$ for the "generalized forces" $Q_i = \sum_{\alpha=1}^{N} \mathbf{F}_\alpha \cdot \frac{\partial x^\alpha}{\partial q^i}$.

Now for

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - F(r),$$

Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0, \qquad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0,$$

become

$$\frac{d}{dt}(m\dot{r}) - \big(mr\dot{\theta}^2 - f(r)\big) = 0, \qquad \frac{d}{dt}(mr^2\dot{\theta}) = 0$$

$$\Downarrow$$

$$m(\ddot{r} - r\dot{\theta}^2) + f(r) = 0.$$

[The simplest approach to these calculations, and really the whole point of the procedure, is just to forget about what anything means, and simply regard $r, \dot{r}, \theta, \dot{\theta}$ as "independent variables" and work formally—until we have to take the derivative d/dt, at which point it behooves us to remember that $\dot{r}$ means $(r \circ c)'$, and $\ddot{r}$ means $(r \circ c)''$.]

We've left the second equation undifferentiated, because we immediately get

$$r^2\dot{\theta} = h \quad \text{for a constant } h,$$

the basic equation at the bottom of page 120, which we previously had obtained as a separate step in the analysis. When we then substitute this into the first equation we obtain

$$\ddot{r} = -\frac{f(r)}{m} + \frac{h^2}{r^3},$$

which is equation (B) on page 121; to obtain equation (A) on that page, from which we originally derived (B), we would have to integrate once.

This simple example already illustrates the basic attractive feature of this approach. We just write down the equations, manipulate them formally, and see what they end up saying, thereby obtaining a set of second order equations that will determine the functions (the kinetic energy T plays such a central role precisely because its *first* derivatives give us the second derivatives of the q^i).

One downside is the fact that the second order equations obtained may be more difficult to work with than those obtained in a more elementary way, as illustrated by the fact that we obtained (B) on page 121 instead of (A)—and, similarly, it would have been foolish to expand out $\dfrac{d}{dt}(r^2\dot{\theta}) = 0$.

In addition, the geometric significance of equations like $d/dt\big(r^2\dot{\theta}\big) = 0$ may by somewhat obscured. But then again, we don't need the geometric arguments to obtain them. Instead, we have an automatic way of obtaining such geometric relations. We ended up with the equation $d/dt\big(r^2\dot{\theta}\big) = 0$ precisely because L involves only $\dot{\theta}$, not θ itself. In general, if q^i does not appear in L, but only $\dot{q}^i$, the coordinate q^i is called *ignorable* or *cyclic* [like θ in polar coordinates]. In this case, the Lagrange equation will be $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^i}\right) = 0$ and hence $\dfrac{\partial L}{\partial \dot{q}^i} = \text{constant}$.

In particular, consider the heavy top of Chapter 9, with kinetic energy T given by the second equation on page 346, except that now we also expand out ω_3:

$$T = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + \frac{I_3}{2}(\dot{\phi}\cos\theta + \dot{\psi})^2, \qquad \text{while } V = Mgl\cos\theta.$$

We note that both ψ and ϕ are cyclic, so that along a solution c

$$\frac{\partial L}{\partial \dot{\psi}} \text{ and } \frac{\partial L}{\partial \dot{\phi}} \quad \text{are constant, i.e.,} \quad \frac{\partial L}{\partial \dot{\psi}}(c'(t)) \text{ and } \frac{\partial L}{\partial \dot{\phi}}(c'(t)) \quad \text{are constant.}$$

The first gives equation (B) on page 346, the second equation (C), fulfilling our promise of obtaining these equations without worrying about what they mean.

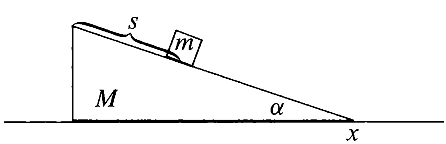
When q^i are standard rectangular coordinates, we have

$$\frac{\partial L}{\partial \dot{q}} = \frac{\partial T}{\partial \dot{q}^i} = m_i \dot{q}^i, \qquad \text{a component of the momentum } m\mathbf{v},$$

and in general $p_i = \partial L/\partial \dot{q}^i$ is called the (*generalized*) *conjugate momentum* to q^i (so for the heavy top, we've seen that the conjugate momenta p_ϕ and p_ψ are constants). Classically, momentum $m\mathbf{v}$ was written as $\mathbf{p}$, which is what suggested the notation p_i, and the accompanying q^i for the coordinates.

The fact that Lagrange's equations can be written using any convenient coordinate system brings up an interesting point. Although we were led to these equations from the perspective of mechanics, one can consider an arbitrary "Lagrangian" function $L\colon TM \to \mathbb{R}$ and the Lagrange equations for curves $c : \mathbb{R} \to M$, and it turns out that in this general case the equations are again "invariant"—if they hold in one coordinate system, they hold in any other. We will prove this by a method that avoids the unpleasantness of a messy computation in Chapter 13, and in a much better way in Part IV.

Constraint problems. As an illustration of the use of Lagrange's equations for a constraint problem, we consider the example from Chapter 6 of a block of mass m sliding on a wedge of mass M, as in the figure below (*pace* Lagrange).



We use the same obvious coordinate system (x, s) as before, where x is the position of one end of the wedge, and s the distance from the top of the wedge to (the center of) the block. We clearly have

$$V = \text{constant} + mg(S - s)\sin\alpha,$$

where S is the total length of the slanted side of the wedge (the constant, included because of the additional height of the center of the block above the wedge, is of course irrelevant, as indeed is S). The contribution of the wedge to the kinetic energy T is clearly $\frac{1}{2}M\dot{x}^2$. The one thing needing a bit of thought (and maybe even the figure!) is the contribution of the block to the kinetic energy T. The velocity of the block is given by

$$\mathbf{v} = (\dot{x} + \dot{s}\cos\alpha,\ -\dot{s}\sin\alpha),$$

so we find that

$$T = \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m(\dot{x}^2 + 2\dot{x}\dot{s}\cos\alpha + \dot{s}^2).$$

Then the non-zero partials of interest are

$$\frac{\partial T}{\partial \dot{x}} = (M+m)\dot{x} + m\dot{s}\cos\alpha$$

$$\frac{\partial T}{\partial \dot{s}} = m\dot{x}\cos\alpha + m\dot{s} \qquad\qquad \frac{\partial V}{\partial s} = -mg\sin\alpha,$$

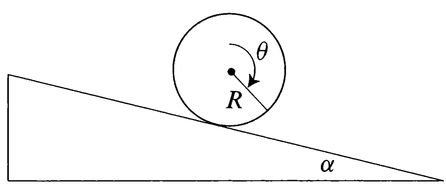
and Lagrange's equations become

$$\frac{d}{dt}\big((M+m)\dot{x} + m\dot{s}\cos\alpha\big) = 0$$

$$\frac{d}{dt}\big(m\dot{x}\cos\alpha + m\dot{s}\big) - mg\sin\alpha = 0,$$

immediately giving the two equations that we obtained at the top of page 220. Although one can easily correlate the steps in this analysis with those used on page 219, the solution via Lagrange's equations can be carried out in a much more systematic and straightforward way.

An even more important simplification is illustrated by another example from Chapter 6, a wheel of radius R and mass M rolling down a (stationary) inclined plane, with the coordinates s and θ. Strictly speaking, we should regard this as a problem in $\mathbb{R}^N$ for some very large N, with rolling being one constraint, and



the fact that the wheel is a rigid body the result of a multitude of constraints. In Chapter 6 we handled this situation, perhaps somewhat mysteriously, by saying that the mass m of the wheel is associated with s, while the moment of inertia I of the wheel should be associated with the angle θ by which the wheel has rotated. In the context of Lagrange's equations, we simply note that the kinetic energy T of the wheel is the sum of its translational and rotational energy, and if I is the moment of inertia of the wheel about its center, then Problem 5-7 shows that

$$T = \tfrac{1}{2}m\dot{s}^2 + \tfrac{1}{2}I\dot{\theta}^2$$

(see Problem 2 for the more general case where the disc is moving on a plane, even including the more general situation where it is not necessarily upright).

Our rolling constraint gives us $\theta = s/R$ so that

$$T = \tfrac{1}{2}m\left(\dot{s}^2 + \frac{I\dot{s}^2}{mR^2}\right)$$

and V is as before, giving the single Lagrange equation

$$\frac{d}{dt}\left(m\dot{s}\left(1 + \frac{I}{mR^2}\right)\right) = mg\sin\alpha \qquad \Longrightarrow \qquad \ddot{s} = \frac{g\sin\alpha}{1 + \dfrac{I}{R^2m}},$$

the same equation we obtained previously.

Admittedly, this particular example is hardly any different than the solution in Problem 6-8 (b). Addendum A and Problem 2 give more interesting uses.

Conservation of energy; action. The formulas for T in the previous examples,

$$\begin{aligned}
T &= \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) \\
T &= \frac{I_1}{2}(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{I_3}{2}(\dot{\phi}\cos\theta + \dot{\psi})^2 \\
T &= \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m(\dot{x}^2 + 2\dot{x}\dot{s}\cos\alpha + \dot{s}^2) \\
T &= \tfrac{1}{2}m\dot{s}^2 + \tfrac{1}{2}I\dot{\theta}^2
\end{aligned}$$

are all homogeneous quadratic functions of the $\dot{q}^i$, that is, if all the $\dot{q}^i$ are multiplied by k, then T is multiplied by k^2.

It is easy to see that this will always be the case, since the third from last formula on page 440 can be written as

$$\dot{x}^\alpha(\mathbf{v}) = \sum_{i=1}^{n} a_{\alpha i}\dot{q}^i(\mathbf{v})$$

for certain functions $a_{\alpha i}$, so each $[\dot{x}^\alpha(\mathbf{v})]^2$ is a homogeneous quadratic function of the $\dot{q}^i$. And this turns out to have an important consequence.

As a matter of notation, for a curve c in M we should write $L(c'(t))$, since L is a function on TM, with the function V really standing for $V \circ \pi$. However, it is often helpful to write $L(c(t), c'(t))$ when we are dealing with a coordinate system, or even $L(c(t))$ when we considering partial derivatives with respect to the q^i, and we will allow ourselves to slip back and forth between these expressions without comment.

Since we can write

$$\frac{d}{dt}L(c'(t)) = \sum_{i=1}^{n} \frac{\partial L}{\partial q^i}(c(t)) \cdot \dot{q}^i(c'(t)) + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot (q^i \circ c)''(t),$$

we see that if c satisfies Lagrange's equations, then

$$\begin{aligned}\frac{d}{dt}L(c'(t)) &= \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(c'(t))\right) \cdot \dot{q}^i(c'(t)) + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot (q^i \circ c)''(t) \\ &= \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t))\right),\end{aligned}$$

which implies that

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t)) - L = E \qquad \text{for a constant } E.$$

Since V does not depend on the $\dot{q}^i$, we have

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t)) = \sum_{i=1}^{n} \frac{\partial T}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t)),$$

and since T is a homogeneous quadratic function of the $\dot{q}^i$, Euler's theorem on homogeneous functions[1] says that

$$\sum_{i=1}^{n} \frac{\partial T}{\partial \dot{q}^i} \cdot \dot{q}^i = 2T,$$

so that

$$E = 2T - L = T + V,$$

the usual expression for energy.

---

[1] If $F(tx) = t^k F(x)$ for $x \in \mathbb{R}^n$, then $\sum_i x_i D_i F(x) = kF(x)$. *Proof:* Take the derivative with respect to t, and set $t = 1$.

The first part of our expression for E, the term

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t)),$$

is called the **action** of c (a name whose origin will be explained in Chapter 15), and merits further consideration in the next section.

Time-dependent Lagrangians. In some cases we have to consider Lagrangians $L\colon TM \times \mathbb{R} \to \mathbb{R}$ that depend explicitly on the time t. An obvious example would arise if V itself depended on t, reflecting an outside force, like the forced oscillations of Chapter 8.

Another common example occurs whenever we have time-dependent constraints, as in the case of a bead sliding on a rotating wire, which we considered in Chapter 6, with gravity ignored (or equivalently, with the wire rotating in a horizontal plane).

In the case of time-dependent constraints, the equations that we started with on page 440 need to be written as

$$\begin{aligned} x^\alpha &= X^\alpha \circ (q^1, \dots, q^n, t), \\ \dot{x}^\alpha &= \dot{X}^\alpha \circ (q^1, \dots, q^n, \dot{q}^1, \dots \dot{q}^n, t), \end{aligned}$$

and the first gives

$$(x^\alpha \circ c)'(t) = \text{previous terms} + \frac{\partial X^\alpha}{\partial t}(\dots).$$

This gives an extra term involving $\partial X^\alpha / \partial t$ in the formula for $\dot{X}^\alpha$, but it ends up having no effect on our equation $\partial \dot{x}^a / \partial \dot{q}^i = \partial x^\alpha / \partial q^i$, or any subsequent equations in our derivation of Lagrange's equations.

In the case of the bead on a rotating wire, the calculations of T on page 443 still hold, giving

$$T = \tfrac{1}{2} m(\dot{r}^2 + r^2 \theta'^2),$$

except that now there is the single coordinate r, while θ is a given function describing how the wire rotates. In this problem we have $V = 0$, and just the one Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{r}}\right) - \frac{\partial T}{\partial r} = 0 \qquad \Longrightarrow \qquad m\ddot{r} = r\theta'^2.$$

Although Lagrange's equations for time-dependent Lagrangians remain the same, the situation for conservation of energy is quite different. The first equation on page 448 now has to be written with an extra term,

$$\frac{d}{dt}L(c'(t)) = (\dots\dots\dots) + \frac{\partial L}{\partial t}(c'(t)),$$

so we cannot expect conservation of energy to hold, which is not surprising for a bead of non-negligible mass m on a rotating wire, since some extra source of energy must be supplied to rotate the wire.

However, that is only half the story. Consider a wire that is being rotated with constant angular velocity, $\dot{\theta} = \omega$ for a constant ω, in which case the Lagrangian actually doesn't depend explicitly on t. Now we get

$$m\ddot{r} = r\omega^2,$$

and since

$$T + V = T = \tfrac{1}{2}m(\dot{r}^2 + r^2\omega^2)$$

we find that

$$\frac{d}{dt}(T + V) = m(\dot{r}\ddot{r} + r\dot{r}\omega^2) = 2mr\dot{r}\omega^2,$$

which is non-zero for $\omega \neq 0$, as we would expect even in this special case.

This doesn't contradict the results of the previous section, because in cases involving time-dependent constraints, when we write the kinetic energy $T = \frac{1}{2}\sum_\alpha m_\alpha \dot{x}^\alpha$ in terms of the q^i and $\dot{q}^i$ there will be extra terms involving $\partial X^\alpha/\partial t$, even if the Lagrangian doesn't happen to depend explicitly on time. For example, in the present case, although the kinetic energy $T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)$ for a single particle is a homogeneous quadratic form in $\dot{\theta}$ and $\dot{r}$, the T for the bead on a rotating wire,

$$T = \tfrac{1}{2}m(\dot{r}^2 + r^2\theta'^2) \qquad \text{for a given function } \theta,$$

is not a homogeneous quadratic form in the single coordinate $\dot{r}$. Thus, our identification of the constant E with $T + V$ cannot be carried through.

So the result of the previous section can best be expressed by saying that for Lagrangians that do not depend explicitly on t, the action minus L,

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t)) \;-\; L(c'(t)),$$

is always a constant; however, this combination will usually not be the same as the energy $T + V$ when our constraints depend on time.

Lagrange multipliers. Not surprisingly, Lagrange multipliers can also be used with Lagrange's equations, now using d'Alembert's Principle for Differential Constraints, from page 233:

> If the constraints on a system require the tangent vector of the motion to lie in the subspace $\ker(\omega_1) \cap \cdots \cap \ker(\omega_L)$, then there are Lagrange multipliers $\lambda_1, \ldots, \lambda_L$ such that the motions of the system under the external forces $\mathbf{F}$ satisfy
> $$\langle \mathbf{F} - m\mathbf{c}'', \mathbf{v} \rangle = \lambda_1 \omega_1(\mathbf{v}) + \cdots + \lambda_L \omega_L(\mathbf{v})$$
> for all tangent vectors $\mathbf{v}$ at $\mathbf{c}$.

We can write this as

$$\sum_{\alpha=1}^{N} \Big(\mathbf{F}_\alpha(c(t)) - m_\alpha c_\alpha{}''(t) \Big) \cdot dx^\alpha(c(t)) = \sum_{l=1}^{L} \lambda_l \omega_l(c(t)),$$

while each ω_l, when restricted to TM, can be written as

$$\omega_l = \sum_{j=1}^{n} a_{li}\, dq^i$$

for certain functions a_{li} describing the constraints. So on $M_{c(t)}$ we also have the complicated looking equations

$$\sum_{\alpha=1}^{N} \sum_{i=1}^{n} \Big(\mathbf{F}_\alpha(c(t)) - m_\alpha c_\alpha{}''(t) \Big) \cdot \frac{\partial x^\alpha}{\partial q^i}\, dq^i(c(t)) = \sum_{l=1}^{L} \sum_{i=1}^{n} \lambda_l a_{li}\, dq^i(c(t)),$$

leading to the set of equations

$$\sum_{\alpha=1}^{N} \Big(\mathbf{F}_\alpha(c(t)) - m_\alpha c_\alpha{}''(t) \Big) \cdot \frac{\partial x^\alpha}{\partial q^i}(c(t)) = \sum_{l=1}^{L} \lambda_l a_{li}(c(t)), \qquad i = 1, \ldots, n.$$

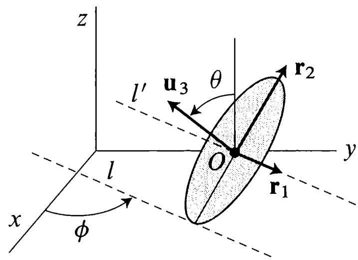
Then the remaining parts of the argument for Lagrange's equations lead to

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \sum_{l=1}^{L} \lambda_l a_{li}.$$

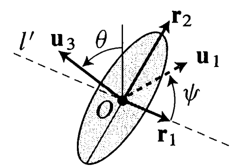
In practice this works out a lot simpler than it may look! Note that the index l simply refers to constraint number l, while the index i refers to the i^{th} coordinate, so, for example, when using coordinates r, θ one would write things like a_{lr} and $a_{l\theta}$ for constraint number l (cf. Problem 2). Remember also that one generally needs to differentiate the relations implied by the constraints, as was done on page 234 for the constraint equations (1) at the top of that page.

ADDENDUM 12A
LAGRANGE'S ROLLING DISC

In contrast to the treatment in Addendum 9A, we will now use Lagrange's equations, and Lagrange multipliers, to obtain equations for the general case of a disc rolling on a plane[1] (you may wish to delay this complicated analysis until after doing the Problems). We will be using the Euler angles as coordinates, as in the figure, which looks very much like the figure of Addendum 9A, except



that we will not be using Euler's equations in any form, and the two vectors now labeled $\mathbf{r}_1$ and $\mathbf{r}_2$ are simply convenient reference axes. Since l' is the line of nodes, when we do consider $\mathbf{u}_1$, we see that the Euler angle ψ measures



the rotation from the horizontal to $\mathbf{u}_1$, so $\dot{\psi}$ is the rate at which the wheel is rotating.

We first want to find the components of the angular velocity vector $\boldsymbol{\omega}$ of the disc with respect to the axes $\mathbf{r}_1, \mathbf{r}_2, \mathbf{u}_3$. The same sort of considerations that we used to get equation (b) at the bottom of page 363 shows that the desired components p, q, and r of $\boldsymbol{\omega}$ are

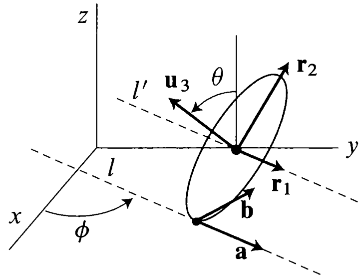
(A) $$p = \dot{\theta}, \qquad q = \dot{\phi}\sin\theta, \qquad r = \dot{\phi}\cos\theta + \dot{\psi}.$$

In addition to θ, ϕ, ψ, we will use the coordinates x, y of O, while the essentially redundant coordinate z of O is $z = a\sin\theta$. Using the moments of inertia of the disc given by Problem 5-6, together with Problem 2 (c) of this chapter, we see that

(B) $$\begin{aligned} T &= \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + a^2\dot{\theta}^2\cos^2\theta) \\ &\qquad + \tfrac{1}{4}m\dot{\theta}^2 + \tfrac{1}{4}m\dot{\phi}^2\sin^2\theta + \tfrac{1}{2}ma^2(\dot{\phi}\cos\theta + \dot{\psi})^2 \\ V &= mga\sin\theta. \end{aligned}$$

[1] From Cabannes [1].

Although Lagrange's equations are supposed to simplify the solution of mechanics problems, a lot depends on choosing good coordinates. In our problem, it will be convenient to express the rolling condition in terms of: the directions of the vector $\mathbf{a}$ along l, parallel to $\mathbf{r}_1$, making an angle of ϕ with respect to



the x-axis; the vector $\mathbf{b}$ in the (x, y)-plane perpendicular to it, making an angle of $\phi + \pi/2$ with the x-axis (and pointing in the direction of the horizontal projection of $\mathbf{r}_2$); and the vertical direction.

Then $\boldsymbol{\omega}$ and the vector $-a\mathbf{r}_2$ from the center O to the contact point have components

$$(\dot\theta,\ -\dot\psi\sin\theta,\ \dot\psi\cos\theta + \dot\phi) \qquad \text{and} \qquad (0,\ -a\cos\theta,\ -a\sin\theta)$$

so that the components of $\boldsymbol{\omega} \times (-a\mathbf{r}_2)$ are

$$(a(\dot\phi\cos\theta + \dot\psi),\ \ a\dot\theta\sin\theta,\ \ -a\dot\theta\cos\theta),$$

and the rolling conditions become (compare pages 240 and 363)

$$\text{(C)} \qquad \begin{aligned} \dot x\cos\phi + \dot y\sin\phi + a(\dot\phi\cos\theta + \dot\psi) &= 0 \\ -\dot x\sin\phi + \dot y\cos\phi + a\dot\theta\sin\theta &= 0, \end{aligned}$$

leading us to consider the 1-forms

$$\begin{aligned} &\lambda_1(\cos\phi\, dx + \sin\phi\, dy + a\cos\theta\, d\phi + a\, d\psi), \\ &\lambda_2(-\sin\phi\, dx + \cos\phi\, dy + a\sin\theta\, d\theta). \end{aligned}$$

Ignoring the irrelevant coordinate z, the equations on page 451 become

(1) [for x] $\quad m\ddot x = \lambda_1\cos\phi - \lambda_2\sin\phi$

(2) [for y] $\quad m\ddot y = \lambda_1\sin\phi + \lambda_2\cos\phi$

(3) [for θ]
$$\begin{aligned} &ma^2(\dot\theta\cos^2\theta)' + \tfrac{1}{2}ma^2\ddot\theta + ma^2\dot\theta^2\sin\theta\cos\theta \\ &\quad - \tfrac{1}{2}ma^2\dot\phi^2\sin\theta\cos\theta + ma^2\dot\phi(\dot\phi\cos\theta + \dot\psi)\sin\theta \\ &\quad = \lambda_2 a\sin\theta - mga\cos\theta \end{aligned}$$

(4) [for ϕ] $\quad \tfrac{1}{2}ma^2(\dot\phi\sin^2\theta)' + ma^2[\cos\theta(\dot\phi\cos\theta + \dot\psi)]' = \lambda_1 a\cos\theta$

(5) [for ψ] $\quad ma^2(\dot\phi\cos\theta + \dot\psi)' = \lambda_1 a.$

Equations (1) and (2) give

$$\lambda_1 = \tfrac{1}{2}m(\ddot{x}\cos\phi + \ddot{y}\sin\phi)$$
$$\lambda_2 = \tfrac{1}{2}m(-\ddot{x}\sin\phi + \ddot{y}\cos\phi),$$

and by using equations (C) and their derivatives, we find that

$$\lambda_1 = ma\big[\dot\theta\dot\phi\sin\theta - (\dot\phi\cos\theta + \dot\psi)'\big]$$
$$\lambda_2 = -ma\big[\dot\phi(\dot\phi\cos\theta + \dot\psi) + (\dot\theta\sin\theta)'\big].$$

Writing things in terms of the abbreviations p, q, and r introduced in equation (A), we find from (5) that $2\dot{r} - pq = 0$, and, making use of this, we obtain the three equations

$$\text{(D)}\qquad \begin{aligned} \dot q + p(q\cot\theta - 2r) &= 0 \\ 2\dot r - pq &= 0 \\ 3\dot p + q(4r - q\cot\theta) &= -2\frac{g}{a}\cos\theta, \end{aligned}$$

which are really equations in only three unknown functions, since $p = \dot\theta$.

A disc rolling vertically along a straight line,

$$\theta = \pi/2, \qquad \phi = \text{constant}, \qquad \psi = \text{constant},$$

has $p = q = r = 0$ and $\cos\theta = 0$, so the equations are satisfied.

For a solution with the disc rolling along a circle, inclined at a fixed angle,

$$\theta = \text{constant}, \qquad \phi' = \text{constant}, \qquad \psi = \text{constant},$$

with $p = 0$, $q = \text{constant}$, and $r = \text{constant}$, the first two equations are automatic, while the third equation,

$$q(4r - q\cot\theta) = -2\frac{g}{a}\cos\theta$$

now gives the condition connecting the angle at which the disc is inclined and the centripetal force that must be exerted in order for the disc to move in a circle.

Our three equations are actually amenable to a bit of manipulation. Dividing the first two equations of (D) by $p = d\theta/dt$, we obtain

$$\frac{dq}{d\theta} + q\cot\theta - 2r = 0, \qquad 2\frac{dr}{d\theta} - q = 0,$$

leading to the single equation

$$\frac{d^2r}{d\theta^2} + \cot\theta \cdot \frac{dr}{d\theta} - r = 0,$$

from which we could also work backwards.

It turns out that this equation can, in a sense, be solved. The substitution $s = \cos^2\theta$ changes it into the "hypergeometric equation"

$$s(s-1)\frac{d^2r}{ds^2} + \left(\frac{1}{2} - \frac{3}{2}s\right)\frac{dr}{ds} - \frac{1}{4}r = 0,$$

one solution of which is given by the infinite sum

$$r = \sum_{n=0}^{\infty} a_n s^n, \qquad \text{with} \qquad \frac{a_n}{a_{n-1}} = \frac{4n^2 - 6n + 3}{2n(2n-1)}.$$

PROBLEMS

1. Consider a pendulum for which the bob of mass m is suspended by a spring with spring constant k, unstretched length l, and length $l+x(t)$ at time t, giving a radial force of $-kx$. Find the Lagrangian, obtain the equations

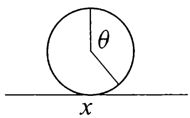
$$m\ddot{x} = m(l+x)\dot{\theta}^2 + \underline{mg\cos\theta} - kx$$

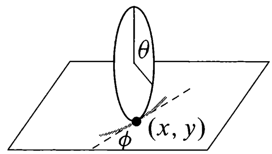
$$m(l+x)\ddot{\theta} + \underline{2m\dot{x}\dot{\theta}} = -mg\sin\theta,$$

and interpret them, and the underlined expressions, in terms of the rotating coordinate system determined by the spring.

$L=m/2[\dot{x}^2+(l+x)^2\dot{\theta}^2]+mg(l+x)\cos\theta-kx^2/2$. The first equation is the radial component of the equation $\mathbf{F}=m\mathbf{a}$ with the centrifugal force added in, and the second is the tangential component with the Coriolis force added in.

2. (a) On page 447 we used the fact that for a disc moving along a straight line (whether rolling or not), the kinetic energy T is given by $T = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\dot{\theta}^2$.





For the case of a disc moving upright on a plane, with I_ϕ the moment of inertia about the vertical axis, the obvious guess for the formula for T is

$$T = \tfrac{1}{2}(m\dot{x}^2 + m\dot{y}^2 + I\dot{\theta}^2 + I_\phi\dot{\phi}^2),$$

corresponding to the analysis on page 234. Prove this formula, again using the formula for T_{rot} on page 194.

(b) When the disc is rolling, so that we have

$$\begin{aligned} \dot{x} &= R\cos\phi\ \dot{\theta} \\ \dot{y} &= R\sin\phi\ \dot{\theta}, \end{aligned} \qquad \text{expressed by} \qquad \begin{aligned} dx - R\cos\phi\ d\theta &= 0 \\ dy - R\sin\phi\ d\theta &= 0, \end{aligned}$$

we want to use Lagrange multipliers to form

$$\begin{aligned} &\lambda_1(dx - R\cos\phi\ d\theta) \\ &\lambda_2(dy - R\sin\phi\ d\theta); \end{aligned}$$

in the notation on page 451, we have, e.g., $a_{1x} = \lambda_1$ and $a_{2\theta} = -\lambda_2 R\sin\phi$. Show that we now obtain, up to sign, the same equations as those on page 234. (The case of a disc rolling down an inclined plane may be handled in exactly the same manner.)

(c) Extend part (a) to the case where the disc is not necessarily upright, resulting in another term involving the moment of inertia about a diameter.

3. (a) For a pendulum with a bob of mass m on a string of length l, with θ the angle from the vertical, we have

$$L = \tfrac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta.$$

Use Lagrange's equation to obtain the standard pendulum equation

$$l^2\ddot{\theta} + lg\sin\theta = 0.$$

(b) To get the tension, consider the Lagrangian

$$\bar{L} = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos\theta$$

on the whole plane, where r is radial distance from the pivot point, together with the constraint $r = l$, expressed by $dr = 0$ (so $a_{1r} = 1$) and use the equations derived on page 451 to get

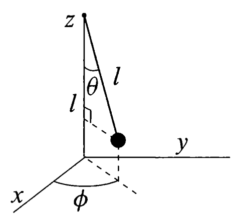
$$m\ddot{r} - mr\dot{\theta}^2 - mg\cos\theta = \lambda,$$

or for $r = l$ (and $\dot{r} = \ddot{r} = 0$) simply

$$-ml\dot{\theta}^2 - mg\cos\theta = \lambda.$$

4. (a) Consider the spherical pendulum of Problem 3-5, with the path parameterized by θ and ϕ. Show that

$$L = \tfrac{1}{2}ml^2(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2) + mgl\cos\theta$$



and conclude that

$$l^2\sin^2\theta\,\dot{\phi} \quad \text{is constant.}$$

Compare with Problem 3-5 (direct use of the coordinates x and y with $V = mg\sqrt{l^2 - x^2 - y^2}$ makes the problem more difficult).
(b) To find the tension of the string of the pendulum, consider the full set of coordinates (r, θ, ϕ), with

$$\bar{L} = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2) + mgr\cos\theta,$$

and use a Lagrange multiplier for the constraint $r = l$ to obtain

$$mg\cos\theta + ml(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) = -\lambda.$$

5. (a) For the sliding particle of Problem 6-11, we have

$$L = \tfrac{1}{2}ml^2\dot{\theta}^2 - mgl\sin\theta.$$

Conclude from the Lagrange equation for θ that

$$l^2\dot{\theta}^2 + gl\cos\theta = 0.$$

(b) Now consider the Lagrangian for the coordinates (r, θ) in the plane,

$$\bar{L} = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgr\cos\theta,$$

with the constraint $r = l$. Obtain

$$m\ddot{r} - mr\dot{\theta}^2 + mg\sin\theta = \lambda$$

so that $\lambda = 0$ precisely when $l\dot{\theta}^2 = g\sin\theta$. As in Problem 6-11, this can be used to determine when the particle leaves the path, though we still need the remainder of the analysis in that problem.

6. (a) For the sliding stick of Problem 6-13, we have, using the value $ml^2/3$ for the moment of inertia of the stick, as mentioned on page 263,

$$L = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}m\dot{y}^2 + \tfrac{1}{6}ml^2\dot{\theta}^2 - mgl\sin\theta.$$

Writing this totally in terms of $\dot{\theta}$, use the Lagrange equation to obtain

$$\ddot{\theta} + \frac{3g}{4l}\cos\theta = 0.$$

(b) For L in the original form, in terms of the coordinates x, y, θ, the conditions that the stick touches the wall and the floor are

$$\begin{aligned} x &= l\cos\theta \\ y &= l\sin\theta, \end{aligned} \qquad \text{or as differential constaints} \qquad \begin{aligned} dx - l\cos\theta\, d\theta &= 0 \\ dy - l\sin\theta\, d\theta &= 0. \end{aligned}$$

Using Lagrange multipliers λ_1 and λ_2 to form

$$\begin{aligned} &\lambda_1(dx - l\cos\theta\, d\theta) \\ &\lambda_2(dy - l\sin\theta\, d\theta), \end{aligned}$$

obtain the equations

$$\begin{aligned} m\ddot{x} &= \lambda_1 \\ m\ddot{y} &= \lambda_2 \\ \tfrac{1}{3}ml^2\ddot{\theta} + mlg\cos\theta &= -l(\lambda_1\cos\theta + \lambda_2\sin\theta) \\ &= ml\dot{\theta}^2, \end{aligned}$$

which can then be combined with our previous equation.

7. A 1-form α on M is, strictly speaking, a function on M whose value $\alpha(p)$ for $p \in M$ is a function on M_p, but we can obviously also think of α simply as a function on TM, which in a coordinate system has the form $\sum_{j=1}^{n} A_j\dot{q}^j$ for certain functions A_j on the domain of the coordinate system.

(a) If $L\colon TM \to \mathbb{R}$ and we let $\overline{L} = L + \alpha + C$ for a constant C, then for any curve c in M we have

$$\frac{d}{dt}\left(\frac{\partial \overline{L}}{\partial \dot{q}^i}(c'(t))\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(c'(t))\right) + \sum_{j=1}^{n} \frac{\partial A_i}{\partial q^j}(c(t))(q^j \circ c)'(t)$$

$$\frac{\partial \overline{L}}{\partial q^i}(c(t)) = \frac{\partial L}{\partial q^i}(c(t)) + \sum_{j=1}^{n} \frac{\partial A_j}{\partial q^i}(c(t))(q^j \circ c)'(t).$$

(b) Conclude that if α is a closed 1-form, $d\alpha = 0$, then for each c the Lagrange equations for the Lagrangians L and $\overline{L}$ are exactly the same.
(c) Conversely, if they are always exactly the same, then α must be closed.

This is expressed, in terminology first introduced for equations of electromagnetism, by saying that $L \mapsto L + \alpha + C$ is a "gauge transformation" of the Lagrangian. Note that saying that the Lagrange equations for L and $\overline{L}$ are exactly the same for all c is not the same as saying that the equations merely have the same solutions. For example, for $TM = T\mathbb{R}^1$, with standard coordinates x and $\dot{x}$, note that the Lagrangians $L(x,\dot{x}) = a\dot{x}^2$ for various constants a all have the same solution curves, even though they lead to different equations.

For a formal statement of the equations being "the same", see Abraham and Marsden [1; §3.5].

(d) For the converse, note that if Lagrange's equations for every curve c are the same for $\overline{L}$ and L, then

$$\frac{d}{dt}\left(\frac{\partial(\overline{L} - L)}{\partial \dot{q}^i}(c'(t))\right) = \frac{\partial(\overline{L} - L)}{\partial q^i}(c(t))$$

for all c. Conclude that we must have separately

$$\frac{d}{dt}\left(\frac{\partial(\bar{L}-L)}{\partial \dot{q}^i}(c'(t))\right) = 0 \qquad \text{and} \qquad \frac{\partial(\bar{L}-L)}{\partial q^i}(c(t)) = 0,$$

and then that

$$\bar{L} - L = \sum_{i=1}^{n} A_i \dot{q}^i + C = \alpha + C,$$

and finally by (c) that $d\alpha = 0$.

Remark. In mechanics books, the usual, virtually incomprehensible, statement is that

$$\bar{L} = L + \frac{d\phi}{dt}$$

for some function ϕ on M. There are two points to unravel here, aside from the omission of the constant C. First, the closed 1-form α is being written as $d\phi$, which it can be locally. Then, the equation that we would write as $\bar{L} = L + d\phi$ is implicitly being applied to a tangent vector $c'(t)$, so that we have

$$\begin{aligned}\bar{L}(c'(t)) &= L(c'(t)) + d\phi(c'(t)) \\ &= L(c'(t)) + \frac{d(\phi \circ c)(t)}{dt},\end{aligned}$$

except that, as usual in physics notation, the c is being suppressed, so that the final term has to be written as $d\phi/dt$! The result is often stated for the general case of Lagrangians $L(c'(t), t)$ that might depend explicitly on t, probably in order to disguise the meaninglessness of writing $d\phi/dt$ for a function ϕ on M.

CHAPTER 13
VARIATIONAL PRINCIPLES

This is like déjà vu all over again.
— Yogi Berra

Lagrange's equations will seem eerily familiar to any one acquainted with differential geometry, because they so closely resemble the Euler equations, from the calculus of variations, which are used to investigate geodesics on Riemannian manifolds (DG, Chap. 9). And, in fact, Lagrange's equations *are* the Euler equations for the Lagrangian $L\colon TM \to \mathbb{R}$.

The Euler equations. As a brief review, we recall that if we have a suitably differentiable function

$$F\colon \mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$$

we can seek, among all functions $f\colon [a,b]\to\mathbb{R}$ with $f(a)=a'$ and $f(b)=b'$, for some fixed a' and b', one which will maximize or minimize the quantity

$$J(f)=\int_a^b F(f(t),f'(t),t)\,dt.$$

We do this by considering a "variation" of f, that is, a function $\alpha\colon(-\varepsilon,\varepsilon)\times[a,b]\to\mathbb{R}$ such that $\alpha(0,t)=f(t)$. The functions $t\mapsto\alpha(u,t)$ are then a family of functions on $(-\varepsilon,\varepsilon)$ which pass through f for $u=0$, and we denote these functions by $\bar\alpha(u)$, so that $\bar\alpha$ is a function from $(-\varepsilon,\varepsilon)$ to the set of functions $f\colon[a,b]\to\mathbb{R}$. If each $\bar\alpha(u)$ satisfies $\bar\alpha(u)(a)=a'$, $\bar\alpha(u)(b)=b'$, in other words, if $\alpha(u,a)=a'$ and $\alpha(u,b)=b'$ for all $u\in(-\varepsilon,\varepsilon)$, then we call α a variation of f keeping endpoints fixed.

For a variation α keeping endpoints fixed we want to compute

$$\left.\frac{dJ(\bar\alpha(u))}{du}\right|_{u=0}=\left.\frac{d}{du}\right|_{u=0}\int_a^b F\left(\alpha(u,t),\frac{\partial\alpha}{\partial t}(u,t),t\right)dt$$

and seek a *critical point* or *extremal* f of J, for which this derivative is 0 for all α.

We first move the differentiation inside the integral sign to get

$$\begin{aligned}&\int_a^b\left[\left.\frac{d}{du}\right|_{u=0}F\left(\alpha(u,t),\frac{\partial\alpha}{\partial t}(u,t),t\right)\right]dt\\ &\qquad=\int_a^b\left[\frac{\partial\alpha}{\partial u}(0,t)\frac{\partial F}{\partial x}(f(t),f'(t),t)+\frac{\partial^2\alpha}{\partial u\partial t}(0,t)\frac{\partial F}{\partial y}(f(t),f'(t),t)\right]dt,\end{aligned}$$

and apply integration by parts to the second term to obtain

$$\left.\frac{dJ(\bar{\alpha}(u))}{du}\right|_{u=0} = \int_a^b \frac{\partial \alpha}{\partial u}(0,t)\left[\frac{\partial F}{\partial x}(f(t), f'(t), t) - \frac{d}{dt}\left(\frac{\partial F}{\partial y}(f(t), f'(t), t)\right)\right] dt + \left.\frac{\partial \alpha}{\partial u}(0,t)\frac{\partial F}{\partial y}(f(t), f'(t), t)\right|_a^b .$$

Since the last term is 0 for variations α keeping endpoints fixed, and since $\partial\alpha/\partial u(0,t)$ can be any function vanishing at a and b we find that f must satisfy "Euler's Equation"

$$\frac{\partial F}{\partial x}(f(t), f'(t), t) - \frac{d}{dt}\left(\frac{\partial F}{\partial y}(f(t), f'(t), t)\right) = 0,$$

where $\partial F/\partial x$ is officially $D_1 F$, the partial derivative of F with respect to the second argument, and $\partial F/\partial y$ is officially $D_2 F$ (and $\partial/\partial t$ would be D_3).

The considerations are easily generalized to $f\colon [a,b] \to \mathbb{R}^n$ with

$$J(f) = \int_a^b F(f(t), f'(t)), t\, dt \qquad \text{for } F\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}.$$

We obtain

$$(*) \qquad \left.\frac{dJ(\bar{\alpha}(u))}{du}\right|_{u=0} = \int_a^b \sum_{i=1}^n \frac{\partial \alpha^i}{\partial u}(0,t)\left[\frac{\partial F}{\partial x^i}(f(t), f'(t), t) - \frac{d}{dt}\left(\frac{\partial F}{\partial y^i}(f(t), f'(t), t)\right)\right] dt + \left.\sum_{i=1}^n \frac{\partial \alpha^i}{\partial u}(0,t)\frac{\partial F}{\partial y^i}(f(t), f'(t), t)\right|_a^b$$

and we find that a critical point f of J must satisfy the n equations

$$\frac{\partial F}{\partial x^i}(f(t), f'(t), t) - \frac{d}{dt}\left(\frac{\partial F}{\partial y^i}(f(t), f'(t), t)\right) = 0,$$

where now $\partial/\partial x^i$ denotes D_i and $\partial/\partial y^i$ denotes D_{n+i}.

A surprisingly significant role will be played by a side-effect of this derivation:

THE BOUNDARY TERM COROLLARY. For any critical point f for J we have

$$\left.\frac{dJ(\bar{\alpha}(u))}{du}\right|_{u=0} = \left.\sum_{i=1}^n \frac{\partial \alpha^i}{\partial u}(0,t)\frac{\partial F}{\partial y^i}(f(t), f'(t), t)\right|_a^b .$$

Hamilton's principle. If $(q^1, \dots, q^n)$ is a coordinate system on M, and we apply these results to the coordinate functions for $L\colon TM \times \mathbb{R} \to \mathbb{R}$, the Euler equations translate exactly into Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial q^i}(c(t), t)\right) - \frac{\partial L}{\partial \dot{q}^i}(c'(t), t) = 0,$$

which are also called the Euler-Lagrange equations, so that the solutions c to Lagrange's equations for $L\colon TM \to \mathbb{R}$ are precisely the critical functions for

$$J(c) = \int_{t_1}^{t_2} L(c(t), c'(t), t)\, dt.$$

The shows immediately that for any $L : TM \times \mathbb{R} \to \mathbb{R}$, if Lagrange's equations for c hold in one coordinate system, then they will also hold in an overlapping one. Of course, this is still a somewhat indirect proof of the "invariance" of these equations. We have not provided a direct interpretation of Lagrange's equations, nor will we be in a position to do so until Part IV. This is not so surprising when one notes that if $L(\mathbf{v}) = T(\mathbf{v}) = \frac{1}{2}\langle \mathbf{v}, \mathbf{v}\rangle^2$ for a Riemannian metric $\langle\ ,\ \rangle$ on M, then $J(c)$ is precisely the "energy" $E(c)$ (DG, pg. 324), whose critical points are, by definition, the geodesics for $\langle\ ,\ \rangle$, and an invariant description of their equations requires the additional machinery of connections, as described, *ad nauseum*, in DG, Vol. 2. Problem 1 investigates these same equations for the more general case where $L = \frac{1}{2}\langle\ ,\ \rangle^2 - V$.

For Lagrangians that arise from mechanics problems, the identification of the Euler equations with Lagrange's equations for L is called

> **Hamilton's principle**. The solutions $c\colon [t_1, t_2] \to M$ for a system with Lagrangian $L = T - V$ are the critical values for
>
> $$J(c) = \int_{t_1}^{t_2} L(c(t), c'(t), t)\, dt.$$

Such so-called variational principles, which may be given an interpretation that is teleological, saying that "Nature" always chooses some sort of optimal path, have often led to philosophical clashes, but modern physicists are interested in them mainly because they often provide a useful path for extending investigations to other areas beyond classical mechanics, like continuum and fluid mechanics, and quantum mechanics (and physicists tell us that quantum mechanics explains how, in a sense, particles do try out all the different paths in order to choose the optimal one).

The name "Hamilton's principle" is actually something of an anachronism, since Hamilton's work, the subject of Part IV, appeared nearly half a century after *Mécanique Analytique* was published. Lagrange stated the principle, but didn't provide a name for it, describing it only as a principle "which I view not as a metaphysical principle but as a simple and general result of the laws of mechanics."

Maupertuis and the Principle of Least Action. Another variational principle, the Principle of Least Action, had been formulated by Maupertuis, quite definitely as a metaphysical principle, even supposedly proving the existence of God, in connection with refraction of light (cf. Chapter 15), and then applied with appropriate modifications to mechanics.

For a Lagrangian $L\colon TM \to \mathbb{R}$ we consider a path $c : [t_1, t_2] \to M$ and the integral of its action (top of page 449),

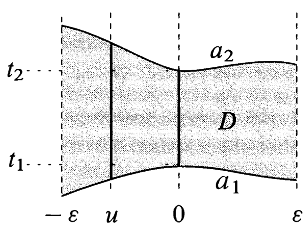
$$\mathcal{A}(c) = \int_{t_1}^{t_2} \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}(c'(t)) \cdot \dot{q}^i(c'(t))\, dt.$$

We want to show that c is a critical point for $\mathcal{A}$ if and only if it satisfies Lagrange's equations, except that the paths and the variations allowed will be quite different from those used in Hamilton's principle.

To being with, we will now consider only Lagrangians L that do not depend explicitly on time.

In addition, we restrict the class of paths by requiring that the energy $E = A - L$ of c has a constant value E_0, and consider only curves whose energy likewise have the constant value E_0.

On the other hand, though we restrict our attention to curves that begin and end at the same points as c, we do not require that they are defined on the same time interval $[t_1, t_2]$. So instead of a variation $\alpha\colon (-\varepsilon, \varepsilon) \times [t_1, t_2] \to M$, we will have α defined on a region $D \subset (-\varepsilon, \varepsilon) \times \mathbb{R}$ bounded by the graphs of functions a_1 and $a_2 > a_1$, where $\bar{\alpha}(u)$ is defined on the interval $[a_1(u), a_2(u)]$, and $a_1(0) = t_1$ and $a_2(0) = t_2$. [We could just as well assume that $a_1(u) = t_1$ for all u, with only $a_2(u)$ varying.]



We still want to have each $\bar{\alpha}(u)$ go from $c(t_1)$ to $c(t_2)$, so that for all u the components $\alpha^i = q^i \circ \alpha$ satisfy

$$\alpha^i(u, a_\nu(u)) = q^i(c(t_\nu)), \qquad \nu = 1, 2,$$

and differentiating with respect to u gives

$$\frac{\partial\alpha^i}{\partial u}(0,t_\nu)+\frac{\partial\alpha^i}{\partial t}(0,t_\nu)\cdot(q^i\circ a_\nu)'(0)=0$$

or

$$\text{(A)}\qquad \frac{\partial\alpha^i}{\partial u}(0,t_\nu)=-\dot q^i(c'(t_\nu))\cdot\dot q^i(a_\nu{}'(0)).$$

Since all the curves $\bar\alpha(u)$ are assumed to have constant energy E_0 and $A = E + L$, we have

$$\mathcal{A}(\bar\alpha(u))=\int_{a_1(u)}^{a_2(u)}[L(c'(t))+E_0]\,dt.$$

When we compute the derivative of $\mathcal{A}(\bar\alpha(u))$ at $u = 0$, the result corresponding to $(*)$ on page 462 has an extra term coming from differentiating the limits on the integral, namely

$$a_2{}'(0)\cdot[L(c'(t_2))+E_0]\;-\;a_1{}'(0)\cdot[L(c'(t_1))+E_0].$$

Since each $L + E$ term can be replaced by the action $A = \sum_i(\partial L/\partial\dot q^i)\dot q^i$, the $\dot q^i$ component of this additional term is

$$\text{(B)}\quad \dot q^i(a_2{}'(0))\cdot\frac{\partial L}{\partial\dot q^i}(c'(t_2))\cdot\dot q^i(c'(t_2))\;-\;\dot q^i(a_1{}'(0))\cdot\frac{\partial L}{\partial\dot q^i}(c'(t_1))\cdot\dot q^i(c'(t_1)).$$

On the other hand, the final boundary term in the result corresponding to $(*)$ will be

$$\text{(C)}\qquad \sum_{i=1}^{n}\frac{\partial\alpha^i}{\partial u}(0,t)\frac{\partial L}{\partial\dot q^i}((c'(t))\Bigg|_{t_1}^{t_2}$$

Using (A), we see that (B) and (C) cancel each other out, so we finally obtain the same expression as before,

$$\int_{t_1}^{t_2}\sum_{i=1}^{n}\frac{\partial\alpha^i}{\partial u}(0,t)\left[\frac{\partial L}{\partial q^i}(c(t))-\frac{d}{dt}\left(\frac{\partial L}{\partial\dot q^i}(c'(t))\right)\right]dt,$$

for the quantity that should be 0.

We wish to conclude, as before, that the terms in brackets must be 0, but a little caution is needed, as we have to choose α's for which the $\bar\alpha(u)$ all have constant energy E_0. In fact, we have to add the extra hypothesis that c is not a critical path for E. The sufficiency of this hypothesis follows from rather general notions concerning infinite dimensional manifolds; Addendum A, already referred to in Problem 6-21, contains a lemma giving the classical argument. **Q.E.D.**

(A very different approach to Maupertuis' Principle of Least Action is mentioned on page 592.)

It was Euler who correctly formulated the principle in a way that applied to mechanics, rather than light, with the additional crucial hypothesis of constant energy, although Maupertuis simply referred to this as "a beautiful application of my principle".[1]

Despite Maupertuis' fervent advocacy for the importance of the Principle of Least Action, nowadays it definitely plays second fiddle to Hamilton's principle, but it is not that hard to understand why his principal was once so popular in mechanics. In the expression for $\mathcal{A}$, the term $\partial L/\partial \dot{q}^i$ is essentially momentum (page 445), so we are basically looking at the product of mass, velocity, and distance (the $\dot{q}^i(t)\,dt$ factor). In Chapter 15 we will see how this strange combination was hit upon, but one can at least imagine wanting to minimize this quantity, whereas there doesn't really seem to be any reason at all why one would want to minimize the Lagrangian in Hamilton's principle! In fact, nowadays the alluring name of "the principle of least action" is often used for Hamilton's principle, with "Maupertuis' principle of least action" used to specifically refer to Maupertuis' version.

Jacobi's form of the principal of least action. The inherent inelegance in the proper statement of the Principal of Least Action may account for the murky presentations that were often given. Jacobi, in his *Vorlesungen über Dynamik*, Jacobi [1] (now available in English translation), bluntly states that "In almost all textbooks, even in the best, those of *Poisson*, *Lagrange* and *Laplace*, the principle has been so presented that, in my view, it is impossible to understand." Jacobi then gave a geometric interpretation completely eliminating the time t.

When $L = T - V = \frac{1}{2}\langle\ ,\ \rangle - V$ for a Riemannian metric $\langle\ ,\ \rangle$ on M, our result states that a curve $c\colon [t_1, t_2] \to M$ is a solution of Lagrange's equations if and only if c is a critical point, under the allowed variations, for $\int_{t_1}^{t_2} T(c'(t))\,dt$. Since $T \geq 0$, it turns out (cf. DG, pp. 331–333 for details) that c is also a critical point for

$$\int_{t_1}^{t_2} \sqrt{T(c'(t))}\,dt = \text{arclength of } c \text{ from } c(t_1) \text{ to } c(t_2),$$

and conversely, given any such critical point c, reparameterization by arclength will then be a critical point for the original integral, giving a formulation of the principle that emphasizes the dependence only on the shape of the curve.

[1] Maupertuis' biography, entailing a somewhat scandalous scientific battle with socio-political overtones, appears to be nearly as convoluted as his principle. See Terrall [1] (and for contrasting views on König and the disputed letters of Leibniz see the `MacTutor History of Mathematics` articles "Johann Samuel König" and "The Berlin Academy and forgery"), as well as remarks in Dugas [1]. Thomas Carlyle's humorously condescending assessment of Maupertuis, written in 19$^{\text{th}}$ century high literary style, is quoted in Chandrasekhar [2; pp. 382–384].

We can also write our integral as

$$\int_{t_1}^{t_2} \sqrt{T(c'(t))}\,\sqrt{T(c'(t))}\,dt = \int_{t_1}^{t_2} \sqrt{(E_0 - V(c(t))}\,\sqrt{T(c'(t))}\,dt$$

[assuming that E_0, the constant energy of the curve c, satisfies $E_0 - V > 0$ on the image of c], showing that the solutions to Lagrange's equations can be identified as reparameterizations of the curves of constant energy $E = 1$ that are geodesics in the "Jacobi metric" $(E_0 - V)\langle\ ,\ \rangle$.

Noether's theorem. One application of Hamilton's principle provides a very simple proof of an important classical result.

Let $L\colon TM \to \mathbb{R}$ be any smooth map on the tangent space of M, in terms of which we can write Lagrange's equations, and consider a smooth one-parameter family of smooth maps $\phi_s\colon M \to M$; more precisely, we have a smooth map $\phi\colon (-\varepsilon, \varepsilon) \times M \to M$, and each ϕ_s denotes $p \mapsto \phi(s, p)$ for $p \in M$.

The maps ϕ_s can also be used to produce the maps $(\phi_s)_*$ from TM to TM. In order to avoid a superfluity of $*$'s, it will be more convenient simply to call these maps $\Phi_s\colon TM \to TM$. Then we can make the following definition:

> We say that the ϕ_s *preserve* L if for every tangent vector $\mathbf{v} \in TM$ we have $L(\Phi_{s*}(\mathbf{v})) = L(\mathbf{v})$.

[To find $\Phi_{s*}(\mathbf{v})$ for any fixed s, we take a curve γ with $\gamma'(0) = \mathbf{v}$, and consider the tangent vector at $t = 0$ of $t \mapsto \phi_s(\gamma(t))$.]

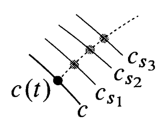
As a trivial example, which will at least show what we are talking about, if $M = \mathbb{R}^3$, with the standard coordinates (x^1, x^2, x^3), and

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + V(x^2, x^3),$$

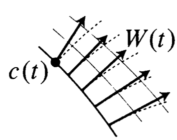
then $\phi_s(x^1, x^2, x^3) = (x^1 + s, x^2, x^3)$ preservc L. If the V term is absent, then $\phi_s(x^1, x^2, x^3) = (x^1 + as, x^2 + bs, x^3 + cs)$ preserve L if (a, b, c) has length 1.

Given a curve c in M we can consider the curves c_s defined by

$$c_s(t) = \phi_s(c(t)) = \phi(s, c(t)),$$



and if the ϕ_s preserve L and $c\colon \mathbb{R} \to M$ is a solution to Lagrange's equations for L, then each curve c_s will be also. We also have the "variation vector field"



$W = \partial\phi/\partial s$ along c determined by ϕ, with $W(c(t))$ being the tangent vector at 0 of the curve $s \mapsto \phi_s(c(t))$.

At each t consider

$$\Phi_c(t) = \lim_{h\to 0} \frac{L\big(c'(t) + h \cdot W(c(t))\big) - L(c'(t))}{h}.$$

(This can be thought of as the directional derivative of L restricted to $M_{c(t)}$ in the direction given by $W(t)$.) Noether's Theorem says that this quantity is constant along c (it is an "integral" for the solutions of Lagrange's equations).

NOETHER'S THEOREM. If the ϕ_s preserve L, then Φ_c is constant along any solution c of Lagrange's equations for L.

PROOF. Since the ϕ_s preserve L, each of the curves

$$c_s(t) = \phi_s(c(t)) = \phi(s, c(t))$$

is also a solution of Lagrange's equations for L, and thus an extremal for $\int_a^b L(c(t), c'(t), t)\, dt$, for all a and b in the interval under consideration.

The Boundary Term Corollary on page 462 then says that for all such a and b, we have

$$0 = \sum_{i=1}^{n} \frac{\partial q^i}{\partial x} \cdot \frac{\partial L}{\partial \dot{q}^i}\bigg|_a^b, \qquad \text{so that} \qquad \sum_{i=1}^{n} \frac{\partial q^i}{\partial x} \cdot \frac{\partial L}{\partial \dot{q}^i} \quad \text{is constant.} \quad ❖$$

The main problem with this proof is that it makes one question why the statement should be considered an important result! (for a more direct proof, see Arnold [2; §20]). Noether's original paper, Noether [1], was actually concerned with questions of continuum mechanics rather than the mechanics of particles, and the version for ordinary mechanics is really a "toy example", serving only as a template for extending the ideas to other areas, not only to continuum mechanics and fluid mechanics, but also to areas like quantum mechanics.

One simple illustration of our elementary version of Noether's theorem concerns the cyclic coordinates discussed on page 445. If q^1 is cyclic, then the ϕ_s given in coordinates by $(q^1, \dots, q^n) \mapsto (q^1 + s, q^2, \dots, q^n)$ will preserve L. In this case we have $\partial \dot{q}^1/\partial s = 1$, and $\partial \dot{q}^i/\partial s = 0$ for all other i, and Noether's theorem says that

$$\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i} \cdot \frac{\partial \dot{q}^i}{\partial s} = \frac{\partial L}{\partial \dot{q}^1} \quad \text{is a constant,}$$

so the invariance equation obtained for cyclic coordinates is a special case of Noether's theorem.

The lures of symmetry, advanced version. The idea behind Noether's theorem, of looking at the change of a Lagrangian $L = T - V$ under a one-parameter family of maps, is sometimes applied in the following manner.

Consider a closed system, where there are no outside acting forces, and hence V comes only from the interacting forces of the particles. For a closed system, the choice of the origin of our inertial system should make no difference. To write equations for this fact it will be convenient, instead of using the index α, to use an index $a = 1, \dots, K$, each of which is a triple, with an expression like $\partial L/\partial x^a$ also standing for a triple. For any unit vector $\mathbf{u} \in \mathbb{R}^3$, we can consider the map $\phi_s \colon \mathbb{R}^{3K} \to \mathbb{R}^{3K}$ that moves each particle over by $s \cdot \mathbf{u}$, conveniently written as $x^a \mapsto x^a + s \cdot \mathbf{u}$. The maps ϕ_s, which don't change any velocities, preserve the Lagrangian. So, with $\langle\ ,\ \rangle$ now denoting the inner product of $\mathbb{R}^3$, we have

$$0 = \frac{d}{ds}\bigg|_{s=0} L(\phi_{s*}(c'(t)) = \sum_a \left\langle \frac{\partial L}{\partial x^a}(c(t)), \mathbf{u} \right\rangle.$$

By Lagrange's equations, we then have

$$\sum_a \left\langle \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{x}^a}(c'(t))\Big), \mathbf{u} \right\rangle = 0,$$

and since $\mathbf{u}$ is an arbitrary unit vector we can conclude that

$$\sum_a \frac{\partial L}{\partial \dot{x}^a}(c'(t)) \quad \text{is constant.}$$

Since

$$\frac{\partial L}{\partial \dot{x}^a} = \frac{\partial T}{\partial \dot{x}^a} = m_a \dot{x}^a,$$

this just says that

$$\sum_a m_a (x^a \circ c)'(t) \quad \text{is constant,}$$

i.e., that momentum is conserved. (Instead of these calculations, we could have appealed to Noether's Theorem, with equivalent calculations then needed to express the invariant.)

Though it is hardly surprising that the conserved quantity just obtained is one that we already knew about, this derivation is sometimes regarded as a proof of conservation of momentum that relies only on the "homogeneity" of space. But the allure of this so-called proof is significantly diminished when we realize that the third law is already built into Lagrange's equations, though, to be sure, it slips into the equations in a subtle way that is easy to miss.

The equation $(*)$ on page 441, originally motivated by our consideration of constraints, can of course be applied in the case where there are no constraints, to get Lagrange's equation for a closed system, in particular for a collection of two particles with no outside forces. However, this equation is not a pair of equations on $\mathbb{R}^3$ for the two particles, separately involving the forces $\mathbf{F}_{12}$ and $\mathbf{F}_{21}$ of each particle on the other, but a single equation on $\mathbb{R}^6$, and this single equation holds *precisely because* we have $\mathbf{F}_{12} = -\mathbf{F}_{21}$. For the case of constraints, and in particular, for the case of rigid bodies, equation $(*)$ may appear to have been derived quite naturally, but again, this was only because somewhere along the way we used the third law.

Needless to say, the more elegant, invariant, and sophisticated the presentation of Lagrangian mechanics and Noether's theorem, the easier it is to hide this fact, and thus appear to demonstrate that the basic laws of mechanics can be derived magically from symmetry.

Invariance of the Lagrangian can be applied in the same way to the rotational symmetry of space, and we should not be too surprised to find (Problem 1) that we simply obtain the law of conservation of angular momentum, which also must be implicit in Lagrange's equations, though now there is an added nasty twist, since conservation of angular momentum actually requires the *strong form* of the third law.

ADDENDUM 13A

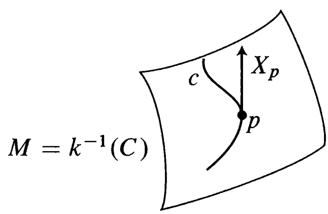
LAGRANGE MULTIPLIERS FOR CONDITIONAL CRITICAL POINTS

This material is taken from DG, Vol. 4, pp. 295–300.

Suppose we are given two differentiable functions $j, k\colon \mathbb{R}^n \to \mathbb{R}$, and we seek a critical point of j on the set $k^{-1}(C)$. The method of Lagrange multipliers states that if p is a critical point of j on $k^{-1}(C)$, and p is not a critical point of k, then there is a number λ such that

$$\frac{\partial j}{\partial x_i}(p) = \lambda \frac{\partial k}{\partial x_i}(p) \qquad i = 1, \dots, n. \tag{1}$$

Problem 5-2 states a more general result, but for the moment we will simply repeat the argument, which is so crucial to the present discussion, for this special case. We note that the hypotheses on k imply that in a neighborhood of p, the set $k^{-1}(C) \subset \mathbb{R}^n$ is a hypersurface M, and that $k_*(X_p) = 0$ for $X_p \in \mathbb{R}^n{}_p$ precisely when $X_p \in M_p$. Every such X_p is $c'(0)$ for some curve c in M. It



follows that $j(c(t))$ has a critical point at $t = 0$, which means that $j_*(X_p) = 0$. Thus the two linear functions $j_*, k_*\colon \mathbb{R}^n{}_p \to \mathbb{R}$ have the property that $\ker k_* \subset \ker j_*$. This implies that $j_* = \lambda k_*$ for some λ, which is equivalent to equation (1).

Now given two functions

$$F\colon \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \qquad \text{and} \qquad G\colon \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R},$$

we define, for a function $f\colon [a, b] \to \mathbb{R}$,

$$J(f) = \int_a^b F(t, f(t), f'(t))\, dt$$

$$K(f) = \int_a^b G(t, f(t), f'(t))\, dt.$$

Among all functions $f\colon [a,b] \to \mathbb{R}$ with fixed values at a and b, and a fixed value $K(f) = C$, we seek one that is a critical point for J.

We want to apply the idea of the above analysis to our two functions J and K; for simplicity, we assume that all functions are suitably differentiable, without worrying about the exact degree of differentiability required. Suppose that f is a critical point for J on $K^{-1}(C)$ but f is not a critical point of K. Consider any variation $\alpha\colon (-\varepsilon,\varepsilon)\times[a,b] \to \mathbb{R}$ of f that keeps endpoints fixed. We know from the formula at the top of page 462 that $dJ(\bar{\alpha}(u))/du\big|_{u=0}$ depends only on the function $\partial\alpha/\partial u(0,t)$ on $[a,b]$. For W on $[a,b]$ with $W(a) = W(b) = 0$, we define

$$J_*(W) = \left.\frac{dJ(\bar{\alpha}(u))}{du}\right|_{u=0} \qquad \begin{array}{l}\text{for any variation } \alpha \text{ of } f\\ \text{with } \partial\alpha/\partial u(0,t) = W(t)\end{array}$$

(we should really write something like $J_{f*}(W)$, but leave out the f for convenience). Notice that there always is a variation α with this property, for example,

$$\alpha(u,t) = f(t) + uW(t).$$

Defining $K_*(W)$ in the same way, we have functions $J_*, K_*\colon \mathcal{V} \to \mathbb{R}$, where $\mathcal{V}$ is the vector space of all functions W on $[a,b]$ with $W(a) = W(b) = 0$. We claim that J_* (and likewise K_*) is linear. To see this we choose two variations α_1 and α_2 with

$$\frac{\partial\alpha_i}{\partial u}(0,t) = W_i(t),$$

and define the variation α by

$$\alpha(u,t) = \alpha_1(u,t) + \alpha_2(u,t).$$

Then

$$\frac{\partial\alpha}{\partial u}(0,t) = W_1(t) + W_2(t),$$

so

$$\begin{aligned} J_*(W_1 + W_2) &= \left.\frac{dJ(\bar{\alpha}(u))}{du}\right|_{u=0} \\ &= \left.\frac{dJ(\bar{\alpha}_1(u))}{du}\right|_{u=0} + \left.\frac{dJ(\bar{\alpha}_2(u))}{du}\right|_{u=0}, \\ &\qquad \text{as one sees by inspecting the formula on page 462,} \\ &= J_*(W_1) + J_*(W_2). \end{aligned}$$

Homogeneity is proved similarly.

LEMMA. If $K(f) = C$, where the function f is not a critical point of K, and $K_*(W) = K_{f*}(W) = 0$, then $W = \partial\alpha/\partial u(0,t)$ for some variation α with the property that each $\bar{\alpha}(u)$ is in $K^{-1}(C)$.

PROOF. Since f is not a critical point, there is W_1 with $K_*(W_1) \neq 0$. Let $L\colon \mathbb{R}^2 \to \mathbb{R}$ be

$$L(r,s) = K(f + rW + sW_1).$$

If we define

$$\beta(u,t) = f(t) + uW_1(t),$$

then β is a variation of f with $\partial\beta/\partial u(0,t) = W_1(t)$ and $\bar{\beta}(u) = f + uW_1$. So

$$K_*(W_1) = \lim_{u\to 0} \frac{K(f + uW_1) - K(f)}{u} = \frac{\partial L}{\partial s}(0,0).$$

Similarly,

$$K_*(W) = \frac{\partial L}{\partial r}(0,0).$$

Since

$$\begin{cases} L(0,0) = K(f) = C \\ \dfrac{\partial L}{\partial s}(0,0) = K_*(W_1) \neq 0, \end{cases}$$

the implicit function theorem shows that there is a function $r \mapsto s(r)$, from a neighborhood of 0 in $\mathbb{R}$ to a neighborhood of 0 in $\mathbb{R}$, such that

$$(1) \qquad C = L(r, s(r)) = K(f + rW + s(r)W_1) \qquad \text{for small } r.$$

Notice that the first part of the equation gives, upon differentiating with respect to r,

$$0 = \frac{\partial L}{\partial r}(0,0) + \frac{\partial L}{\partial s}(0,0)s'(0) = K_*(W) + K_*(W_1)s'(0) = K_*(W_1)s'(0),$$

and hence

$$s'(0) = 0.$$

Thus, if we define the variation α by

$$\alpha(u,t) = f(t) + uW(t) + s(u)W_1(t),$$

then each $\bar{\alpha}(u) = f + uW + s(u)W_1$ is in $K^{-1}(C)$ by (1), and also

$$\frac{\partial\alpha}{\partial u}(0,t) = W(t) + s'(0)W_1(t) = W(t). \ ❖$$

THEOREM (Euler's Rule). If f is a critical point of J on $K^{-1}(C)$ and f is not a critical point of K, then there is a number λ such that f is a critical point of $J - \lambda K$ (and consequently the Euler equations for $J - \lambda K$ hold for f).

PROOF. Consider the two linear functions $J_*, K_* \colon \mathcal{V} \to \mathbb{R}$. If $K_*(W) = 0$, let α be the variation given by the Lemma, with all $\bar{\alpha}(u)$ in $K^{-1}(C)$. Since f is a critical point of J on $K^{-1}(C)$, function $u \mapsto J(\bar{\alpha}(u))$ has a critical point at 0, and consequently

$$J_*(W) = \left.\frac{dJ(\bar{\alpha}(u))}{du}\right|_{u=0} = 0.$$

Thus $\ker K_* \subset \ker J_*$. The vector space $\mathcal{V}$ is infinite dimensional, but it still follows (Problem 5-2) that there is a number λ with $J_* = \lambda K_*$, which is equivalent to the assertion that f is a critical point of $J - \lambda K$. ❖

The straightforward extension to the case of functions $f \colon [a, b] \to \mathbb{R}^m$ is left to the reader.

PROBLEM

1. (a) Consider rotations $A(s)$ for which $\boldsymbol{\omega}$, the vector corresponding to the skew-symmetric matrix $A'(0)$, satisfies $|\boldsymbol{\omega}| = 1$, and let $\phi_s : \mathbb{R}^{3K} \to \mathbb{R}^{3K}$ be defined by

$$\phi_s(x^a) = \boldsymbol{\omega} \times x^a.$$

Show that

$$\begin{aligned}
0 &= \left.\frac{d}{ds}\right|_{s=0} L(\phi_{s*}(c'(t))) \\
&= \sum_a \left\langle \frac{\partial L}{\partial x^a}(c(t)),\ \boldsymbol{\omega} \times c_a(t) \right\rangle + \sum_a \left\langle \frac{\partial L}{\partial \dot{x}^a}(c'(t)),\ \boldsymbol{\omega} \times {c_a}'(t) \right\rangle \\
&= \left\langle \boldsymbol{\omega},\ \sum_a \frac{d}{dt}\left(c_a(t) \times \frac{\partial L}{\partial \dot{x}^a}(c'(t)) \right) \right\rangle,
\end{aligned}$$

leading to conservation of angular momentum.
(b) Note that here we are using the standard form of Lagrange's equations, which involve *conservative forces*, and the result is not so surprising when we note that the most common conservative forces are radially symmetric ones! Compare the situation for the more general form of Lagrange's equations given in the footnote on page 443.

CHAPTER 14

SMALL OSCILLATIONS

Small oscillations about equilibrium points, briefly alluded to in Chapter 6, and investigated in special cases in Chapter 8, can be given a simple unified approach in terms of Lagrangian mechanics, supplying a short coda to Part III.

We will be considering a Lagrangian $L\colon TM \to \mathbb{R}$ of the form

$$L = T - V = \tfrac{1}{2}\langle\ ,\ \rangle - V$$

for a Riemannian metric $\langle\ ,\ \rangle$ on M, written, say, as

$$L = \tfrac{1}{2}\sum_{i,j=1}^{n} g_{ij}\dot q^i \dot q^j - V.$$

We first seek equilibrium points for the Lagrange equations, that is, we ask when a constant curve $c(t) = p \in M$ can be a solution of the equations

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot q^i}(c'(t))\right) - \frac{\partial L}{\partial q^i}(c(t)) = 0,$$

where $c'(t)$ is now always the zero vector $0_p \in M_p$. Since $\partial T/\partial q^i$ and $\partial T/\partial \dot q^i$ will be a sum of terms involving at least one $\dot q^j$, each of which is 0 at 0_p, the part of Lagrange's equation involving T will automatically be 0. So, analogously to the situation in Chapter 6 (page 213ff.), we have an equilibrium point precisely when each

$$\frac{\partial V}{\partial q^i}(p) = 0.$$

If p is an equilibrium point, the function $T \geq 0$ on TM has a minimum value at 0_p, so if V has a strict local minimum at p, then the energy $E = V + T$ also has a strict local minimum at 0_p, this minimum E_0 being equal to $V(p)$. For small $\varepsilon > 0$, the set of points in TM with $E \leq E_0 + \varepsilon$ will be a small neighborhood of 0_p (we need a strict local minimum at p to insure that the neighborhood is small). If we consider the component of this neighborhood that contains 0_p, then conservation of energy says that any solution curve starting at a point of this component must stay in this component, which shows that p is a point of stable equilibrium.

(On the other hand, even if V has a strict local maximum at $\mathscr{p}$, it does not follow that $\mathscr{p}$ is a point of unstable equilibrium. It is easy to see that one can't expect this in the C^∞ case, but even in the analytic case it is apparently only known for the easy case of dimension 1, and for dimension 2.)

We now want to "linearize" the Lagrange equations near an equilibrium point, our linearization of the equations for the spherical pendulum in Chapter 8 (page 291) being a special case. We note that

$$V(q^1,\dots,q^n) = V(\mathscr{p}) + \sum_{i=1}^{n} \frac{\partial V}{\partial q^i}(\mathscr{p})q^i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial q^i \partial q^j}(\mathscr{p})q^i q^j + \cdots,$$

where we might as well choose $V(\mathscr{p}) = 0$, and where the first sum is 0 because $\mathscr{p}$ is an equilibrium point, while

$$g_{ij}(q^1,\dots,q^n) = g_{ij}(\mathscr{p}) + \sum_{k=1}^{n} \frac{\partial g_{ij}}{\partial q^k}(\mathscr{p})q^i + \cdots,$$

and since T is quadratic in the $\dot{q}^i$, the lowest nonvanishing approximation to T is simply

$$\tfrac{1}{2}\sum_{i,j=1}^{n} g_{ij}(\mathscr{p})\dot{q}^i\dot{q}^j,$$

giving us the linearization

$$(*)\qquad \sum_{j=1}^{n} g_{ij}(\mathscr{p})(c^j)''(t) + \sum_{j=1}^{n} \frac{\partial^2 V}{\partial q^i q^j}(\mathscr{p})c^j(t) = 0, \qquad i = 1,\dots,n.$$

[More formally, linearization for a set of first order equations was already introduced in Addendum 8C, and a set of second order equations, like the Lagrange equations, is linearized by converting it into a system of first order equations in $2n$ variables (as in Addendum 10A). Choosing a coordinate system q so that all $q^i(\mathscr{p}) = 0$ for our equilibrium point $\mathscr{p}$, Problem 1(b) shows that $(*)$ is indeed the appropriate linearization.]

Thus, the linearized system is described by the two symmetric $n \times n$ matrices $\mathbf{T} = \big(g_{ij}(\mathscr{p})\big)$ and $\mathbf{V} = \big(V_{ij}(\mathscr{p})\big) = \big(\partial^2 V/\partial q^i \partial q^j(\mathscr{p})\big)$, where the matrix $\mathbf{T}$ is positive definite.

Symmetry of $\mathbf{T}$ and $\mathbf{V}$ means that the corresponding linear transformations $\boldsymbol{T}, \boldsymbol{V}\colon \mathbb{R}^n \to \mathbb{R}^n$ are self-adjoint with respect to the usual inner product on $\mathbb{R}^n$:

$$\langle \boldsymbol{T}(v), w\rangle = \langle v, \boldsymbol{T}(w)\rangle, \quad \langle \boldsymbol{V}(v), w\rangle = \langle v, \boldsymbol{V}(w)\rangle \qquad \text{for all } v, w \in \mathbb{R}^n.$$

It follows that there is a diagonalizing basis $w_1, \dots, w_n$ of $\mathbb{R}^n$ such that

$$\text{(D)} \qquad T(w_i) = w_i, \qquad V(w_i) = \lambda_i w_i$$

for certain $\lambda_1, \dots, \lambda_n$. In fact, since $\mathbf{T}$ is positive definite, we can choose a orthonormal basis for the inner product $\langle\!\langle v, w \rangle\!\rangle = \langle T(v), w\rangle$. When we write $\mathbf{V}$ in this basis it is still symmetric, since that is still equivalent to the condition $\langle V(v), w\rangle = \langle v, V(w)\rangle$ for all $v, w \in \mathbb{R}^n$. So $\mathbf{V}$ has a basis of eigenvectors $v_1, \dots, v_n$ which are orthonormal for $\langle\!\langle\ ,\ \rangle\!\rangle$, with corresponding eigenvalues λ_i.

If we let $C : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation that takes the standard basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ to $v_1, \dots, v_n$, then in the coordinate system $C \circ q$, equations $(*)$ reduce to

$$(c^i)''(t) + \lambda_i c^i(t) = 0.$$

When $\mathbf{V}$ is positive definite, implying that V has a strict local minimum at p, and thus that p is a stable equilibrium point, the general solution is the sum of harmonic oscillations, along axes that are orthogonal with respect to $\mathbf{T}$. If $\mathbf{V}$ instead has some $\lambda_i < 0$, we have to allow exponential "oscillations" for each such λ. For $\lambda_i = 0$ we might have an oscillation only discoverable by looking at higher derivatives of V, and not approximable by a harmonic oscillation, or it might indicate uniform linear motion (see the example at the end of this chapter). Including all these under the rubric of oscillations, we can then state the result:

> All small oscillations of a system about an equilibrium point can be written as the sum of small oscillations along a set of $\mathbf{T}$-orthogonal axes.

It is important to remember, however, that the standard short-hand terminology "small oscillations of a system", means nothing other than "the oscillations of the first order approximation to a system". As in the case of the spherical pendulum, we are not actually studying small oscillations of the system itself, and, in fact, for the general case, the system may not have *any* small motions that are oscillations. Moreover, the sum of two or more of the small oscillations along the $\mathbf{T}$-orthogonal axes will not be periodic if the corresponding λ_i are not commensurable. (a situation similar to that for Lissajous figures in Chapter 8).

Note that equation (D) implies that the λ_i are the roots of the equation

$$\det(\mathbf{V} - \lambda\mathbf{T}) = 0,$$

which connects this description of the solutions with the normal modes studied in Chapter 8. That is, if we look for solutions of $(*)$ of the form

$$c^j(t) = A_j e^{i\omega t},$$

with all q^j components having the same period ω, we obtain

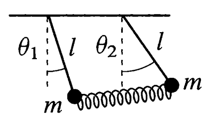
$$\sum_{j=1}^{n}\bigl(V_{ij}(p) - \omega^2 g_{ij}(p)\bigr)c^j(t) = 0,$$

which shows that $\det(\mathbf{V} - \omega^2\mathbf{T}) = 0$, so the normal modes correspond to $\omega_i = \sqrt{\lambda_i}$. Note, by the way, that in the normal modes, all q^j components not only have the same period, but they are also in sync (have the same phase) for each pair with A_j having the same sign, or in anti-sync (having a phase difference of π) for pairs having opposite sign.

Thus we see that small oscillations of a system near an equilibrium point can always be written as sums of normal modes, as promised in Chapter 8, where we investigated the special case of N harmonic oscillators all interacting in a linear way. However, the method of our general proof may not provide the best route for analyzing small oscillations of any particular system, because the natural choice of coordinates for finding the Lagrangian may not end up being the best choice for the analysis of small oscillations.

An extreme example is afforded by the spherical pendulum. Recall that for small oscillations, the coordinates x and y on page 291 seemed preferable to the coordinates θ and ϕ. As a matter of fact, in this case we *can't* use θ and ϕ, because, as we see from the formula for v^2 on page 290, $\mathbf{T}$ is not even positive definite at the equilibrium point. (On the other hand (Problem 3-5), this only eliminates cases where the pendulum is actually swinging in a plane.)

In our investigation of two coupled oscillators on page 302, the angles θ_1 and θ_2 that the pendulums make with the vertical are the obvious choice for writing the Lagrangian, but our choice of the linear coordinates x_1 and x_2



worked out much better when looking for normal forms, as we did in a more general case on page 305, because it is easier to make approximations in terms of these coordinates. We didn't actually write down $T - V$ in that case, but simply passed directly to approximating equations. Even when the general method is used, approximations are usually made as early as possible, as in Problem 2, rather than obtaining exact formulas and taking derivatives. Of course, such approximations can be tricky, while the longer straightforward method can always be relied upon.

In this regard, let us consider again the case of the double pendulum discussed on page 307. Using the components indicated in the figure on that page, we have

$$\text{(A)} \qquad \begin{aligned} x_1 &= l_1 \sin\theta_2, & x_2 &= l_1 \sin\theta_1 + l_2 \sin\theta_2 \\ y_1 &= l_1 \cos\theta_1, & y_2 &= l_1 \cos\theta_1 + l_2 \cos\theta_2, \end{aligned}$$

and we obtain

$$\begin{aligned} T &= \frac{m_1}{2}(\dot{x}_1{}^2 + \dot{y}_1{}^2) + \frac{m_2}{2}(\dot{x}_2{}^2 + \dot{y}_2{}^2) \\ &= \frac{m_1+m_2}{2} l_1{}^2\dot{\theta}_1{}^2 + \frac{m_2}{2} l_2{}^2\dot{\theta}_2{}^2 + m_2 l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 \\ V &= -m_1 g y_1 - m_2 g y_2 \\ &= -(m_1+m_2) g l_1 \cos\theta_1 - m_2 g l_2 \cos\theta_2 . \end{aligned}$$

The need to calculate the tension on the string for the lower pendulum has been eliminated because d'Alembert's principle for constraints has been incorporated into Lagrange's equations. There is also no need to estimate the $\dot{\theta}_i$, for we can simply calculate the derivatives at $(0, 0)$, finding that

$$\mathsf{T} = \begin{pmatrix} (m_1+m_2) l_1{}^2 & m_2 l_1 l_2 \\ m_2 l_1 l_2 & m_2 l_2{}^2 \end{pmatrix}, \qquad \mathsf{V} = \begin{pmatrix} (m_1+m_2) g l_1 & 0 \\ 0 & m_2 g l_2 \end{pmatrix},$$

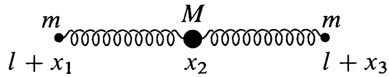
which leads to the equations

$$\begin{aligned} \ddot{\theta}_1 + \frac{g}{l_1}\theta_1 &= -\frac{m_2}{m_1+m_2}\frac{l_2}{l_1}\ddot{\theta}_2 \\ \ddot{\theta}_2 + \frac{g}{l_2}\theta_2 &= -\frac{l_1}{l_2}\ddot{\theta}_1 . \end{aligned}$$

We then obtain the equations $(*)$ on page 308 when we switch back to x_1 and x_2 using (A), which for small θ_i simplify to

$$\theta_1 = \frac{x_1}{l_1}, \qquad \theta_2 = \frac{x_2 - x_1}{l_2}.$$

A final point is illustrated by the oft-presented example of the linear triatomic molecule, where two atoms of mass m are symmetrically located on each side of an atom of mass M, all three atoms lying on a straight line, with the forces between the atoms approximated by those produced by springs with spring constant k (the molecules H_2O and CO_2 fit this model reasonably well). If l



is the unstretched length for these springs, it is convenient to use x_2 as the coordinate of the middle atom, with $l + x_1$ and $l + x_3$ being the coordinates of the outer atoms. We then have

$$T = \frac{m}{2}(\dot{x}_1{}^2 + \dot{x}_3{}^2) + \frac{M}{2}\dot{x}_2{}^2$$

$$V = \frac{k}{2}[(x_1 - x_2)^2 + (x_3 - x_2)^2],$$

so that

$$\mathbf{T} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}, \qquad \mathbf{V} = \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}.$$

The characteristic polynomial

$$0 = \det(\mathbf{V} - \lambda\mathbf{T}) = \det\begin{pmatrix} k - \lambda m & -k & 0 \\ -k & 2k - \lambda M & -k \\ 0 & -k & k - \lambda m \end{pmatrix}$$

reduces to

$$\lambda(k - \lambda m)[k(M + 2m) - \lambda m M] = 0,$$

so that the $\omega_i = \sqrt{\lambda_i}$ are

$$\omega_1 = 0, \qquad \omega_2 = \sqrt{\frac{k}{m}}, \qquad \omega_3 = \sqrt{\frac{k}{m}\left(1 + \frac{2m}{M}\right)}.$$

In this particular case, the zero value of ω_1 simply indicates that the whole molecule can be moving horizontally with uniform velocity, and it arises because we have used three coordinates, while there are really only two degrees of freedom. We could remove this redundancy by adding an extra constraint, for example the condition that the center of mass of the molecule should remain stationary, but the coordinates we chose are useful for discussing the nature of the oscillations for the two normal modes ω_2 and ω_3, which can be found by solving for the eigenvectors, or in this case simply by enlightened guessing:

(1) For ω_2, the middle atom remains fixed, while the outer atoms oscillate with frequency ω_2 and equal amplitude, but in anti-synch.

(2) For ω_3, the two outer atoms oscillate in synch with frequency ω_3, with equal amplitude, while the middle atom oscillates in anti-synch with them, with $2m/M$ times the amplitude.

PROBLEMS

1. (a) Lagrange's equations for a curve c are a set of second order equations for the $c^i = q^i \circ c$, but they are not written in the standard form

$$\frac{d^2c^i}{dt^2} = A_i(t, c(t), dc/dt),$$

and in fact it may not be possible to write them this way, because the matrix of coefficients of the c^i,

$$\frac{\partial^2 T}{\partial \dot{q}^i \dot{q}^j}$$

might be singular (a case normally discarded by considering only "regular" Lagrangians, see Chapter 16). For the case where

$$L(\mathbf{v}) = \tfrac{1}{2}\langle \mathbf{v}, \mathbf{v}\rangle^2 - V$$

for a Riemannian metric $\langle\ ,\ \rangle$ on M, so that

$$L(c'(t)) = \sum_{i,j=1}^{n} \langle (c^i)'(t), (c^j)'(t)\rangle - V,$$

the matrix, being positive definite, is certainly not singular, and we can put them in standard form by mimicking the procedure for putting the equations for geodesics into standard form (DG, pp. 326–328): If the Riemannian metric $\langle\ ,\ \rangle = \sum_{i,j=1}^{n} g_{ij}\, dq^i \otimes dq^j$, with corresponding Γ_{ij}^k, show that Lagrange's equations can now be written as

$$\frac{d^2c^k}{dt^2} + \sum_{i,j=1}^{n} \Gamma_{ij}^k(c(t)) \frac{dc^i}{dt}\frac{dc^j}{dt} = -\sum_{l=1}^{n} g^{kl} \frac{\partial V}{\partial q^l}.$$

In terms of the covariant derivative ∇ (DG, Vol. 2, Chap. 6), the left side of the equation can be written as $\nabla_{c'(t)} c'$, and if grad V is defined by the equation

$$\langle \operatorname{grad} V, X\rangle = dV(X) = X(V) \qquad \text{for all tangent vectors } X,$$

our equation can be written as

$$\nabla_{c'(t)} c' = -\operatorname{grad} V(c(t))$$

or, using the D/dt notation,

$$\frac{Dc'(t)}{dt} = -\operatorname{grad} V(c(t)).$$

(b) Use part (a) to show that (∗) on page 477 is the appropriate linearization of the Lagrange equations near a equilibrium point.

2. (a) For the coupled oscillators on page 479, where the i^{th} pendulum bob has coordinates (x_i, y_i) with

$$\begin{aligned} x_i &= l \sin \theta_i \\ y_i &= l(1 - \cos \theta_i), \end{aligned}$$

one can write, or at least imagine the mess one would obtain by writing, the Lagrangian in terms of the θ_i. Instead, write an approximation directly in terms of the θ_i and m, l, and the spring constant k.

Answer:

$T = \frac{ml^2}{2}(\dot{\theta}_1{}^2 + \dot{\theta}_2{}^2)$

$V = \frac{mgl}{2}(\theta_1{}^2 + \theta_2{}^2) + \frac{kl^2}{2}(\theta_2 - \theta_1)^2$

(b) One can now determine **T** and **V**, and then proceed to find the solutions of $\det(\mathbf{V} - \lambda\mathbf{T}) = 0$, but we already know the obvious normal modes for this case, so instead simply express **T** and **V** in terms of $\phi_1 = \theta_1 + \theta_2$ and $\phi_2 = \theta_1 - \theta_2$ and determine the oscillation periods for the two modes.

INTERLUDE

CHAPTER 15
LIGHT

Light!? This is supposed to be a book about mechanics. What is a chapter about light doing here?

Well, perhaps it doesn't seem all that strange, since light plays a role in special relativity, though we don't cover that topic in this volume, and quantum mechanics has joined particles and light waves in a bond that no physicist may put asunder. But the studies of light and mechanics have actually been closely intertwined throughout their history, and this interconnection helps explain the development of Hamiltonian mechanics, and its relation to quantum mechanics.

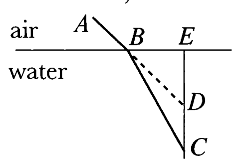
Optics in antiquity. The darkness shrouding early investigations of light is amply described in Ronchi [1]. Euclid (∼ 330–260 B.C.), in his generally valuable *Optics*, accepted the view that the eye emitted rays (the Greeks focused on *vision*, not light, presuming that the eye must take an active part—as a later commentator asked, if light rays emanated from objects rather than the eye, how was it that one often couldn't find something one was searching for, even though it was it plain sight?). Euclid [1] even explained that far away objects couldn't be seen because they were situated between two adjacent rays spread out from the eye!

The law of reflection, "Angle of incidence equals angle of reflection" was simply taken as an axiom, and the renowned mathematician, engineer, and inventor of ancient times Hero of Alexandria (∼ 10–70 A.D.), is credited with the observation that reflected light, in going from A to B via a point C on a mirror, always follows a path that makes the total distance $AC+CB$ shortest. This may well be the earliest "variational principle" ever enunciated, aside from the fact that light travels in straight lines, to be followed 16 centuries later by yet another variational principle involving light.

Islamic scholars. Though familiar with Euclid's *Optics*, Islamic scholars believed that light was emitted from the objects one saw, the best evidence being the *camera obscura* (basically a pin-hole camera). The important scholar Abu Ali Mohammed Ibn Al Hasan Ibn Al Haytham (∼ 965–1039), known to the western world as Alhazen, published a book in which he compared the reflection of light to a body's motion, noting that if an arrow with a small spherical body at the tip is shot toward a mirror on a wall "at an angle to the perpendicular we shall see that the arrow is reflected back not on the same line by which it came, but in a direction which is symmetrical to the first with reference to the perpendicular to the mirror." Moreover, Alhazen explained this by resolving the motion into components parallel and normal to the wall, as Galileo would do centuries later, with the parallel component remaining unchanged and the normal component reversed. A Latin translation of Alhazen's book in 1572 played an important part in overthrowing the dominance of ancient Greek thought in the west.

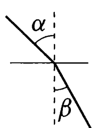
Kepler and Galileo. In 1604, Kepler (1571–1630), best known to most of us for his astronomical investigations, published a book about vision, explaining for the first time how the image of an object is focused by a lens, and confidently writing "I say that vision takes place when the image of the whole hemisphere of the world in front of the eye and even a little more, is formed upon the concave reddish surface of the retina". But it was a mystery how the lens of the eye could change its focal point, and for a long time Kepler's book was barely noticed, lenses being considered a matter for craftsmen, rather than scientists (it is not even known who invented spectacles), though the interest in telescopes of Galileo (1564–1642) forced a revision to the scholastic view that had regarded telescopes not only as déclassé, but downright misleading, totally unacceptable devices for use in science. The third chapter of Ronchi [1], "The downfall of ancient optics", gives an account of these intriguing developments, which inaugurated the science of optics, and swept away centuries of misunderstanding and confusions about light and vision, leaving us free to examine more modern misunderstandings and confusions.

Descartes. The first accurate law of refraction is usually attributed to work in the early 1600's by Willebrord Snel van Royen (1580–1626), a Dutch mathematician whose Latin name "Snellius" led to the spelling Snell, used by everyone except au courant scholars.[1] Considering a ray of light AB in air that is refracted along BC after hitting water at B, and letting BD be the extension of AB to the same horizontal distance as C, Snell found that the ratio BC/BD was



always the same. Writing $BC/BD = (BE/BD)/(BE/BC)$, we can express the result as the "sine law"

$$\frac{\sin\alpha}{\sin\beta} = \text{constant.}$$



Descartes (1596–1650) reached exactly this conclusion, and the suggestion that he might have borrowed from Snell excited some nationalistic feelings, so that

[1] Recent scholarship has also called attention to the work of Ibn Sahl in 984 and Thomas Harriot in 1602.

the sine law of refraction is generally known as Snell's Law, but in France as Descartes' Law. In any case, Descartes gave an argument for the law, although his views on light were rather unformed and obscure, generally regarding light as a pressure, rather than a movement of particles. Like Alhazen, he resolved the motion (or whatever) of the light into components parallel and normal to the line separating the air and water, with the parallel component remaining unchanged, while the other component, rather than being reflected, changed magnitude so that the resultant velocity would be the speed of light in water. This led to the result that the ratio $\sin\alpha/\sin\beta$ for the incident and refracted angles is a constant, namely, the ratio v_2/v_1 of the speed of light in water to the speed of light in air.

Alhazen had proposed a somewhat similar explanation for refraction, but it seems he didn't recognize the difficulty that this explanation entailed: since the angle of refraction is less than the angle of incidence, the speed of light would have to be greater in water than in air. Descartes did recognize the difficulty, admitting "Perhaps you would be surprised if you carried out the experiment", and resorted to considerable extemporizing (Ronchi [1; pg. 117]) to support the idea that the speed should be greater in water than in air, which everyone, including Descartes himself and his supporters, found rather unnatural.

Fermat. Fermat (1601–1695) rejected these arguments completely. He not only regarded Descartes' view of the velocities as absurd, but also objected to reliance on mechanical analogies, saying that one should start "from the principle, so common and so well-established, that Nature always acts in the shortest ways", nowadays ensconced in Fermat's principle: the path followed by light is always a critical path for the time. But there were obstacles to overcome.

Assured that experiments confirmed Descartes' law—"All the difficulty was therefore reduced to the fact that it appeared that I had to fight not only men but also Nature"—Fermat only reluctantly decided to see if his principle at least gave a law that agreed with Descartes' law within experimental error. His reluctance was partly due to the fact that finding the path of least time was not so simple as in the case of reflection. Fermat had "my method of *maximis* and *minimis*, which is rather successful for expediting this kind of problem", basically setting a derivative to zero [Newton read, and generalized, this method], but finding the equivalent of the derivative is not so simple when the square roots have to be dealt with directly, without benefit of the chain rule and basic facts about derivatives. However, at the end of the calculations he found that "the reward for my effort has been the most extraordinary, the most unforseen, and the happiest that ever was", namely that the path taking the least time followed exactly Descartes' law for diffraction in the form $\sin\alpha/\sin\beta = v_1/v_2$, but with the assumption that the speed of light in water is *less* than in air, $v_2 < v_1$.

Fermat's calculations, in an unpublished paper *Analysis ad Refractiones* attached to correspondence, is given in Sabra [1; pp. 144ff.], where one can clearly see the precursor of the derivative. (In some later correspondence, a paper *Synthesis ad refractiones* took the classically preferred opposite tack, giving a direct, rather complicated, argument that the path satisfying the sine law takes less time than any other, cf. Sabra [1; pp. 150–152] or Dugas [1; pp. 255–257].)

The supporters of Descartes opposed Fermat's explanation as disdainfully as Lagrange, a century later, would dismiss the significance of the principle of Least Action, which, as we will see later on, stemmed from Fermat's principle in a perverted sort of way; and they gave the same sort of criticisms that most of us would give to such teleological principles today. Feynman [1; pp. 26-7 to 26-8] has a discussion of Fermat's principle from a viewpoint informed by modern physics, which gives it something of an explanatory role, rather than a mere metaphysical one.

Huygens. Christiaan Huygens (1629–1695), one of the men mentioned in the *Principia* as "the foremost geometers of the previous generation" (cf. page 22), was once referred to by Newton as "Summus Hugenius" for his exposition, and Chapter 8 gives some examples of his mastery of geometry.

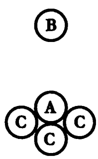
In 1678, Huygens published a short book, *Traité de lumiere*, available in English translation as *Treatise on Light*, Huygens [2], and a comparison with the *Principia* reveals the extent of the generational divide. The *Principia* is difficult to read not only because Newton uses his geometric prowess to provide many proofs, but also because he states and derives almost all formulas in words, often accompanied by geometric figures. In Huygens work, by contrast, there basically are no formulas at all—everything is geometry.

Huygens objected to the notion that light consists of streams of particles, because there are obviously light rays simultaneously traveling in essentially all directions, including directly opposing directions, as when two people see each other, so it would seem impossible for these particles not to interfere with each other; and Descartes' vague ideas had similar problems.

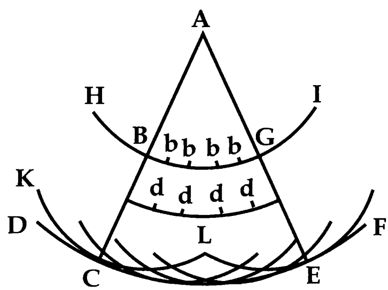
Huygens instead proposed that light must spread something like sound in the air, produced "by a movement which is passed on successively from one part of the air to another; and that the spreading of this movement, taking place equally rapidly on all sides, ought to form spherical surfaces ever enlarging and which strike our ears". For the analogous surfaces in the case of light, he chose the word 'waves' [what we would call wave fronts] "from their resemblance to those which are seen to be formed in water when a stone is thrown into it, and which present a successive spreading as circles, though these arise from another cause, and are only in a flat surface", perhaps the first use of the word 'waves' in any situation other than water waves. Though differing in nature, waves in water give a very nice example of waves in different directions passing through each other to supplement the evidence from sound waves.

The speed of light had been estimated a little earlier, by Roemer in 1676, from the delay in the eclipses of Jupiter's satellites, and Huygens explained the possibility of such a great, but finite, speed as Roemer's estimate gave by pointing out that when a single spherical object hits a line of identical touching ones, the motion passes "as in an instant to the last of them", though it is certainly not instantaneous, but successive, for if the movement "did not pass successively through all these spheres, they would all acquire the movement at the same time, and hence would all advance together". He thus decided that the "particles of the ether", whose motion presumably resulted in the phenomenon of light, must be "as nearly approaching to perfect hardness and possessing a springiness as prompt as we choose." [I.e., they are perfectly elastic rigid bodies.]

Huygens also pointed out that the particles of the ether are not arranged in straight lines, like a row of spheres "but confusedly, so that one of them touches several others", as in (a), so when B hits A it will come to a stop, while A will impart its motion to "all the spheres CCC which touch it". So each particle in the path of a wave will create its own secondary wave, a popular idea of the time for water waves, and a basic tenet of his theory. Having established this



(a)



(b)

foundation, Huygens next had to explain why light, if it is a wave motion, travels in straight lines, unlike sound. Considering a light source A and an opening BG delimited by opaque barriers HB and GI, as in (b), he points out that the motion should remain in the region bounded by CE, except that we also have to consider the fact that new waves are created at each point 'b' of BG, not to mention points 'd' further along. Huygens reasoned that the part of the motion outside the region is very feeble compared to the part along CE, the envelope of all these secondary waves (for a review of envelopes, see Addendum 8B).

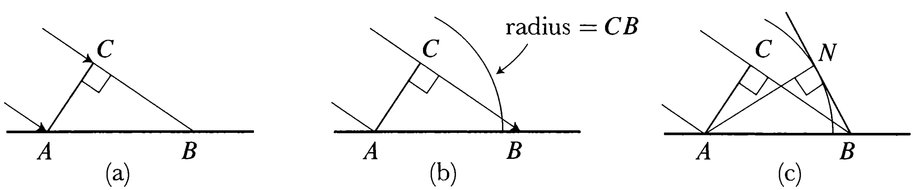
Noting that in the case of light, BG may be taken extremely small compared to the distance to the source A, "since this opening is always large enough to contain a great number of particles of the ethereal matter, which are of an inconceivable smallness", he concludes "then we may take the rays of light as if they were straight lines". At this point, the "rays" have essentially become the orthogonal trajectories of the wave fronts, which are the envelopes of all the emitted spheres. (The concept of a light "ray" had a tortuous history, which informs the whole first chapter of Buchwald [1].)

The creation of a new wave front by taking the envelope of secondary waves, "Huygens' construction", is the second basic idea of Huygens' theory, and he adds wistfully, "And all this ought not to seem fraught with too much minuteness or subtlety, since we shall see in the sequel that all the properties of Light, and everything pertaining to its reflection and its refraction, can be explained in principle by this means."

Despite the highly speculative nature of the whole discussion, which more closely resembles a stream of consciousness novel than a scientific exposition, this vague mechanistic description is actually rather concrete compared to many of the fantastical views of light until then.[1] Moreover, after this discussion, which ends the first chapter of Huygens' book, the ideas are used in the short second chapter to give a simple explanation of the law of reflection.

For simplicity, we draw diagrams in two dimensions; the proper additional considerations for three dimensions will be left to the reader—or they may be found in Huygens [2; pp. 31-32].

We consider (a) two rays, perpendicular to the wave front AC, one of which has just hit the mirror at A, while the other will not hit the mirror until it gets to B. The ray that has hit the mirror at A now starts the creation of a circle of

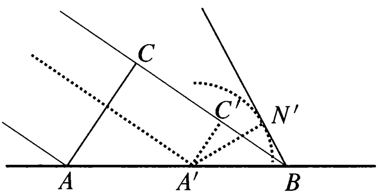


increasing radius centered at A, and at the moment (b) that the other ray hits the mirror at B, the radius of this circle will be CB. If we draw the line BN tangent to this circle, as in (c), then $AN = CB$, so the right triangles ACB and ANB

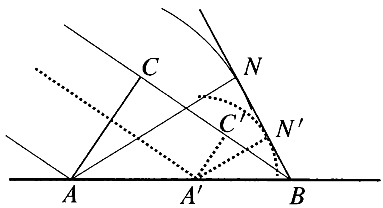
[1] In fact, the whole history of the theory of light is a comedy of errors (and perhaps a prophetic warning), involving many strange philosophical fantasies and extravagances, with one theory after another being replaced by a new theory with its own difficulties, culminating in an exquisitely detailed theory with great predictive powers, which was nevertheless soon supplanted by a different theory—see Addendum A.

with common side AB are congruent, and thus we have $\angle CBA = \angle NAB$, and also $\angle NBA = \angle CAB$.

But the same argument holds for any other ray perpendicular to the wave front, as in (d), intersecting the mirror at A', say: when our ray CB hits the



(d)



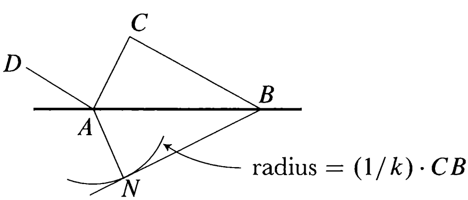
(e)

mirror at B, the tangent BN' from B to the circle started at A' will make an angle $\angle N'BA' = \angle C'A'B = \angle CAB$, and thus (e) this tangent line will be the same as the tangent line BN.

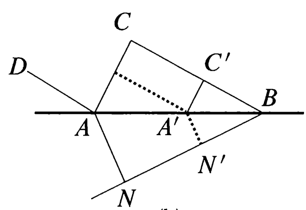
In other words, BN is the envelope of all the new circles produced by the various rays, so it is the reflected wave front, with rays making the angle $\angle NAB$ with the mirror.

The third chapter begins with further speculative arguments, which we will pass over, to show that the speed of light in denser matter should be slower, and then explains the law of refraction by the following geometric argument, using the same principles as that for the law of reflection.

We consider a wave front AC again, where the ray DA hits the boundary between the air and water at A, producing an expanding series of circles (a).



(a)



(b)

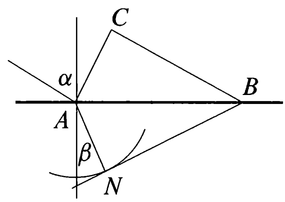
At the time that the ray CB hits the boundary at B the radius of the circle is $(1/k) \cdot CB$, where the speed of light in water v_2 is $1/k$ times v_1, the speed of light in air. As before, we draw the tangent line BN to this circle.

If we consider another ray hitting the boundary at A', as in (b), and draw $A'N'$ perpendicular to BN, then by similar triangles, $A'N' = (1/k) \cdot C'B$, which is just the radius of the circle centered at A' when the ray CB hits the boundary at B. Thus we see that BN is now the envelope of all the circles produced by the refracted rays. Finally, for the angle of incidence α and the angle of refraction β we then have

$$\sin\alpha = \sin\angle BAC = CB/AB$$

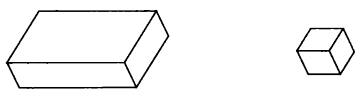
$$\sin\beta = \sin\angle ABN = AN/AB$$

$$\frac{\sin\alpha}{\sin\beta} = \frac{CB}{AN} = \frac{v_1}{v_2}.$$



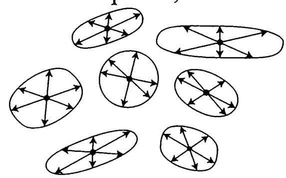
In the short fourth chapter, Huygens discusses phenomena arising from the varying density of the atmosphere (though not mirages, first widely known in Europe after Napoleon's Egypt campaign). Because of the variation in the speed of light at different levels of the atmosphere, the diameters of the secondary waves will vary as light descends toward the earth, causing the light rays to depart from straight lines, though Huygens' analysis is merely descriptive, without any formulas that would allow one to check predictions.

The longest, most difficult, and most impressive, chapter is the fifth, which is an investigation of Iceland spar, a crystal with three pairs of parallel faces, all



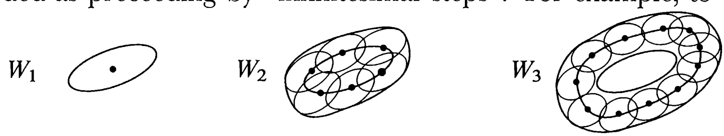
parallelograms having the same obtuse angles (101°52″); it is easily split along planes parallel to the faces, so that one can even get a piece with all six faces being congruent rhombuses. The distinctive property of Iceland spar is that it produces a double image, especially noticeable when it is placed on a printed page. Huygens discovered that a ray of light hitting a face of the crystal is split into two rays, one of which obeys Snell's law, while for the other, the ratio of the sines depends on the inclination of the initial ray, and he decided that this second ray must therefore not come from spherical surfaces but from a more general shape, for which he tried an ellipsoid of revolution. The analysis that he then gave, explaining a phenomenon earlier theories could not, is now available in Huygens [2]; a critical discussion of this analysis may be found in Appendix 1 of Buchwald [1].

Mathematically, to accommodate both "non-homogeneous" media, like the atmosphere, and "non-isotropic" media like Iceland spar, we must allow the wave fronts created at various points to have non-spherical shapes and also vary in size. We could do this by considering a Riemannian metric on some region of $\mathbb{R}^3$, with the ellipsoid representing the unit sphere at each point indicating the distance light travels in unit time in the various directions from that point. More generally, we can consider an "indicatrix" at each point, a surface bounding a body radially symmetric about the point, indicating the distance light travels



in various directions from that point in unit time; for simplicity we will assume the indicatrix is a smooth surface bounding a convex body. (By declaring vectors lying on the indicatrices to be unit vectors, we can then assign lengths to vectors at each point—basically a Finsler metric, as in DG, Vol. 2—on some region of $\mathbb{R}^3$.)

When the indicatrices vary from point to point, Huygens' construction must be regarded as preceeding by "infinitesimal steps". For example, to construct



the wave fronts for light emanating from a single point p, we first consider the wave front W_1 around p after time ε, which will just be ε times the unit sphere of our metric at p. Then we take the envelope of the wave fronts after time ε around each point of W_1 to get the wave front W_2, and then take the envelope of the wave fronts after time ε around each point of W_2 to get the wave front W_3, ... And then, finally, we take the limit of this sequence of constructions as $\varepsilon \to 0$. Huygens' explanations of reflection and refraction basically used such constructions—for the case of discontinuously changing indicatrices.

One clear reason why Huygens' theory found few adherents was that its totally geometric formulation made computations almost impossible. In addition, however, at the end of his analysis of Iceland spar, he added "one more marvellous phenomenon which I discovered after having written all the foregoing", which was inexplicable on the basis of his theory—he had basically discovered polarization, see Addendum A. Until Huygens' original picture was given a complete overhaul later on, this provided yet another strong argument for the main competition, the corpuscular theory, formulated by our final luminary.

Newton. Color had always been regarded as something different from light, perhaps a companion to light, or something that light elicited from a body when it illuminated it. This view was demolished by Newton (1642–1727), who first achieved fame around 1671 with his reports on his investigations into light, showing that white light was a mixture of the various colors. These investigations had begun around 1666, although his famous *Opticks* wasn't published until 1704, and his experimental work was carried out with such great accuracy and ability that his views, and his corpuscular theory of light, attained almost incontestable authority, especially after the publication of the *Principia.*

Newton also made many discoveries—often not further investigated—that did not fit in very well with his corpuscular theory, cleverly ending his *Opticks* with 31 'Queries', taking 60 pages, which leave the task of examining these questions to his successors, so that, as Ronchi [1] puts it, "no one could have worked better than Newton, not to build, but rather to demolish, the corpuscular theory."

At this point we merely want to note that Newton's argument for Snell's law was a mechanical analogy similar to Descartes', again involving a greater speed of light in more dense material—Newton even suggested that the light corpuscules were attracted by the denser material. Because of this, both sides of the debate between wave theories and corpuscular theories felt that accurately comparing the speed of light in water and in air (which at the time probably seemed close to hopeless) would decisively decide which theory would prevail.

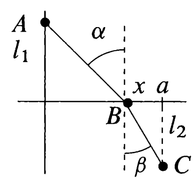
As it turns out, this determination, made independently by Foucault and by Fiseau and Breguet (cf. Tobin [1; Chap. 8]), occurred in 1850, by which time the wave theory had essentially already triumphed, though, as briefly described in Addendum A, it was greatly changed from Huygens' original conception.

In the meantime, however, Newton's authority for the speed of light being greater in denser matter led, by a circuitous route, to a most modern-looking variational principal.

Maupertuis. Maupertuis (1698–1759) shared Fermat's taste for laws of physics derived from first principles, but the basis for Fermat's argument was destroyed by Newton's theory, as Maupertuis pointedly noted in a paper of 1744, *The agreement between the different laws of Nature that had, until now, seemed incompatible.*

If one is bent on deducing the law of refraction from a variational principle that assumes the speed v_1 of light in air is less than the speed v_2 of light in water, it is not hard to concoct one. If we consider

$$\begin{aligned} v_1 \cdot AB + v_2 \cdot BC &= v_1\sqrt{l_1{}^2 + x^2} + v_2\sqrt{l_2{}^2 + (a-x)^2} \\ &= f(x), \end{aligned}$$



then setting $f'(x) = 0$ immediately gives the desired result

$$\frac{\sin\alpha}{\sin\beta} = \frac{v_2}{v_1} > 1,$$

so we minimize the sum of velocities times distances on the path from A to C.

Maupertuis justified this as follows (Dugas [1; pp. 262–263]):

> In meditating deeply on this matter, I thought that, since light has already forsaken the shortest path when it goes from one medium to another—the path which is a straight line—it could just as well not follow that of the shortest time. Indeed, what preference can there be in this matter for time or distance? Light cannot at once travel along the shortest path and along that of the shortest time—why should it go by one of these paths rather than by the other? Further, why should it follow either of these two? It chooses a path which has a very real advantage—*the path which it takes is that by which the quantity of action is the least.*
>
> It must now be explained what I mean by the quantity of action. When a body is carried from one point to another a certain action is necessary. This action depends on the velocity that the body has and the distance that it travels, but is neither the velocity nor the distance taken separately. The quantity of action is the greater as the velocity is the greater and the path which it travels is the longer. It is proportional to the sum of the distances, each one multiplied by the velocity with which the body travels along it.
>
> *It is this quantity of action which is Nature's true storehouse, and which it economises as much as possible in the motion of light.*

For a system of particles, with the velocity replaced by momentum to take into account the mass of the particles, the "action" defined by Maupertuis will be the action integral $\mathcal{A}$ (pages 464 and 466). Thus, in this roundabout way, Maupertuis succeeded in producing an incorrect principle for light that turned out to give a correct principle for mechanics—when given the proper formulation by Euler, with paths of constant energy.

Since the quantity to be minimized for Fermat's principle, the total time, is simply $(1/v_1) \cdot AB + (1/v_2) \cdot BC$, Maupertuis had basically just replaced $1/v$ by v to get a principle consistent with the assumption $v_2 > v_1$. Conflicts between the two views could thus be neatly finessed by minimizing the quantity $r_1 \cdot AB + r_2 \cdot BC$, where r_i is the "refractive index" of the i^{th} medium, leaving aside the question of whether this is proportional to velocity, for Maupertuis, or to its reciprocal, for Fermat.

Maupertuis certainly had no doubts on the matter. Unconstrained by the defect of false modesty, he wrote much later in his *Essai de Cosmologie* of 1751:

> After so many great men have worked on this matter, I hardly dare say I have discovered the principle on which all the laws of motion are founded; ... Our principle, more in conformity with the ideas of things

> that we should have, leaves the world in its natural need of the Creator, and is a necessary result of the wisest doing of that same power. ... What satisfaction for the human mind, in contemplating these laws—so beautiful and so simple—that they may be the only ones that the Creator and the Director of things has established in matter in order to accomplish all the phenomena of the visible world.

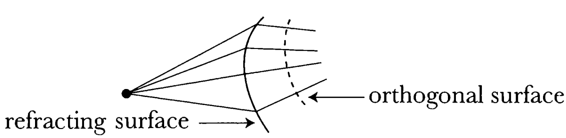
To bring the story full-circle, we temporarily trespass into the domain of modern physics. Maupertuis' action has dimensions Mass × Velocity × Length, or $\mathsf{MVL} = \mathsf{ML}^2\mathsf{T}^{-1}$, but nowadays usually written as $\mathsf{MV{\cdot}VT} = \mathsf{MV}^2\mathsf{T} = \mathsf{ET}$ (where E = Energy), like the fundamental constant $h = 6.626 \cdot 10^{-27}$erg sec, which Planck introduced to solve a very specific classical problem, but which eventually became a cornerstone of quantum mechanics, somewhat to his chagrin.

And, undoubtedly to the chagrin of many others, Planck, like Maupertuis, believed that an argument for the existence of a supreme being can be found in teleological principles, with special reference to the principle of least action (except that Planck was using this term to refer not to Maupertuis' principle of least action, but to what we now call Hamilton's principle)—see his lengthy essay "Religion and Natural Science" in Planck [1].

Malus. After Kepler, Galileo, Descartes, Fermat, Huygens, and Newton, with a detour to Maupertuis, we come to a much lesser known scientist whose work will provide us with a link between the study of light and modern mechanics, Etienne-Louis Malus (1775–1812), never mentioned in mechanics books, and not all that often in texts on optics.

He receives much more attention in Buchwald [1, pp. 23ff.], which describes the horrific circumstances leading to his study of optics, emphasizes his role in turning experimental optics from an activity using hardly any mathematics and unconcerned with obtaining or presenting hard data, into a modern science, and describes his work in detail, especially his work on polarization.

A graduate of the newly founded Ecole Polytechnique in Paris, where the chemistry of the time was one of the subjects of instruction, his original hope was to prove that light is a compound of caloric and oxygen (I kid you not), but happily the first paper he presented in Paris was not concerned with the physical structure of light at all. He considered rays of light emanating from a single point that encounter the surface of a body that either reflects or refracts them, and proved that the resulting rays are all normal to some surface. Though sounding



similar to Huygens' ideas, this surface was not thought of as a wave front, but was simply a mathematical entity that happened to exist as a consequence of the standard laws of reflection and refraction, and Malus's proof was a complicated analytic derivation from these laws.

The interest for optics was presumably in following the rays in the opposite direction, to deduce that they can all be focused at a point only if they were all orthogonal to some surface. In any case, Malus' paper was apparently received very favorably, and must have generalized many specific optical investigations, for the committee of Laplace, Lagrange, Monge, and Lacroix chosen by the Académie des Sciences to read the paper concluded:

> To apply thus, without any limitation on its generality, calculation to phenomena;—to deduce, from a single consideration of a very general kind, all the solutions which before were only obtained from particular considerations,—is truly to write a treatise on analytical optics, which, considering the whole science in a single point of view, cannot but contribute to the extension of its domain.

Malus did not try to show that if this new collection of rays undergoes another reflection or refraction then there will again be a surface normal to the resulting rays, let alone the more general claim that any system of rays orthogonal to a surface becomes another such system under a reflection or refraction. In fact, an error in his reasoning led Malus to the conclusion that his result definitely could not be extended to a second reflection or refraction.

But in 1816 Dupin proved the result for reflection of any given system of rays orthogonal to a surface (which so upset the Académie that they appointed a special investigative committee, which discovered Malus' error), and he conjectured that it was also true in the case of refractions.

Proofs of this were given in 1825 by Quetelt and Gergonne (further details may be found in Atzema [1]), establishing "Malus' theorem" for any number of reflections and refractions.

* * *

Fortunately for the science of mechanics, all these later papers were unknown to an Irish mathematical prodigy who had read Malus' paper with interest.

ADDENDUM 15A

BATTLING TO A DRAW

"Geometrical optics" is the term for the analysis of light that depends only on the laws of reflection and refraction—particularly important for the design of mirrors and lenses—by investigating the way that a family of light rays behaves as it hits a mirror or passes through a lens, ignoring as much as possible the many subtle phenomena concerning light.

Geometrical optics is still an important and challenging field of study, but in the first half of the 19th century a theory of light was finally created that could explain the various subtleties, a wave theory of light that eventually totally displaced all previous theories, including Huygens' wave theory, though aspects of Huygens' theory were incorporated into this new theory.

The first important phenomenon about light that presented a challenge was *diffraction*, the fact that light does not travel in exact straight lines, but does bend a bit, just like sound waves, although observing this is much more difficult because of the short wave lengths involved, and many complicated effects are observed, especially because different colors diffract differently. It was first discovered, probably accidentally, by Father Francesco Maria Grimaldi (1618–1663), whose book on the subject of light was published in 1665, and ignored for a long time, though his work involving colors may have instigated Newton's original experiments.

On the other hand, the second important phenomenon that has to be accounted for is the "marvellous phenomenon" that Huygens discovered about Iceland spar (page 495). He noted that when a ray of light was split into two rays as it passed through the crystal, these rays no longer acted like ordinary light rays in terms of their passage through another crystal, and in particular acted quite differently when the second crystal was perpendicular to the first rather than parallel to it. This phenomenon of polarization, where light rays seemed to have a specific orientation in the plane perpendicular to the direction of their propagation, seemed wholly incompatible with Huygens' picture of wave prorogation.

Thus Newton's corpuscular theory had remained dominant, though experts knew of all sorts of troubling difficulties, until the beginning of the 19th century, when the third important phenomenon was discovered. In 1802, the English physician Thomas Young (1773–1829) published his first account of new experiments that called the corpuscular theory into question, and he soon produced the famous Young interference experiment, where light from a point source passing through two small holes close to each other produces an interference pattern, from which Young could even calculate wave lengths of different colors, these waves being assumed transversal to the direction of propagation, rather than longitudinal like sound waves, to which light had so often been compared over the centuries. These heretical ideas were naturally denounced by the stalwart defenders of the Newtonian dogma, and his ideas were roundly derided.

Fortunately, somewhat similar ideas put forth at about the same time by the French civil engineer Augustin Fresnel (1788–1827) elicited a much more encouraging reception, and Fresnel rapidly began to develop a complex new theory involving intricate calculations, also invigorating Young to add further contributions. The assumption of transversal waves enabled the phenomenon of polarization to be accounted for, and Fresnel was able to account for diffraction phenomena by adding in Huygens' construction, together with the complicating factor that the various secondary waves would interfere with each other.

Thus, in the brief space of the quarter century between Young's initial paper and Fresnel's early death, an English physician and a French civil engineer succeeded in creating a detailed elaborate theory destined to quickly displace all previous theories, and the experiments carried out in 1850, establishing that the speed of light in air is greater than the speed in water, were simply the coup de grâce that finally silenced the last stubborn adherents of Newton's theory.

And then the whole apparatus of the new wave theory was subsumed, at least theoretically, around 1864, by Maxwell's theory of electromagnetic waves, with investigations by Kirchhoff and others establishing that optics may be deduced as the limiting case of very small wave lengths, and the entire idea of secondary waves was upended (though calculations using Fresnel's theory are still often the only reasonable mathematical approach to finding solutions). One can now look at the telescope of history from the other end, as in Born and Wolf [1], which begins with Maxwell's equations on page 1, and then goes on to give detailed and thorough discussions of a large class of optical phenomena from the point of view of the wave equation.

That wasn't the end, of course, since modern quantum theory says that light is a wave *and* a particle. Perhaps some might still say, with Grimaldi, "let us be honest we do not really know anything about the nature of light and it is dishonest to use big words [or equations] which are meaningless."

ADDENDUM 15B

HUYGENS' PRINCIPLE

We have followed Born and Wolf [1] in referring to "Huygens' construction", because once the wave equation became the basis for the theory of light it was possible to pose explicit mathematical questions, quite unrelated to Huygens' original mechanistic ideas, which led to results about partial differential equations that now, somewhat inexplicably, have Huygens' name attached to them.

In Chapter 8 we considered the 1-dimensional wave equation

$$u_{tt} - v^2 u_{xx} = 0,$$

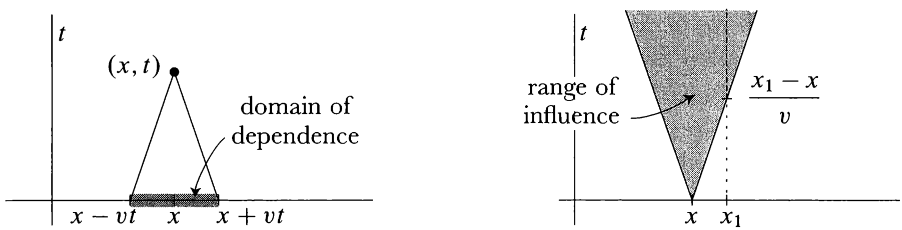
and (Problem 8-5) we found that in terms of initial conditions

$$u(x,0) = \phi(x), \qquad u_t(x,0) = \psi(x)$$

we have d'Alembert's formula

$$u(x,t) = \frac{\phi(x+vt) + \phi(x-vt)}{2} + \frac{1}{2v}\int_{x-vt}^{x+vt} \psi(s)\,ds.$$

This formula shows (a) that for a point x and a time t, the value of $u(x,t)$ depends on the values of ϕ and ψ on the interval $[x - vt, x + vt]$, the *domain of dependence* of (x,t). Or (b), we can consider the *range of influence* of x, the



(a) (b)

set of all points (x_1, t) whose domain of dependence contains x. The inverted triangular shape indicates a *finite propagation speed*: for $x_1 > x$ the value of $u(x,t)$ at points (x_1,t) with $t < (x_1 - x)/v$ is completely independent of $\phi(x)$ and $\psi(x)$. On the other hand, notice that the value of $u(x_1,t)$ for *all* $t \geq (x_1 - x)/v$ does depend on these values; in a 1-dimensional world a brief noise at x at time $t = 0$ will continue to have reverberations at x_1 forever after time $(x_1 - x)/v$, which is certainly not what we experience in our 3-dimensional world.

There is a neat analytical trick, Poisson's method of *spherical means*, that enables us to obtain the solution for the 3-dimensional wave equation,

$$u_{tt} = v^2 \Delta u = v^2(u_{xx} + u_{yy} + u_{zz}),$$

from those of the 1-dimensional wave equation. We follow the exposition in McOwen [1] pretty closely.

(a) **Spherical means.** For a continuous function f on $\mathbb{R}^n$, its spherical mean over a sphere with center x and radius r is

$$M_f(x,r) = \frac{1}{\omega_{n-1}} \int_{|\xi|=1} f(x + r\xi)\, dS(\xi), \tag{1}$$

where dS denotes the $(n-1)$-dimensional volume element on the unit sphere $S^{n-1} \subset \mathbb{R}^n$, and ω_{n-1} is the total $n-1$-dimensional volume of S^{n-1}. Although $M_f(x,0)$ isn't defined, it is easy to see that

$$\lim_{r\to 0} M_f(x,r) = f(x). \tag{2}$$

From (1) we obtain

$$\frac{\partial}{\partial r} M_f(x,r) = \frac{1}{\omega_{n-1}} \int_{|\xi|=1} \sum_{i=1}^{n} f_{x_i}(x + r\xi)\xi_i \, dS(\xi),$$

and the integrand can be written as $\langle X, \nu\rangle$, where ν is the unit outward normal on the unit n-disk and X is the vector field

$$X = \left(\frac{\partial f(x + r\xi)}{\partial x_1}, \dots, \frac{\partial f(x + r\xi)}{\partial x_n}\right).$$

The divergence of X (page 138) is

$$\operatorname{div} X = r \sum_{i=1}^{n} f_{x_i x_i}(x + r\xi) = r\Delta_x f(x + r\xi),$$

so the divergence theorem (page 139) gives

$$\frac{\partial}{\partial r} M_f(x,r) = \frac{r}{\omega_{n-1}} \Delta_x \int_{|\xi|\le 1} f(x + r\xi)\, d\xi. \tag{3}$$

The substitution $\xi' = r\xi$, $d\xi' = r^n d\xi$ gives

$$\int_{|\xi|\le 1} f(x + r\xi)\, d\xi = \frac{1}{r^n} \int_{|\xi'|\le 1} f(x + \xi')\, d\xi',$$

and in spherical coordinates we can write

$$\frac{1}{r^n}\int_{|\xi'|\le 1} f(x+\xi')\,d\xi' = \frac{1}{r^n}\int_0^r \rho^{n-1}\int_{|\xi|=1} f(x+\rho\xi)\,dS(\xi)\,d\rho$$
$$= \frac{\omega_{n-1}}{r^n}\int_0^r \rho^{n-1} M_f(x,\rho)\,d\rho.$$

Substituting back into (3) we obtain

$$\frac{\partial}{\partial r}M_f(x,r) = \frac{1}{r^{n-1}}\Delta_x\int_0^r \rho^{n-1}M_f(x,\rho)\,d\rho.$$

Finally, multiplying by r^{n-1}, taking $\partial/\partial r$, and dividing by r^{n-1}, we get the *Darboux equation*

(4) $$\left(\frac{\partial^2}{\partial r^2} + \frac{n-1}{r}\frac{\partial}{\partial r}\right)M_f(x,r) = \Delta_x M_f(x,r).$$

(b) Application to the wave equation. Now consider the equation

(5) $$u_{tt} = v^2\Delta u$$

on $\mathbb{R}^n$ for $t>0$, with the initial conditions

(6) $$u(x,0) = \phi(x), \qquad u_t(x,0) = \psi(x).$$

If u satisfies (5), then

$$\frac{\partial^2}{\partial t^2}M_u(x,r,t) = \frac{1}{\omega_{n-1}}\int_{|\xi|=1} u_{tt}(x+r\xi,t)\,dS(\xi)$$
$$= \frac{1}{\omega_{n-1}}\int_{|\xi|=1} v^2\Delta u(x+r\xi,t)\,dS(\xi)$$
$$= v^2\Delta M_u(x,r,t),$$

and (4) implies the *Euler-Poisson-Darboux equation*

(7) $$\frac{\partial^2}{\partial t^2}M_u(x,r,t) = v^2\left(\frac{\partial^2}{\partial r^2} + \frac{n-1}{r}\frac{\partial}{\partial r}\right)M_u(x,r,t),$$

while the initial conditions (6) give

(8) $$M_u(x,r,0) = M_\phi(x,r), \qquad \frac{\partial M_u}{\partial t}(x,r,0) = M_\psi(x,r).$$

If we can solve (7) and (8) to find $M_u(x,r,t)$, then by (1), we will have

$$u(x,t) = \lim_{r\to 0} M_u(x,r,t).$$

(c) **The wave equation in dimension 3**. When $n = 3$, a little calculation shows that equation (7) works out to be equivalent to

$$\frac{\partial^2}{\partial t^2} rM_u(x,r,t) = v^2 \frac{\partial^2}{\partial r^2} rM_u(x,r,t). \tag{9}$$

This means that for each x, the function

$$U^x(r,t) = rM_u(x,r,t)$$

is a solution of the 1*-dimensional wave equation* for $r, t > 0$:

$$\frac{\partial^2}{\partial t^2} U^x(r,t) = v^2 \frac{\partial^2}{\partial r^2} U^x(r,t),$$

while (8) shows that we have

$$\begin{aligned} U^x(r,0) &= rM_\phi(x,r) = \Phi^x(r), \quad \text{say,} \\ U^x{}_t(r,0) &= rM_\psi(x,r) = \Psi^x(r), \quad \text{say,} \end{aligned}$$

and equation (2) gives

$$U^x(0,t) = \lim_{r\to 0} rM_u(x,r,t) = 0 \cdot u(x,t) = 0.$$

Since $\Phi^x(0) = 0 = \Psi^x(0)$, Problem 8-5 (b) shows that if we extend Φ^x and Ψ^x to be odd functions of r, we can use d'Alembert's formula to obtain

$$U^x(r,t) = \frac{\Phi^x(r+vt) + \Phi^x(r-vt}{2} + \frac{1}{2v}\int_{r-vt}^{r+vt} \Psi^x(\rho)\,d\rho.$$

Since Φ^x and Ψ^x are odd functions, for $r < vt$ we have

$$\Phi^x(r-vt) = -\Phi^x(vt-r), \qquad \int_{r-vt}^{r+vt} \Psi^x(\rho)\,d\rho = \int_{vt-r}^{vt+r} \Psi^x(\rho)\,d\rho,$$

and thus

$$\begin{aligned} M_u(x,r,t) = \frac{1}{r}U^x(r,t) &= \frac{\Phi^x(vt+r) - \Phi^x(vt-r)}{2r} + \frac{1}{2vr}\int_{vt-r}^{vt+r} \Psi^x(\rho)\,d\rho \\ &= \frac{(vt+r)M_\phi(x,vt+r) - (vt-r)M_\phi(x,vt-r)}{2r} + \frac{1}{2vr}\int_{vt-r}^{vt+r} \rho M_\psi(x,\rho)\,d\rho. \end{aligned}$$

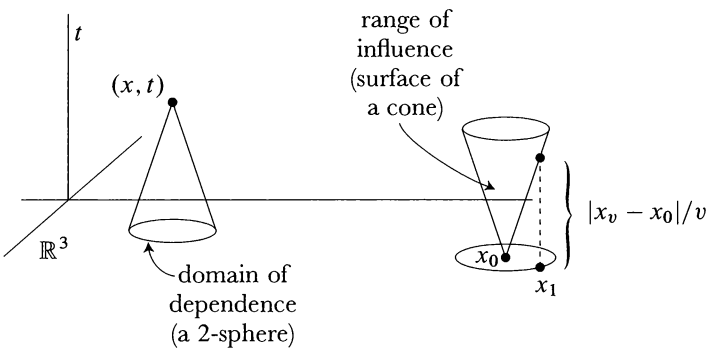
Letting $r \to 0$ we obtain

$$u(x,t) = \left.\frac{\partial(\tau M_\phi(x,\tau)}{\partial \tau}\right|_{\tau=vt} + tM_\psi(x,vt)$$
$$= \frac{\partial}{\partial t}(tM_\phi(x,vt)) + tM_\psi(x,vt),$$

leading to *Kirchhoff's formula*

$$u(x,t) = \frac{1}{4\pi}\frac{\partial}{\partial t}\left(t\int_{|\xi|=1} \phi(x+vt\xi)\,dS(\xi)\right) + \frac{t}{4\pi}\int_{|\xi|=1} \psi(x+vt\xi)\,dS(\xi). \tag{10}$$

From Kirchhoff's formula, we obtain the following picture for the domain of dependence and range of influence in the case of the 3-dimensional wave



equation, in contrast to the picture on page 502 for the 1-dimensional wave equation. We again have a finite propagation speed, but now we have *sharp signals.* Noise at x_0 at time 0 is experienced at x_1 only at time $|x_1 - x_0|/v$.

(d) The wave equation in dimension 2. Although this trick doesn't work for the 2-dimensional wave equation, we can analyze that case by another trick, Hadamard's *method of descent*, where we view the 2-dimensional wave equation as the special case of the 3-dimensional problem where the initial conditions are independent of x_3.

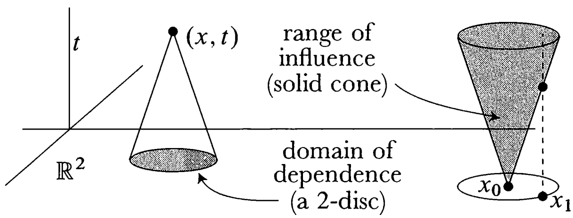
Letting ξ_1, ξ_2 be coordinates in $\mathbb{R}^2$, the upper half of the unit sphere in $\mathbb{R}^3$ is the graph of $f(\xi_1,\xi_2) = \sqrt{1-\xi_1{}^2-\xi_2{}^2}$, so we can write

$$dS(\xi, f(\xi)) = \sqrt{1+(\partial f/\partial\xi_1)^2+(\partial f/\partial\xi_2)^2}\;d\xi_1\,d\xi_2 = \frac{d\xi_1\,d\xi_2}{\sqrt{1-\xi_1{}^2-\xi_2{}^2}}.$$

To use (10) we just have to multiply the integral over the upper half of the unit sphere by 2, to obtain

$$u(x_1,x_2,t) = \frac{1}{4\pi}\frac{\partial}{\partial t}\left(2t\int_{\xi_1{}^2+\xi_2{}^2\le 1}\frac{\phi(x_1+ct\xi_1,x_2+ct\xi_2)\,d\xi_1\,d\xi_2}{\sqrt{1-\xi_1{}^2-\xi_2{}^2}}\right)$$
$$+\frac{t}{4\pi}\left(2\int_{\xi_1{}^2+\xi_2{}^2\le 1}\frac{\psi(x_1+ct\xi_1,x_2+ct\xi_2)\,d\xi_1\,d\xi_2}{\sqrt{1-\xi_1{}^2-\xi_2{}^2}}\right).$$

From this formula we now get the following picture for the domain of dependence and range of influence in the case of the 2-dimensional wave equation,



where, once again, a brief disturbance will continue to be felt after it first reaches another point. A physical example is given by water waves. Dropping a pebble into a large bowl of water produces expanding circular waves that continue to be felt after the first time they reach another point.

It turns out (McCowen [1; §3.2 e]) that this pattern persists for all $n > 1$: we have sharp signals for odd n but not for even n. This fact is sometimes referred to as *Huygens' Principle*, which would seem to be an overly generous attribution. Sometimes the use of the term Huygens' Principle is meant to bring attention to the domain of dependence, rather than the range of influence. For example, in three dimensions, the fact that the domain of dependence is the *surface* of a sphere, rather than the 3-disk it bounds, might be regarded as related to Huygens' construction of a new wave front by taking the envelope of secondary waves along the old wave front, although the formulas deduced so far say nothing at all about envelopes (and it isn't quite clear from Huygens' descriptions whether he would have regarded the old wave front as the entire domain of dependence). In fact, in connection with this interpretation one will sometimes find the statement that although Huygens' Principle is true in three and higher odd dimensions, it is false in even dimensions, as if Huygens had pronounced a principle that happened to be correct by a lucky accident.

On the other hand, there is a theorem, true in all dimensions, that really does formalize Huygens' construction, which we will briefly discuss later, in Addendum 18B.

PART IV

HAMILTONIAN MECHANICS

FROM ARAGONITE TO THE SCHRÖDINGER WAVE EQUATION

CHAPTER 16

THE COTANGENT BUNDLE

After Lagrangian mechanics, which takes place on the tangent bundle TM of a manifold, we move to Hamiltonian mechanics, which occurs on the cotangent bundle T^*M, where each tangent space M_p is replaced by its dual space $M_p{}^*$. For convenience, we use the same map π to indicate the projection $\pi\colon T^*M \to M$ as we used for TM. We generally follow the notation of DG, with some special additions for mechanics.

For a coordinate system $(q^1, \dots, q^n)$ on $U \subset M$, the bases $\partial/\partial q^i$ for the fibres of TM over U determine dual bases for the fibres of T^*M over U, which we will denote by $p_1, \dots, p_n$, so that for $\lambda \in \pi^{-1}(U) \subset T^*M$ we have

$$\text{(a)} \qquad p_i(\lambda) = \lambda(\partial/\partial q^i).$$

On T^*M we then have the coordinate system $(q^1 \circ \pi, \dots, q^n \circ \pi, p_1, \dots, p_n)$, which for the moment will be scrupulously denoted by $(\bar{q}^1, \dots, \bar{q}^n, p_1, \dots, p_n)$.

Special features of the cotangent bundle. For each coordinate system $(\bar{q}, p) = (\bar{q}^1, \dots, \bar{q}^n, p_1, \dots, p_n)$ on T^*M we can write down the 1-form

$$\boldsymbol{\theta} = \sum_{i=1}^{n} p_i \, d\bar{q}^i.$$

This 1-form $\boldsymbol{\theta}$ will be very important in the sequel, and it turns out to be independent of the coordinate system $(\bar{q}, p)$. This can be checked by a direct, slightly confusing, calculation in coordinates, but we can simply note that the definition of the p_i in equation (a) can be written as

$$\lambda = \sum_{i=1}^{n} p_i(\lambda) \, dq^i(\pi(\lambda)),$$

and thus for a tangent vector X on T^*M at the "point" $\lambda \in T^*M$, we have

$$\lambda(\pi_* X) = \sum_{i=1}^{n} p_i(\lambda)(\pi_* X) \, dq^i(\pi_* X) = \sum_{i=1}^{n} p_i(X) \, d\bar{q}^i(X) = \boldsymbol{\theta}(X),$$

so that we have the invariant definition

$$\boldsymbol{\theta}(X) = \pi(X)(\pi_* X).$$

If this also seems slightly confusing, one can just check that this invariantly defined 1-form $\boldsymbol{\theta}$ is equal to $\sum\limits_{i=1}^{n} p_i\, d\bar{q}^i$ by noting that

$$\text{they both give } 0 \text{ for } X = \left.\frac{\partial}{\partial p_i}\right|_\lambda, \quad \text{since } \pi_*\left(\frac{\partial}{\partial p_i}\right) = 0,$$

$$\text{they both give } p_i(\lambda) \text{ for } X = \left.\frac{\partial}{\partial \bar{q}^i}\right|_\lambda, \quad \text{since } \pi_*\left(\frac{\partial}{\partial \bar{q}^i}\right) = \frac{\partial}{\partial q^i}.$$

Having settled this matter, requiring a careful distinction between q^i and $\bar{q}^i$, we henceforth generally ignore the distinction and write $\boldsymbol{\theta} = \sum_{i=1}^n p_i\, dq^i$. Using $\boldsymbol{\theta}$, we now define the even more important 2-form $\boldsymbol{\omega}$ by

$$\boldsymbol{\omega} = d\boldsymbol{\theta} = \sum_{i=1}^{n} dp_i \wedge dq^i .$$

It is easily checked that

$$\begin{aligned}\boldsymbol{\omega}\left(\sum_{j=1}^{n} A_j \frac{\partial}{\partial q^j} + B_j \frac{\partial}{\partial p_j},\ \sum_{k=1}^{n} C_k \frac{\partial}{\partial q^k} + D_k \frac{\partial}{\partial p_k}\right)& \\ = \left[\sum_{i=1}^{n} C_i\, dp_i - D_i\, dq_i\right]&\left(\sum_{j=1}^{n} A_j \frac{\partial}{\partial q^j} + B_j \frac{\partial}{\partial p_j}\right).\end{aligned}$$

It follows, in particular, that $\boldsymbol{\omega}$ is **nondegenerate**: for X on T^*M, if $\boldsymbol{\omega}(X,Y) = 0$ for all Y, then $X = 0$. For the **interior product** or **contraction** $X \lrcorner\, \boldsymbol{\omega}$ defined by

$$X \lrcorner\, \boldsymbol{\omega}\ (Y) = \boldsymbol{\omega}(X,Y)$$

we can say that $X \mapsto X \lrcorner\, \boldsymbol{\omega}$ is one-one from the tangent bundle $T(T^*M)$ of T^*M to its cotangent bundle $T^*(T^*M)$. The notation $i_X\boldsymbol{\omega}$ is often used for $X \lrcorner\, \boldsymbol{\omega}$, but we will stick with the "hook" notation.

Note that since $\boldsymbol{\omega} = d\boldsymbol{\theta}$, we have $d\boldsymbol{\omega} = 0$, which also follows directly from the expression $\boldsymbol{\omega} = \sum_{i=1}^n dp_i \wedge dq^i$.

WARNING. Sometimes $\boldsymbol{\omega}$ is given the opposite sign, $\boldsymbol{\omega} = \sum_{i=1}^n dq^i \wedge dp_i$, which affects multitudes of additional formulas later on.

It is not hard to see that the n-fold wedge product $\boldsymbol{\omega} \wedge \cdots \wedge \boldsymbol{\omega}$ can be written as

$$\boldsymbol{\omega} \wedge \cdots \wedge \boldsymbol{\omega} = (-1)^N n!\, dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n,$$

for some N (the precise value is $N = n(n+1)/2$, but who's counting?). This is non-zero everywhere, so we can use $(1/n!)(-1)^N \boldsymbol{\omega} \wedge \cdots \wedge \boldsymbol{\omega}$ as a volume element on T^*M, which reduces to the standard volume on $\mathbb{R}^{2n}$ when $M = \mathbb{R}^n$; in particular, T^*M is always orientable.

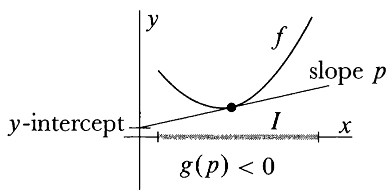
The Legendre transform. A mathematical device now quite important for mechanics appeared in the paper Legendre [1] of 1787, in connection with the partial differential equation for minimal surfaces. Many differential geometry books can be consulted for proofs that (1) the graph of $f\colon \mathbb{R}^2 \to \mathbb{R}$ will be a critical point for the area function if and only if it has mean curvature $H = 0$, and (2) this is equivalent to the equation

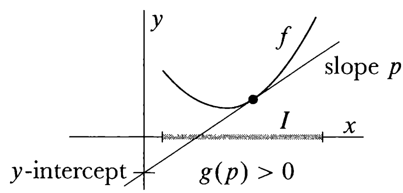
$$(1 + f_y{}^2) f_{xx} - 2 f_x f_y f_{xy} + (1 + f_x{}^2) f_{yy} = 0,$$

where subscripts denote partial derivatives.[1]

Legendre introduced a way of transforming a partial differential equation of this sort into a simpler one by using the partial derivatives of the function f as the new variables. Though often stated as a formal relationship, the definition of the Legendre transformation can most easily be understood geometrically. We start with the simplest case of a function $f\colon \mathbb{R} \to \mathbb{R}$.

Suppose $f'' > 0$ on an interval $I \subset \mathbb{R}$, so that f is convex on this interval, with monotonically increasing derivative f' having values in the interval $f'(I)$. Any number $p \in f'(I)$ is then the derivative of f at some unique point in I, so there is a unique tangent line to the graph of f with slope p. We let $g(p)$ be the negative of the y-intercept of this tangent line. We thus obtain a new





function $g\colon f'(I) \to \mathbb{R}$, the *Legendre transform* of f, which we will sometimes write as $\mathcal{L}f$ or $\mathcal{L}(f)$ [there is no even remotely standard notation for g].

Since the x-coordinate of the point of tangency is $(f')^{-1}(p)$, and thus the y-coordinate is $f\big((f')^{-1}(p)\big)$, the tangent line is the graph of the equation $y - f\big((f')^{-1}(p)\big) = p \cdot [x - (f')^{-1}(p)]$. The y-intercept is the value of y for $x = 0$, and the negative, $g(p)$, is

$$\text{(a)} \qquad g(p) = p \cdot (f')^{-1}(p) - f\big((f')^{-1}(p)\big).$$

This is often stated, for computational purposes, and for those allergic to inverse functions, as

$$\text{(a}'\text{)} \qquad g(p) = p\bar{x} - f(\bar{x}) \qquad \begin{array}{l}\text{for the (unique) } \bar{x} \in I \\ \text{with } f'(\bar{x}) = p.\end{array}$$

---

[1] DG is probably one of the worst references; (2) occurs in Vol. 3, pg. 137, and (1) in Vol. 4, pg. 262! Sic semper "Comprehensive".

Since $f'' > 0$ implies that $(f')^{-1}$ is differentiable everywhere, we can differentiate (a) to obtain

$$g'(p) = (f')^{-1}(p) + p \cdot \big((f')^{-1}\big)'(p) - f'\big((f')^{-1}(p)\big) \cdot \big((f')^{-1}\big)'(p),$$

with the second and third terms canceling, so that

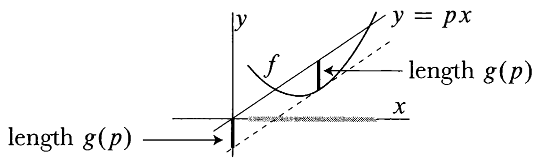
$$\text{(b)} \qquad g'(p) = (f')^{-1}(p).$$

It follows that $g'' > 0$, so we can consider the Legendre transform h of g. And, as (b) might suggest, the Legendre transformation is involutive, $\mathcal{L}(\mathcal{L}(f)) = f$:

$$\begin{aligned} h(x) &= x \cdot (g')^{-1}(x) - g\big((g')^{-1}(x)\big) \\ &= xf'(x) - g(f'(x)) && \text{by (b)} \\ &= xf'(x) - f'(x) \cdot (f')^{-1}(f'(x)) + f\big((f')^{-1}(f'(x))\big) && \text{by (a)} \\ &= f(x). \end{aligned}$$

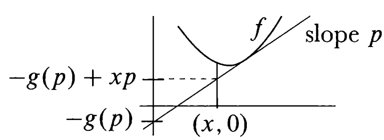
A more revealing argument can be given using an equivalent geometric definition (the original form, essentially due to Fenchel [1]). Note that $g(p)$ can also be described as the maximum vertical distance of the graph of f from the line $y = px$, that is, as the maximum of the function

$$\phi(x) = px - f(x).$$



Naturally this maximum is at the point $\bar{x}$ where $0 = \phi'(\bar{x})$, i.e., $f'(\bar{x}) = p$, so $g(p) = \phi(\bar{x}) = p\bar{x} - f(\bar{x})$, which is just equation (a′).

Now if $h(x)$ is the Legendre transform of $g(x)$, then $h(x)$ is the maximum value of $xp - g(p)$ for all slopes p of tangent lines to f. But for the tangent line of f with slope p, the y-intercept is $-g(p)$, by the first definition of g, so the number $-g(p) + xp$ is the y-coordinate of the intersection of that tangent



line with the vertical line through $(x, 0)$. But the tangent lines all lie below the graph of f, so all values of $xp - g(p)$ are $\le f(x)$, and in fact the maximum value, $f(x)$, is obtained for $p = f'(x)$. **Q.E.D.**

The definition of the Legendre transform can be generalized to the case of a function $f : U \to \mathbb{R}$ for $U \subset \mathbb{R}^n$. Instead of f' we now have the derivative $Df = (D_1 f, \dots, D_n f)$ where the D_i denote the partial derivatives $\partial/\partial x_i$, and now $Df : U \to \mathbb{R}^n$. We want to assume that Df is one-one, with a differentiable inverse, which is always the case if the Jacobian matrix $\left(\partial^2 f/\partial x_i \partial x_j\right)$ is everywhere positive definite (Problem 5). The Legendre transform g of f is defined on the set of $p = (p_1, \dots, p_n)$ for which we have $p_i = \partial f/\partial x_i(\bar{x})$ for some unique $\bar{x} \in U$, with $g(p)$ defined as the negative of the y-intercept of the tangent n-plane to the graph of f at $\bar{x}$.

The formula analogous to (a) is

$$\text{(A)} \qquad \begin{aligned} g(p) &= \langle p, (Df)^{-1}(p)\rangle - f\big((Df)^{-1}(p)\big) \\ &= \sum_{i=1}^{n} p_i \cdot [(Df)^{-1}(p)]_i - f\big((Df)^{-1}(p)\big), \end{aligned}$$

probably most easily seen by using the alternate description of $g(p)$ as the maximum vertical distance of the graph of f from the n-plane with equations $y_i = p_i x$: analogous to (a′) we have

$$\text{(A}'\text{)} \quad g(p) = \sum_{j=1}^{n} p_j \cdot \bar{x}_j - f(\bar{x}) \qquad \begin{array}{l} \text{for the (unique) } \bar{x} \in U \\ \text{with } Df(\bar{x}) = p \text{, i.e., such that} \\ \bar{x} \text{ is a critical point of } \sum p_j x_j - f(x). \end{array}$$

From (A) we easily derive the analogue of (b),

$$\text{(B)} \qquad Dg(p) = (Df)^{-1}(p)$$

and then prove in a straightforward way that $\mathcal{L}$ is involutive.

The classical definitions of the Legendre transformation were not geometric in nature, and were basically presented as manipulations for writing an expression in one set of variables as a new expression in terms of some other set. To see the use to which Legendre put this transformation, and also deal with some more standard classical notation, suppose we have a function that we write as $f(x, y)$, satisfying the minimal surface equation

$$(1 + {f_y}^2) f_{xx} - 2 f_x f_y f_{xy} + (1 + {f_x}^2) f_{yy} = 0.$$

To simplify this nonlinear equation, we want to get rid of the f_x and f_y terms, which is just what the Legendre transform, which we write as $g(\xi, \eta)$, will accomplish, since the condition on $\bar{x}$ in (A′) will now be written as

$$\xi = f_x, \quad \eta = f_y.$$

Since the Legendre transformation is involutive, we also have

$$x = g_\xi, \quad y = g_\eta.$$

Of course, we are leaving out the arguments for the functions, in the classical manner, trusting that everything will fall into place on its own.

Differentiating $\xi = f_x$ we obtain

$$1 = \frac{\partial f_x}{\partial \xi} = \frac{\partial f_x}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial f_x}{\partial y}\frac{\partial y}{\partial \xi} = f_{xx}g_{\xi\xi} + f_{xy}g_{\xi\eta},$$

together with three other equations. All four can be written together as

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}\begin{pmatrix} g_{\xi\xi} & g_{\xi\eta} \\ g_{\xi\eta} & g_{\xi\xi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

from which we obtain

$$\left.\begin{aligned} f_{xx} &= D \cdot g_{\eta\eta} \\ f_{xy} &= -D \cdot g_{\eta\xi} \\ f_{yy} &= D \cdot g_{\xi\xi}, \end{aligned}\right. \qquad \text{for } D = f_{xx}f_{yy} - {f_{xy}}^2.$$

When we apply these general equations to the minimal surface equation we immediately obtain

$$(1+\xi^2)g_{\xi\xi} + 2\xi\eta g_{\xi\eta} + (1+\eta^2)g_{\eta\eta} = 0,$$

a linear second order equation. Addendum A gives another example, involving first order equations.

We now want to consider the Legendre transform of a function $f\colon V \to \mathbb{R}$ for a real n-dimensional vector space V (or an open subset of V). Recall that for each $v \in V$ we have the derivative $Df(v) \in V^*$. Invariantly defined, $Df(v)$ is just the best linear approximation to $f(v+h) - f(v)$ at $h = 0$, i.e., the $\lambda \in V^*$ for which

$$|f(v+h) - f(v) - \lambda(h)| = o(|h|) \qquad h \in V$$

for any norm $|\ \ |$ on V. This map Df is exactly the same as the Df on page 515 when we use a basis $v_1, \dots, v_n$ of V to identify V with $\mathbb{R}^n$, and the dual basis ${v_1}^*, \dots, {v_n}^*$ to identify V^* with $\mathbb{R}^n$.

If $Df\colon V \to V^*$ is one-one with differentiable inverse under one, and hence any, such pair of identifications, we can define the Legendre transform $g = \mathcal{L}(f)\colon V^* \to \mathbb{R}$ of f by the exact analogue of (a),

$$g(\lambda) = \lambda\big((Df)^{-1}(\lambda)\big) - f\big((Df)^{-1}(\lambda)\big).$$

For computations, we simply use all the same formulas as before, except that we work with a basis $\{v_i\}$ of V and the dual basis $\{v_i{}^*\}$ of V^*. The Legendre transformation is again involutive: the maps $f\colon V \to \mathbb{R}$ and $\mathfrak{L}(\mathfrak{L}(f))\colon V^{**} \to \mathbb{R}$ are equal when we identify V^{**} with V by the natural isomorphism.

The fact that we are interested in the Legendre transformation for $f\colon V \to \mathbb{R}$ is, of course, a dead giveaway that we want to apply it to manifolds.

Given a Lagrangian $L\colon TM \to \mathbb{R}$, at each point $p \in M$ the restriction $L_p = L|M_p$, is a smooth function taking the vector space M_p to $\mathbb{R}$. We can therefore consider the derivative

$$D(L_p)\colon M_p \to M_p{}^*.$$

By putting together all the maps $D(L_p)$, we obtain a map from TM to T^*M called the "fibre derivative" of L, though the "fibre-wise derivative" of L might be a better name. For this map we will adopt the notation

$$\mathbf{F}DL\colon TM \to T^*M = \bigcup_{p \in M} D(L_p)\colon M_p \to M_p{}^*.$$

In terms of a coordinate system $(q,\dot q) = (q^1,\dots,q^n,\dot q^1,\dots \dot q^n)$ for TM and the corresponding coordinate system (q,p) for T^*M, the map $\mathbf{F}DL$ is given by $(q,\dot q) \mapsto (q, \partial L/\partial \dot q)$. More precisely, we can write

$$p_i \circ \mathbf{F}DL = \frac{\partial L}{\partial \dot q^i}.$$

We call L a **regular** Lagrangian if each $D(L_p)$ is one-one with a differentiable inverse, in which case can write the above equation for $p_i \circ \mathbf{F}DL$ as

$$p_i = \frac{\partial L}{\partial \dot q^i} \circ (\mathbf{F}DL)^{-1}.$$

Note that L is always regular in the case of mechanics problems, where $L = T - V$ for the positive definite kinetic energy T.

For a regular Lagrangian L we can also apply the Legendre transformation to each L_p, to obtain maps $\mathfrak{L}(L_p)\colon M_p{}^* \to \mathbb{R}$, and putting these together we get the "fibre-wise Legendre transform" of L,

$$H = \mathbf{F}\mathfrak{L}(L)\colon T^*M \to \mathbb{R} = \bigcup_{p \in M} \mathfrak{L}(L_p)\colon M_p{}^* \to \mathbb{R}.$$

All these considerations are easily extended to the case where we have a Lagrangian $L\colon TM \times \mathbb{R} \to \mathbb{R}$, obtaining $H\colon T^*M \times \mathbb{R} \to \mathbb{R}$. Basically, we simply work on each $TM \times \{t\}$ separately, and then put the results together.

ADDENDUM 16A

THE CLAIRAUT EQUATION

The *Clairaut equation* provides an interesting elementary application of the Legendre transformation, and also serves as a simple illustration of the envelope of a family of solutions of an equation, which will play an important role in the optional sections near the beginning of Chapter 18. In dimension 1, a Clairaut equation is one of the form

$$u(x) = xu'(x) - f(u'(x)).$$

Differentiation leads immediately to

$$0 = \big[x - f'(u'(x))\big] \cdot u''(x),$$

and thus "in general" to two different equations:

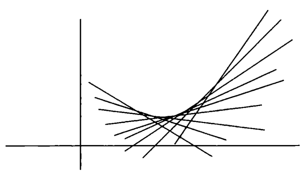
$$(1) \qquad u''(x) = 0$$

$$(2) \qquad x - f'(u'(x)) = 0.$$

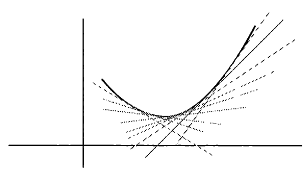
For (1) we must have $u'(x) = c$ for some constant c, and indeed

$$u(x) = cx - f(c)$$

is always a solution. We thus have a 1-parameter family of linear solutions (a),



(a)



(b)

and the envelope (b) of this 1-parameter family of solutions must also be a solution. The standard way of obtaining the envelope (see Addendum 8B) is to "differentiate the equation $u(x) = cx - f(c)$ with respect to c",

$$0 = x - f'(c) \implies c = (f')^{-1}(x),$$

and then substitute back into the equation $u(x) = cx - f(c)$, to get

$$u(x) = x(f')^{-1}(x) - f\big((f')^{-1}(x)\big),$$

so the envelope is $\mathfrak{L}(f)$, which is also just what we get by solving (2) for $u'(x)$ and substituting back into the equation. This whole discussion assumes that we are considering an interval on which which we can define $\mathcal{L}(f)$, which is equivalent to the condition that the corresponding straight lines have an envelope. (We can easily obtain more complicated solutions by piecing together an arc of $\mathcal{L}(f)$ with portions of the straight line solutions at either end, but for simplicity we stick to the general case.)

Another approach to Clairaut's equation is to apply the Legendre transformation directly to u, to obtain information about its transform $v = \mathcal{L}(u)$. To minimize notational complexity, and maximize confusion, we use the classical notation on page 515. We have

$$\begin{aligned} v(\xi) &= \xi x - u(x) \qquad\qquad\qquad \text{with } u_x = \xi \\ &= x u_x - [x u_x - f(u_x)] \\ &= f(\xi), \end{aligned}$$

i.e., the Legendre transform of the solution is f, so by involutivity, the solution is $\mathcal{L}(f)$. In this case, we miss all the straight line solutions, which don't have Legendre transforms.

In a similar way, for the 2-dimensional Clairaut equation

$$u = x u_x + y u_y - f(u_x, u_y),$$

we find, using the notation on page 515, that $v = \mathcal{L}(u)$ satisfies

$$\begin{aligned} v(\xi, \eta) &= \xi x + \eta y - u(x, y) \\ &= x u_x + y u_y - [x u_x + y u_y - f(u_x, u_y)] \\ &= f(\xi, \eta), \end{aligned}$$

so that one solution is $\mathcal{L}(f)$.

The missing solutions are obtained by differentiating the equation, as before, now separately with respect to x and y. Letting $f_1 = D_1 f$ and $f_2 = D_2 f$, we obtain

$$\begin{aligned} (x + f_1) u_{xx} + (y + f_2) u_{xy} &= 0 \\ (x + f_1) u_{xy} + (y + f_2) u_{yy} &= 0. \end{aligned}$$

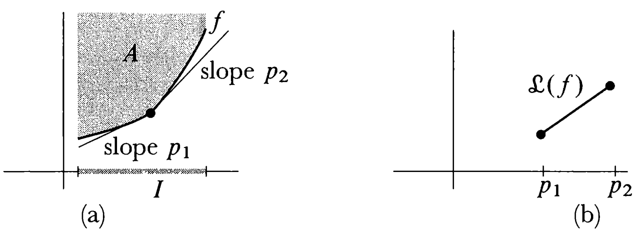
If $u_{xx} u_{yy} - {u_{xy}}^2 \neq 0$, then we have $x = -f_1$ and $y = -f_2$, which gives the solution $\mathcal{L}(f)$ already obtained.

On the other hand, when $u_{xx} u_{yy} - {u_{xy}}^2$ is identically 0, we get the developable surfaces, or in the general case, the tangent planes to the graph of $\mathcal{L}(f)$.

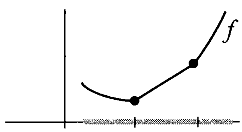
PROBLEMS

1. (a) $\mathcal{L}(cf)(p) = c\mathcal{L}f(p/c)$. (Definition (a′) helps preserve sanity.)
(b) If $f(x) = x^2$, then $\mathcal{L}f(p) = \frac{1}{4}p^2$.
(c) If $f(x) = \frac{1}{2}mx^2$, then $\mathcal{L}f(p) = p^2/2m$.

2. (a) For any continuous f on an interval I, the geometric definition of $\mathcal{L}(f)$ on page 514 makes sense whenever $A = \{(x, y) : x \in I, y \geq f(x)\}$ is convex.



Show that a point where the left- and right-hand derivatives of f are unequal gives rise to a line segment in the graph of $\mathcal{L}(f)$.
(b) Conversely, what happens if the graph of f contains a straight line segment?



We can summarize by saying that if f is a convex polygon, then $\mathcal{L}(f)$ is also, with vertices of f corresponding to edges of $\mathcal{L}(f)$ and edges of f corresponding to vertices of $\mathcal{L}(f)$.

3. (a) If $g = \mathcal{L}(f)$, then we have $mx - f(x) \leq g(m)$, for all x and m, and thus

$$mx \leq f(x) + g(m).$$

(b) If $f(x) = x^a/a$, then $\mathcal{L}(f) = p^b/b$, where $\frac{1}{a} + \frac{1}{b} = 1$. Hence

$$mx \leq \frac{x^a}{a} + \frac{m^b}{b}$$

for all $x, m > 0$ and $\frac{1}{a} + \frac{1}{b} = 1$ (*Young's inequality*).

4. If f is a quadratic function,

$$f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

show that

$$\mathcal{L}(f)\big(D_1 f(x), \ldots, D_n f(x)\big) = f(x).$$

5. For two points $a, b \in U \subset \mathbb{R}^n$, consider

$$\frac{d}{dt} Df(a + t(b - a)),$$

i.e., the vector

$$\left(\frac{d}{dt}\frac{\partial f}{\partial x_1}(a + t(b - a)), \ldots, \frac{d}{dt}\frac{\partial f}{\partial x_n}(a + t(b - a)) \right).$$

Show that

$$\frac{d}{dt} Df(a + t(b - a)) = (b - a) \cdot J(b - a),$$

where J is the Jacobian matrix $\big(\partial^2 f / \partial x_i \partial x_j\big)$.

(a) Conclude that

$$Df(b) - Df(a) = \int_0^1 (b - a) \cdot J(b - a),$$

and using the hypothesis that J is always positive definite, show that if $b - a \neq 0$, then $Df(b) - Df(a) \neq 0$.

CHAPTER 17

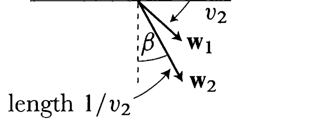
THE INTERPLAY OF MECHANICS AND OPTICS

William Rowan Hamilton (1805–1865) presented his first paper, "Theory of Systems of Rays", to the Royal Irish Academy in 1824, and on the basis of it and its sequels he was appointed to the Professorship of Astronomy at the University of Dublin in 1827. Three Supplements followed the paper, the whole kit and kaboodle taking up pages 1–293 of Hamilton [1]; luckily, we are not studying optics, so we need only consider one or two things near the beginning and the end of from these papers!

Optics emulates mechanics. In the Introduction to his first paper, Hamilton noted that previous investigations of geometrical optics had usually been confined to special cases, with very few general theoretical results. Though Malus' theorem was of a more general nature, Hamilton noted the error in it, and considered it to be too specialized.

In addition, Hamilton wanted to give optics, still buffeted by the battle between wave and corpuscular theories, as secure a foundation as mechanics, by basing it on a principle that was "independent of any hypothesis about the nature or the velocity of light". Maupertuis' principle of least action still held that favored position for mechanics, and Hamilton chose Maupertuis' principle of least action for light as the basis for optics, avoiding questions of the velocity of light by reference to the refractive index, as mentioned on page 497. We now know that Hamilton's choice was in fact Fermat's principle and if we choose units so that the speed of light in a vacuum is 1, the refractive index for any medium is $1/v$ for the speed of light v in that medium, and Fermat's principle states that the path of a light ray is a critical point for $\int (1/v)\, ds$ along the path, that is, for the time taken to traverse the path.

Note that if $\mathbf{w}_1$ and $\mathbf{w}_2$ are vectors along the directions of the incident ray and refracted ray whose lengths are $1/v_1$ and $1/v_2$ for the two media, the horizontal components of these vectors have lengths $\sin\alpha/v_1$ and $\sin\beta/v_2$, and thus the

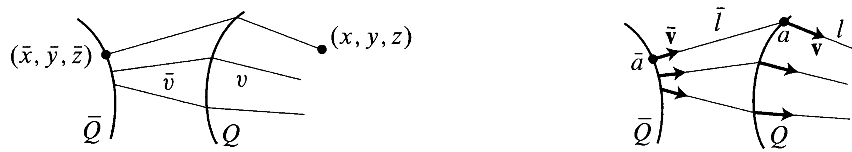


sine law $\sin\alpha/\sin\beta = v_1/v_2$ says these horizontal components are equal, so that

$$\mathbf{w}_2 - \mathbf{w}_1 \text{ is } \textbf{perpendicular to the separating plane},$$

a criterion that can be applied to the tangent plane of a smooth boundary surface at any point.

MALUS' THEOREM. Suppose now that we have, as in (a) of the figure below, a family of light rays emitted from a surface $\bar{Q}$, passing through a medium with refractive index $1/\bar{v}$ until they are refracted by a surface Q which is the boundary of a medium with refractive index $1/v$. The time required for light



(a) (b)

to go from a point $(\bar{x}, \bar{y}, \bar{z})$ on $\bar{Q}$ to a point (x, y, z) is called the *optical length* of the light ray. Now consider (b) a vector field $\bar{\mathbf{v}}$ on $\bar{Q}$ pointing in the direction of the light rays, with all vectors having length $1/\bar{v}$, and a vector field $\mathbf{v}$ on Q pointing in the direction of the refracted rays, with vectors of length $1/v$. If a is the point where the ray from $\bar{a}$ intersects Q, and $\bar{l}$ is the distance from $\bar{a}$ to a, while l is the distance from a to some point b along the refracted ray, then the optical length $\mathcal{O}$ of the ray from $\bar{a}$ to b is $\bar{l}/\bar{v} + l/v$.

For a curve $u \mapsto \bar{a}(u)$ in the surface $\bar{Q}$ with corresponding $\bar{\mathbf{v}}(u)$, as well as corresponding $a(u)$ and $\mathbf{v}(u)$, and $\bar{l}(u)$ and $l(u)$, we have

(1) $$a(u) = \bar{a}(u) + \bar{l}(u)\bar{v} \cdot \bar{\mathbf{v}}(u),$$

(2) $$b(u) = a(u) + l(u)v \cdot \mathbf{v}(u).$$

Since all $\bar{\mathbf{v}}(u)$ have the same length $1/\bar{v}$, differentiating (1) and taking the inner product with $\bar{\mathbf{v}}(u)$ gives

$$\langle a'(u), \bar{\mathbf{v}}(u)\rangle = \langle \bar{a}'(u), \bar{\mathbf{v}}(u)\rangle + \bar{l}'(u)/\bar{v}.$$

But $a'(u)$ is tangent to Q, so, using the criterion for the direction of the refracted ray given at the top of this page, we have

$$\langle \bar{\mathbf{v}}(u) - \mathbf{v}(u), a'(u)\rangle = 0,$$

and we obtain

(3) $$\langle a'(u), \mathbf{v}(u)\rangle = \langle \bar{a}'(u), \bar{\mathbf{v}}(u)\rangle + \bar{l}'(u)/\bar{v}.$$

Differentiating (2), taking the inner product with $\mathbf{v}(u)$, and using (3) we obtain

$$\begin{aligned}\langle b'(u), \mathbf{v}(u)\rangle &= \langle a'(u), \mathbf{v}(u)\rangle + l'(u)/v \\ &= \langle \bar{a}'(u), \bar{\mathbf{v}}(u)\rangle + \bar{l}'(u)/\bar{v} + l'(u)/v \\ &= \langle \bar{a}'(u), \bar{\mathbf{v}}(u)\rangle + \mathcal{O}'(u).\end{aligned}$$

Suppose, in particular, that all the rays are emitted perpendicularly from $\bar{Q}$, so that all $\langle \bar{a}'(u), \bar{\mathbf{v}}(u)\rangle = 0$. Then all $\langle b'(u), \mathbf{v}(u)\rangle = 0$ if and only if $\mathcal{O}$ is a constant. So, given a system of rays all perpendicular to $\bar{Q}$, we can find a surface perpendicular to the refracted rays simply by taking the endpoints of all rays with any constant optical length. In this way (or at any rate, by equivalent considerations), Hamilton was able to prove the most general version of Malus' Theorem for refraction, without any involved calculations; reflections can also be included by choosing $v = -v_0$.

All these considerations can easily be extended to the case of an "optical instrument" where several successive refractions are involved. For a region of the final medium where each point (x, y, z) is on a unique ray starting from a point $(\bar{x}, \bar{y}, \bar{z})$ on $\bar{Q}$, Hamilton called the optical length of that ray from $(\bar{x}, \bar{y}, \bar{z})$ to (x, y, z) the *characteristic function* $V(x, y, z)$ for this instrument, with

$$\text{(C)} \qquad \left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2 = \frac{1}{v^2},$$

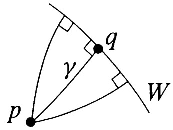
where v is the speed of light in the final medium, and he made V, and related functions, the theoretical basis for optics.

FERMAT'S PRINCIPLE AND HUYGENS' CONSTRUCTION. Hamilton had mainly been concerned with geometrical optics, but by the time he presented the Third Supplement, in 1832, the wave theory of Young and Fresnel was pretty well established, and Hamilton addressed it near the end.

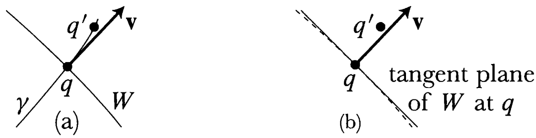
Given indicatrices, as on page 495, determining the waves obtained by Huygens' construction, we can introduce Fermat's principle by identifying our light rays with the geodesics for this method of measuring lengths—restricting our attention to geodesics γ for which every tangent vector $\gamma'(t)$ has "length 1", i.e., lies on the indicatrix at $\gamma(t)$. For light emanating from a single point p, we can then define the wave front $\Phi_p(t)$ for $t > 0$ to consist of points $\gamma(t)$ where γ defined on $[0, t]$ is a minimal geodesic with $\gamma(0) = p$ (more generally, we could consider light rays emanating from a surface, or any closed set).

When our indicatrices are spheres of varying size so that we are dealing with a Riemannian metric $\langle\ ,\ \rangle$ for which orthogonality is the same as the

Euclidean one, each such geodesic γ from p to a point q in the wave front W will intersect W orthogonally. This can be proved by a calculation, basically

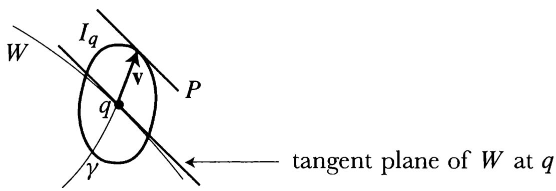


the one used to prove "Gauss' Lemma" (cf. DG, pg. 337), but a geometric limiting argument for this intuitively obvious fact can be given by considering the extension of γ to a point q' beyond q by a very small amount (a). Then (b)



the line from q to q' will be practically the direction of the tangent vector $\mathbf{v}$ of γ at q, and it should also be almost orthogonal to the tangent plane of W at q, since the shortest distance from the point q' to that plane is the perpendicular.

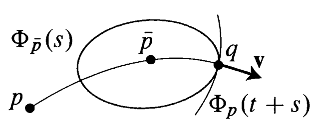
For general indicatrices, of which ellipsoids are a special case, a similar argument shows that if I_q is the indicatrix at q, then the tangent vector $\mathbf{v}$ of γ at q



is the unique vector $\mathbf{v}$ in I_q *conjugate* to the tangent plane of W at q, meaning that the tangent plane P of I_q at $\mathbf{v}$ is parallel to the tangent plane of W at q, so that $\mathbf{v}$ is the point of the indicatrix furthest from thc tangent plane of W at q.

Knowing a specific relationship, it is easy to show that for $s > 0$ we have Huygens' construction,

$$\Phi_p(t+s) \subset \text{envelope of } \{\Phi_{\bar{p}}(s) : \bar{p} \in \Phi_p(t)\}.$$

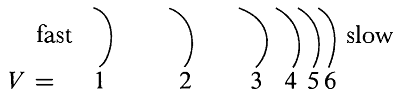


In fact, for $q \in \Phi_p(t+s)$, let γ be a geodesic defined on $[0, t+s]$ with $\gamma(0) = p$ and $\gamma(t+s) = q$, and let $\mathbf{v}$ be the tangent vector at q, so that $\mathbf{v}$ is conjugate to the tangent plane of $\Phi_p(t+s)$ at q. Then $\bar{p} = \gamma(t) \in \Phi_p(t)$ and since $\bar{\gamma} = \gamma|[t, t+s]$ is a geodesic of length s from $\bar{p}$ to q, we have $q \in \Phi_{\bar{p}}(s)$. But $\mathbf{v}$ is also the tangent vector of $\bar{\gamma}$ at q, so it must also be conjugatc to the tangent plane of $\Phi_{\bar{p}}(s)$ at q. It follows that $\Phi_{\bar{p}}(s)$ and $\Phi_p(t+s)$ are tangent at q.

In this way, Hamilton linked Fermat's principle to Huygens' construction for the spread of wavefronts, and conversely, given Huygens' construction, the integral curves of the conjugate directions give the rays for Fermat's principle, formalizing the notion, by then generally recognized, that the two principles were essentially equivalent.

Note, by the way, that the wave fronts are just the level sets of V; since these level sets are far apart when the wave moves fast and close together when they move slowly, Hamilton called the gradient

$$\mathbf{p} = \operatorname{grad} V = \left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial z}\right),$$

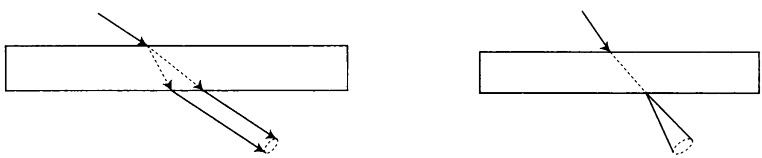


which can also be defined by

$$dV = \sum_{i=1}^{3} p_i \, dx^i,$$

the *vector of normal slowness* of the wave front. The vector $\mathbf{p}$ is perpendicular to the wave front in the usual inner product on $\mathbb{R}^3$, so if the indicatrices are all spherical, and thus perpendicularity isn't changed, we have, using (C) on page 524 that the tangent $\mathbf{t}$ of a ray is given by the equation $\mathbf{t} = \mathbf{p}/|\mathbf{p}| = v\mathbf{p}$.

CONICAL REFRACTION IN ARAGONITE. The third Supplement ended with a demonstration that Hamilton had also mastered the intricacies of Fresnel's theory. He proved that for "bi-axial" crystals, with properties like Iceland spar, when a ray hits the crystal at a certain angle (a) the ray will not split into two rays, but theoretically split instead into a narrow cone of rays, then emerging as



(a) internal conical refraction (b) external conical refraction

a narrow hollow cylinder of rays (internal conical refraction), while at another angle (b) a ray of light will not be split at all until it emerges from the crystal in a narrow cone (external conical refraction).

Hamilton asked the physicist Humphrey Lloyd of the University of Dublin to test this prediction. After obtaining an extremely pure specimen of the crystal aragonite, Lloyd was able to check these results, with data agreeing extremely well with Hamilton's predictions. Fresnel's theory had by then received so much attention that no one expected any startling new theoretical deductions to be made, and the discovery of conical refraction suddenly made Hamilton quite famous, and something of a scientific hero in Ireland.

But Hamilton's lasting fame today rests on a paper published shortly after the third Supplement had been written.

Mechanics returns the compliment. Hamilton's main papers on Optics appeared in the *Transactions of the Royal Irish Academy*, but in 1834 two papers by Hamilton on the subject of Dynamics were published in the *Philosophical Transactions of the Royal Society* of England (perhaps his conical refraction fame had a role in this). These papers eventually turned out to be as important as Hamilton obviously expected. In fact, in ways that were unexpected, and in ways that no one at the time could have anticipated, they turned out to be more important than he could ever have imagined.

As we saw in previous sections, in Hamilton's treatment of optics, the rays were determined by using a variational principle, Maupertuis' principle of least action for light, which we discreetly replaced by Fermat's principle, involving $\int (1/v)\,ds$. Then the "optical length" of a ray, the value of $\int (1/v)\,ds$ along the ray, was used to define a *characteristic function.*

Hamilton aimed to transfer this treatment of optics back to mechanics, and in his first paper he naturally chose Maupertuis' principle of least action to determine the trajectories of a set of particles, and used the value of the action integral $\mathcal{A}$ along this trajectory to define a **characteristic function** for mechanics.

Not surprisingly, in view of the vagaries of Maupertuis' form of the principle of least action, the formulas Hamilton derived had some unpleasant complexities, and right at the end of the paper Hamilton noted that it would be simpler if, instead of the integral of the action, one used the integral $\int L = \int T - V$ to get an "auxiliary function S".

By the time Hamilton wrote up his second paper, S had been promoted to the head of the class, and was now called the **principal function**. So in this paper Hamilton started all over. In particular, the trajectories c of a system of particles were now determined by the principle that

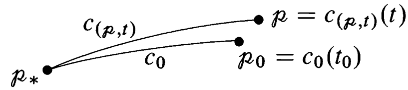
$$c\colon [t_1, t_2] \to M \quad \text{should be a critical value for} \quad \int_{t_1}^{t_2} L(c'(t), t)\,dt.$$

And, as in one of Kipling's *Just So Stories*, that is how "a simple and general result of the laws of mechanics" noted by Lagrange (page 464) ended up being known as *Hamilton's principle.*

For the characteristic function in optics, we had to consider a region where each point is on a unique ray starting from some point on our initial surface $\bar{Q}$, or in the simplest case a unique ray starting from some fixed point; this is basically the same as saying that we need each point in the region to be on a ray with a unique initial direction. The principal function S for mechanics requires more involved considerations, because mechanics trajectories can have different speeds, so we need to specify more than just an initial direction. As we will see, Hamilton was not one to shy away from such involved considerations.

Suppose $\wp_*$ is some fixed point in M, and we have a particular trajectory c_0 from p_*, defined on an interval $[0, t_0]$, with endpoint $\wp_0 \in M$. Rather than simply considering a neighborhood of $\wp_0$ in M, we will consider a neighborhood of $(\wp_0, t_0)$ in $M \times \mathbb{R}$: if $(\wp, t)$ is sufficiently close to $(\wp_0, t_0)$, there will be a unique trajectory $c_{(\wp,t)} : [0, t] \to M$, defined on the interval $[0, t]$, with $c_{(\wp,t)}(0) = \wp_*$ and $c_{(\wp,t)}(t) = \wp$. We then define $S\colon M \times \mathbb{R} \to \mathbb{R}$ by

$$S(\wp, t) = \int_0^t L(c_{(\wp,t)}{}'(\tau), \tau)\, d\tau.$$



The equations on T^*M. Hamilton's first step in studying the principal function S was to transfer Lagrange's equations on TM to equations on T^*M: a curve $c\colon \mathbb{R} \to M$ gives rise to the curve $c'\colon \mathbb{R} \to TM$, and thus to a curve $\gamma = \mathbf{F}DL \circ c' : \mathbb{R} \to T^*M$,

$$\gamma(t) = DL_{c(t)}(c'(t)),$$

and we want to translate Lagrange's equations for c into equations for γ, which turn out to assume a very simple form in terms of the fibre-wise Legendre transform H of L (Hamilton never actually mentions the Legendre transformation, merely introducing its formulas with hardly any explanation). We follow the classical computations, with a bit more explicitness.

A coordinate system q on M gives coordinates $(q, \dot{q})$ on TM and (q, p) on T^*M. So that we don't get tangled up in cumbersome notation, we will let the inverse of $\mathbf{F}DL\colon TM \to T^*M$ be denoted simply by $\phi : T^*M \to TM$. Then the next to last equation on page 517 can be written

$$(1) \qquad p_i = \frac{\partial L}{\partial \dot{q}^i} \circ \phi.$$

For H, we want to apply equation (A) on page 515 for the $DL_\wp$ on the various $M_\wp$, so the coordinates that we apply to the $(Df)^{-1}$ term are now the $\dot{q}^i$, and thus we define

$$(2) \qquad H = \sum_{i=1}^n p_i \cdot (\dot{q}^i \circ \phi) - L \circ \phi.$$

(It's a good idea to consult Problem 1 for examples of what H looks like). Then

$$\begin{aligned} dH = {}& \sum_{i=1}^n p_i\, d(\dot{q}^i \circ \phi) + \sum_{i=1}^n (\dot{q}^i \circ \phi)\, dp_i \\ & - \sum_{i=1}^n \left(\frac{\partial L}{\partial q^i} \circ \phi\right) d(q^i \circ \phi) - \sum_{i=1}^n \left(\frac{\partial L}{\partial \dot{q}^i} \circ \phi\right) d(\dot{q}^i \circ \phi). \end{aligned}$$

Equation (1) shows that the first and fourth terms cancel, so that we have

$$dH = \sum_{i=1}^n (\dot{q}^i \circ \phi)\, dp_i - \sum_{i=1}^n \left(\frac{\partial L}{\partial q^i} \circ \phi\right) d(q^i \circ \phi).$$

But we always have $dH = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q^i} d(q^i \circ \phi)$, so we get

$$\dot{q}^i \circ \phi = \frac{\partial H}{\partial p_i}, \qquad \frac{\partial L}{\partial q^i} \circ \phi = -\frac{\partial H}{\partial q^i}.$$

[For more classical looking computations, simply delete '$\circ\,\phi$' wherever it occurs.]

Now suppose that c satisfies Lagrange's equations,

$$\frac{\partial L}{\partial q^i}(c(t), c'(t)) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}(c(t), c'(t))\right).$$

Taking into account equation (1) again, we find that γ satisfies

$$\text{(H)} \qquad \begin{aligned} \dot{q}^i(\gamma'(t)) &= (q^i \circ \gamma)'(t) = \frac{\partial H}{\partial p_i}(\gamma(t)) \\ \dot{p}^i(\gamma'(t)) &= (p_i \circ \gamma)'(t) = -\frac{\partial H}{\partial q^i}(\gamma(t)) \end{aligned} \qquad \begin{array}{c}\text{or in} \\ \text{"condensed} \\ \text{notation"}\end{array} \qquad \left\{ \begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i}, \end{aligned} \right.$$

a system of $2n$ first order equations, in contrast to Lagrange's system of n second order differential equations.

For a Lagrangian $L\colon M \times \mathbb{R} \to \mathbb{R}$ depending on t, we just apply this analysis to each $M \times \{t\}$ to get equations on $T^*M \times \mathbb{R}$. It is not hard to see that in this case we also have

$$\text{(H}'\text{)} \qquad \frac{\partial H}{\partial t}(\gamma(t), t) = -\frac{\partial L}{\partial t}(c'(t), t) \qquad \text{condensed to} \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.$$

Equation (H) implies that, in slightly abbreviated notation,

$$\frac{d}{dt} H(\gamma(t), t) = \sum_{i=1}^{n} \left[\frac{\partial H}{\partial q^i}\dot{q}^i(t) + \frac{\partial H}{\partial p_i}\dot{p}_i\right] + \frac{\partial H}{\partial t}(\gamma(t), t) = \frac{\partial H}{\partial t}(\gamma(t), t),$$

so when L (and H) do not depend on t, we have the generalized energy integral

$$\text{(E)} \qquad H(\gamma(t)) \text{ is a constant.}$$

More generally, equations (1) and (2) show that

$$H(\gamma(t), t) = \left[\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L\right](c'(t), t).$$

The term in brackets is just the action minus L on page 448, which is a constant if L and H do not depend on t; usually, of course, it is the energy $E = T + V$.

The partial derivatives of S. For Hamilton, the equations (H) were mainly a tool for finding the partial derivatives of the principal function S. For an initial point $\mathscr{p}_* \in M$, the principal function S may be written in coordinates as

$$S(q,t) = S(q^1, \dots, q^n, t) = \int_0^t L\big(c_{(q,t)}{}'(\tau)\big)\, d\tau,$$

where $c_{(q,t)}\colon [0,t] \to M$ is the solution of Lagrange's equations satisfying $c_{(q,t)}(0) = \mathscr{p}_*$ and $c_{(q,t)}(t) =$ the point with coordinates q. For notational convenience, we use the abbreviation γ_q for the curve

$$\gamma_q = \gamma_{(q,t)} = \mathbf{F}DL \circ c_{(q,t)} \quad \text{in } T^*M.$$

In Hamilton's calculations of the partial derivatives of S, the Lagrangian L was replaced, without any preamble, by the Legendre transform of H, thus relying on the fact that the Legendre transformation is involutive, so that we have (after a bit of thought)

$$\text{(3)} \qquad S(q,t) = \int_0^t \sum_{j=1}^n (p_j \circ \gamma_q)(\tau)(q^j \circ \gamma_q)'(\tau) - H(\gamma_q(\tau))\, d\tau.$$

To find $\partial S/\partial q^i$, we take the derivative $\partial/\partial q^i$ inside the integral sign to obtain

$$\begin{aligned}
&\frac{\partial S}{\partial q^i}(q,t)\\
&= \int_0^t \Bigg[\sum_{j=1}^n \frac{\partial}{\partial q^i}(p_j \circ \gamma_q)(\tau)(q^j \circ \gamma_q)'(\tau) + \sum_{j=1}^n (p_j \circ \gamma_q)(\tau)\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)'(\tau)\\
&\quad - \sum_{j=1}^n \frac{\partial H}{\partial q^j}(\gamma_q(\tau))\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)(\tau) - \sum_{j=1}^n \frac{\partial H}{\partial p_j}(\gamma_q(\tau))\frac{\partial}{\partial q^i}(p_j \circ \gamma_q)(\tau)\Bigg]\, d\tau.
\end{aligned}$$

Applying equations (H), we see that the first and fourth sums cancel, and we can also write the $\partial H/\partial q^j$ in the third sum in terms of p_j, to get

$$\begin{aligned}
\frac{\partial S}{\partial q^i}(q,t) &= \int_0^t \Bigg[\sum_{j=1}^n (p_j \circ \gamma_q)(\tau)\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)'(\tau)\\
&\qquad + \sum_{j=1}^n (p_j \circ \gamma_q)'(\tau)\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)(\tau)\Bigg]\, d\tau,\\
&= \int_0^t \sum_{j=1}^n \Big((p_j \circ \gamma_q)(\tau)\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)(\tau)\Big)'\, d\tau\\
&= \sum_{j=1}^n (p_j \circ \gamma_q)(\tau)\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)(\tau)\Bigg|_{\tau=0}^{\tau=t}.
\end{aligned}$$

Since $c_{(q,t)}(t)$ has coordinates q, the same is true of $\gamma_q(t)$, so that

$$\frac{\partial}{\partial q^i}(q^j \circ \gamma_q)(t) = \delta_i^j,$$

and the upper limit term, for $\tau = t$, is just $p_i(\gamma_q(t))$. Moreover, all $\gamma_q(0) = \mathscr{p}_*$, so that the lower limit term, for $\tau = 0$, vanishes. Thus, we have found that

$$\text{(4)} \qquad \frac{\partial S}{\partial q^i}(q,t) = p_i(\gamma_q(t)).$$

For the computationally challenged, a startlingly short alternative proof is given in Problem 2.

These results mean that if we have a formula for S, then we know the $p_i(\gamma_q(t))$, which in the case of mechanics problems with $L = T - V$ means that we know the $\dot{q}^i(c(t))$, and thus we can solve for c in terms of integrals. "The difficulty of mathematical dynamics is therefore reduced to the search and study of this one function S, which may for that reason be called the PRINCIPAL FUNCTION of motion of a system."

A partial differential equation for S. For the study of S, we will first find a formula for $\partial S/\partial t$. Note that equation (3) immediately gives, for fixed q,

$$\frac{d}{dt}S(q,t) = \sum_{j=1}^{n}(p_j \circ \gamma_q)(t)(q^j \circ \gamma_q)'(t) - H(\gamma_q(t)).$$

But also

$$\frac{d}{dt}S(q,t) = \frac{\partial S}{\partial t}(q,t) + \sum_{j=1}^{n}\frac{\partial S}{\partial q^j}(q,t)(q^j \circ \gamma_q)'(t).$$

Therefore

$$\text{(5)} \qquad \begin{aligned} \frac{\partial S}{\partial t}(q,t) &= \sum_{j=1}^{n}(p_j \circ \gamma_q)(t)(q^j \circ \gamma_q)'(t) - H(\gamma_q(t)) \\ &\qquad - \sum_{j=1}^{n}\frac{\partial S}{\partial q^j}(q,t)(q^j \circ \gamma_q)'(t) \\ &= -H(\gamma_q(t)) \qquad \text{using (4).} \end{aligned}$$

Together, (4) and (5) give us the equations

$$\text{(6)} \qquad \begin{aligned} \frac{\partial S}{\partial q^i}(q,t) &= p_i(\gamma_q(t),t) \\ \frac{\partial S}{\partial t}(q,t) &= -H(\gamma_q(t),t) \end{aligned} \qquad \begin{array}{c}\text{or in}\\ \text{condensed}\\ \text{notation}\end{array} \qquad \left\{ \begin{aligned} \frac{\partial S}{\partial q^i}(q,t) &= p_i \\ \frac{\partial S}{\partial t}(q,t) &= -H(q,p,t). \end{aligned} \right.$$

These equations can be combined to give a single first order partial differential equation for the function S on $M \times \mathbb{R}$,

$$\frac{\partial S}{\partial t} + H\left(q^1, \dots, q^n, \frac{\partial S}{\partial q^1}, \dots, \frac{\partial S}{\partial q^n}, t\right) = 0.$$

In Hamilton's paper, this equation is arrived at in short order, and the remaining forty or so pages are devoted to explaining how to get increasingly better approximations to the solution S of the equation, and thus to the equations of the system for which S was constructed. A prototypical use of such a process would be to obtain approximate solutions to the three body problem, starting with the solution of the two body problem, especially if the third body were small, like a moon.

On the other hand, while Hamilton treated the equations (H) simply as a means to obtain the equation for S, these equations became a central object of study in the work of Jacobi, to be discussed in the next chapter, imbuing Hamilton's results with an entirely different flavor, which forms the basis for what we now call "Hamiltonian mechanics". In fact, nowadays equations (H) are usually presented as the basis of Hamiltonian mechanics, with the principal function S usually entering the picture only as a sort of afterthought. So, although Hamilton's viewpoint was the inspiration for a whole new way of looking at mechanics, in the following chapters we will also want to focus attention on the equations (H) themselves.

For the present however, we simply want to access our current situation, and prepare for the future.

Invariant definitions; the interplay of TM and T^*M. So far we have only given somewhat indirect proofs of the invariance of Lagrange's equation. Given a Lagrangian $L\colon TM \to \mathbb{R}$, we have not defined a vector field $\mathbf{X}_L$ on TM such that the (first order) equations for the integral curves of $\mathbf{X}_L$ will be equivalent to Lagrange's (second order) equations. This calls for a few words about second order ODE's on M.

For a coordinate system q on M, and corresponding $(q, \dot{q})$ on TM, consider the second order system of ODE's

$$\text{(a)} \qquad \frac{d^2 q^i}{dt^2} = F_i(q, \dot{q}),$$

with the arguments omitted, as usual, and the case of explicit dependence on t omitted for simplicity. We can write this as a system of $2n$ first order ODE's,

$$\text{(b)} \qquad \frac{dq^i}{dt} = v^i, \qquad \frac{dv^i}{dt} = F_i(q, v),$$

which can be thought of as a first order ODE for a vector field X on TM,

$$\text{(c)} \qquad X(v) = \sum_{i=1}^{n} \left.\frac{\partial}{\partial q^i}\right|_v + \sum_{i=1}^{n} F_i(q,v) \left.\frac{\partial}{\partial \dot{q}^i}\right|_v \qquad \text{for all } v \in M_q.$$

We clearly have

$$\text{(II)} \qquad \pi_*(X(q,v)) = v \qquad \text{for all } v \in TM$$

and we call any X on TM satisfying (II) a second order ODE on M, since it is then of the form (c), equivalent to (b), and thus to (a).

The question then becomes: can we define, directly in terms of L, a vector field $\mathbf{X}_L$ on TM, which is a second order ODE on M giving Lagrange's equations for L? In general, we cannot, because Lagrange's equations are not in the standard form (a), with second derivatives given explicitly. It is only in the case of a *regular* Lagrangian that they can be written in the form (a), so we only expect to find $\mathbf{X}_L$ for a regular Lagrangian L.

Taking this as a hint, we shift our attention to the equations on T^*M into which Lagrange's equations are taken by $\mathbf{F}DL$,

$$\text{(H)} \qquad \dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}.$$

These equations can simply be regarded as a set of first order differential equations for a vector field on T^*M,

$$\mathbf{X}_H = \sum_{i=1}^{n} \left[\left(\frac{\partial H}{\partial p_i}\right) \frac{\partial}{\partial q^i} - \left(\frac{\partial H}{\partial q^i}\right) \frac{\partial}{\partial p_i} \right].$$

Everything has still been written in terms of a coordinate system, but now we can bring in the invariantly defined 2-form $\boldsymbol{\omega} = \sum_i dp_i \wedge dq^i$ on T^*M. We easily check that

$$\boldsymbol{\omega}\left(\mathbf{X}_H, \sum_{j=1}^{n} A_j \frac{\partial}{\partial q^j} + \sum_{j=1}^{n} B_j \frac{\partial}{\partial p_j}\right) = \sum_{i=1}^{n} \left(-B_i \frac{\partial H}{\partial p_i} - A_i \frac{\partial H}{\partial q^i}\right)$$
$$= -dH\left(\sum_{j=1}^{n} A_j \frac{\partial}{\partial q^j} + \sum_{j=1}^{n} B_j \frac{\partial}{\partial p_j}\right),$$

or, using the notation introduced on page 512,

$$\boxed{\mathbf{X}_H \lrcorner\, \boldsymbol{\omega} = -dH}$$

Since $X \mapsto X \lrcorner\, \boldsymbol{\omega}$ is one-one, this determines $\mathbf{X}_H$ uniquely, giving an invariant definition of the vector field $\mathbf{X}_H$. This makes sense for any $H: T^*M \to \mathbb{R}$, and henceforth a "Hamiltonian H on T^*M" just means any smooth function $H: T^*M \to \mathbb{R}$ (with the suggestion that we will be interested in $\mathbf{X}_H$).

Aside from the elegance of an invariant definition on T^*M, we emphasize that on TM we are constrained to consider second order ODE's, which means we can only allow curves of the form $(q(t), \dot{q}(t))$; more general curves of the form $(Q^1(t), \ldots, Q^{2n}(t))$ for an arbitrary coordinate system $(Q^1, \ldots, Q^{2n})$ on the $2n$-dimensional manifold TM never enter the picture. On the other hand, on T^*M, where there is no particular *a priori* relationship between the p and q coordinates, the equations of a trajectory may involve more general sets of coordinates $(Q^1, \ldots, Q^n, P_1, \ldots, P_n)$, with significant consequences, as we shall see in Chapter 19.

The extended Hamilton's principle. A result that will play an important role in Chapter 19 gives an interpretation of the equations (H) as a condition for extremals. Since equations (H) are basically the Legendre transform of Lagrange's equations, which are equivalent to the Euler equations for the Lagrangian, it is not unreasonable to expect that they can be derived directly from the Euler equations for the Hamiltonian $H = \mathbf{F}\mathfrak{L}(L)$. However, there is a bit of a surprise in store.

For two fixed points $\not{p}_0, \not{p}_1 \in M$, consider a curve γ in T^*M, defined on the interval $[t_0, t_1]$, for which $\pi(\gamma(t_0)) = \not{p}_0$ and $\pi(\gamma(t_1)) = \not{p}_1$, where π is the projection $\pi\colon T^*M \to M$. In other words, the "base curve" $c = \pi \circ \gamma$ of γ goes from $\not{p}_0$ to $\not{p}_1$. Consider the integral

$$\begin{aligned} J &= \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} p_j\, dq^j - H\right](\gamma'(t))\, dt \\ &= \int_{t_0}^{t_1} \sum_{j=1}^{n} p_j(\gamma(t))\dot{q}^j(\gamma'(t)) - H(\gamma'(t), t)\, dt. \end{aligned}$$

Since L is the Legendre transform of H, this is essentially the same as

$$\int_{t_0}^{t_1} L(c'(t), t)\, dt,$$

so we might expect that the critical values of J satisfy the equations (H). Note that this is not directly equivalent to Hamilton's principle, since we are now considering arbitrary curves γ in T^*M, not only those of the form $c'(t)$; in other words, as it is usually expressed, we are allowing p and q to vary independently.

PROPOSITION (EXTENDED HAMILTON'S PRINCIPLE). The critical values of J are precisely the curves $\gamma\colon [t_0, t_1]$ in T^*M whose base curves $c = \pi \circ \gamma$ go from $\not{p}_0$ to $\not{p}_1$ and which satisfy the equations (H).

PROOF. Consider a variation $\alpha(u,t)$ of c, and let

$$J(u) = \int_{t_0}^{t_1} \left[\sum_{j=1}^{n} p_j\, dq^j - H\right] \left(\frac{\partial \alpha}{\partial t}(u,t)\right) dt.$$

We leave it to the reader to check that the following computations make sense when the proper arguments of functions are inserted. We start with

$$\left.\frac{dJ(u)}{du}\right|_{u=0} = \int_{t_0}^{t_1} \sum_{i=1}^{n} \frac{\partial p_i}{\partial u}\dot{q}^i + p_i \frac{\partial \dot{q}^i}{\partial u} - \frac{\partial H}{\partial q^i}\frac{\partial q^i}{\partial u} - \frac{\partial H}{\partial p_i}\frac{\partial p_i}{\partial u}\, dt.$$

Now $\partial \dot{q}^i/\partial u = d/dt(\partial q^i/\partial u)$, and using integration by parts we have

$$\begin{aligned}
\left.\frac{dJ(u)}{du}\right|_{u=0} &= \int_{t_0}^{t_1} \sum_{i=1}^{n} \frac{\partial p_i}{\partial u}\dot{q}^i - \dot{p}_i \frac{\partial q^i}{\partial u} - \frac{\partial H}{\partial q^i}\frac{\partial q^i}{\partial u} - \frac{\partial H}{\partial p_i}\frac{\partial p_i}{\partial u}\, dt \\
&= \int_{t_0}^{t_1} \sum_{i=1}^{n} \left[\frac{\partial p_i}{\partial u}\left(\dot{q}^i - \frac{\partial H}{\partial p_i}\right) - \frac{\partial q^i}{\partial u}\left(\dot{p}_i + \frac{\partial H}{\partial q^i}\right)\right] dt.
\end{aligned}$$

It follows that the quantities in parentheses must be 0, since we can vary the p_i and q^i independently. ❖

The paradoxical fact that the critical paths for all variations, which we need to look at for the extended Hamilton's principle, can be found by examining only the critical paths satisfying $p_i = \partial L/\partial \dot{q}^i$, the ones we need to look at for Hamilton's principle, is explained by a fact that arises from the geometric definition of the Legendre function: applying equation (A′) on page 515, and using the fact that the Legendre transformation is involutive, we see that for given $(\dot{q}^1, \ldots, \dot{q}^n)$, the point

$$(p_1, \ldots, p_n) = \left(\frac{\partial L}{\partial \dot{q}^1}, \ldots, \frac{\partial L}{\partial \dot{q}^n}\right)$$

is *already* an extremal for

$$\sum_{i=1}^{n} p_i \dot{q}^i - H.$$

ADDENDUM 17A
LIOUVILLE'S VOLUME THEOREM

There are two important theorems by Liouville concerning Hamiltonian mechanics. Here we present the simpler of the two, Liouville's volume theorem. A slick proof will appear as a fleeting by-product of the material in Chapter 19 (page 576), but Liouville actually proved a more general result than the one obtained there, and it might be interesting to examine arguments similar to the original ones (a predilection of the author that will intrude itself, once again, in Chapter 21). We begin with a lemma due to Liouville, and an old-fashioned proof.

LEMMA (LIOUVILLE). For an n-parameter family of maps $F_t\colon \mathbb{R}^n \to \mathbb{R}^n$, given by

$$(*)\qquad F_t(a_1,\dots,a_n) = \big(x_1(a_1,\dots,a_n,t),\dots,x_n(a_1,\dots,a_n,t)\big),$$

consider the n-parameter family of vector fields X given by

$$X_i(t)(a_1,\dots,a_n) = x_i{}'(a_1,\dots,a_n,t), \quad \text{i.e.,} \quad = \frac{\partial x_i}{\partial t}(a_1,\dots,a_n,t).$$

Set

$$J(t) = \det M(t) = \det \frac{\partial(x_1,\dots,x_n)}{\partial(a_1,\dots,a_n)}(t),$$

and let Δ be

$$\Delta = \frac{\partial X_1}{\partial x_1} + \cdots + \frac{\partial X_n}{\partial x_n}.$$

Then

$$\frac{\partial J}{\partial t} = J\Delta.$$

PROOF 1. The partial derivative $\partial J/\partial t$ is the sum of the determinants of n matrices, of which the first, for example, is

$$\frac{\partial(X_1,x_2,\dots,x_n)}{\partial(a_1,\dots,a_n)}.$$

For the case $n = 2$, we have the matrix

$$\begin{pmatrix} \dfrac{\partial X_1}{\partial a_1} & \dfrac{\partial X_1}{\partial a_2} \\[2ex] \dfrac{\partial x_2}{\partial a_1} & \dfrac{\partial x_2}{\partial a_2} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial X_1}{\partial x_1}\dfrac{\partial x_1}{\partial a_1} + \dfrac{\partial X_1}{\partial x_2}\dfrac{\partial x_2}{\partial a_1} & \dfrac{\partial X_1}{\partial x_1}\dfrac{\partial x_1}{\partial x_2} + \dfrac{\partial X_1}{\partial x_2}\dfrac{\partial x_2}{\partial a_2} \\[2ex] \dfrac{\partial x_2}{\partial a_1} & \dfrac{\partial x_2}{\partial a_2} \end{pmatrix}.$$

The determinant is not changed if we subtract $\partial X_1/\partial x_2$ times the second row from the first row, which gives us the matrix

$$\begin{pmatrix} \dfrac{\partial X_1}{\partial x_1}\dfrac{\partial x_1}{\partial a_1} & \dfrac{\partial X_1}{\partial x_1}\dfrac{\partial x_1}{\partial a_2} \\[2ex] \dfrac{\partial x_2}{\partial a_1} & \dfrac{\partial x_2}{\partial a_2} \end{pmatrix}$$

whose determinant is $\partial X_1/\partial x_1$ times J, and similarly the determinant of the second matrix in the expansion of $\partial J\,\partial t$ is $\partial X_2/\partial x_2$ times J. The whole argument clearly extends to general n. ❖

PROOF 2. We only need to prove this for t close to any given t_0, so we can assume that M is close to the identity, and thus that $\log M$ is well-defined. Then we have

$$J = \det M = e^{\operatorname{trace}(\log M)}$$

(this is clearly true for diagonalizable complex M, which are dense, hence it is true for all M).

Therefore,

$$\begin{aligned} J' &= J \operatorname{trace}(M^{-1}M') \\ &= J\sum_{i,j=1}^{n} M^{-1}{}_{ij}M_{ji}{}' = J\sum_{i,j=1}^{n}\frac{\partial a^i}{\partial x^j}\frac{\partial x_j{}'}{\partial a_i} \\ &= J\sum_{i,j}^{n}\sum_{k=1}^{n}\frac{\partial a^i}{\partial x^j}\frac{\partial x_j{}'}{\partial x_k}\frac{\partial x_k}{\partial a_i} = J\sum_{j,k=1}^{n}\sum_{i=1}^{n}\frac{\partial a^i}{\partial x^j}\frac{\partial x_k}{\partial a_i}\frac{\partial x_j{}'}{\partial x_k} \\ &= J\sum_{j,k=1}^{n}\delta_{jk}\frac{\partial x_j{}'}{\partial x_k} = J\sum_{j=1}^{n}\frac{\partial x_j{}'}{\partial x_j}. \text{ ❖} \end{aligned}$$

COROLLARY. For a bounded open set $U \subset \mathbb{R}^n$, let $U_t = F_t(U)$, and let $A(t)$ be the n-dimensional volume of U_t. Then

$$A'(t) = \int_{U_t} \Delta\, dx_1 \dots dx_n.$$

PROOF. By the change of variable formula for multiple integrals, we have

$$A(t) = \int_{U_t} dx_1 \dots dx_n = \int_{U_0} J\, da_1 \dots da_n,$$

so

$$A'(t) = \int_{U_0} \frac{\partial J}{\partial t} da_1 \dots da_n = \int_{U_0} \Delta J\, da_1 \dots da_n = \int_{U_t} \Delta\, dx_1 \dots dx_n. \text{ ❖}$$

COROLLARY. If $\Delta = 0$, then the F_t are volume preserving.

Now suppose we have a Hamiltonian H on T^*M and we consider the vector field $\mathbf{X}_H$ on page 533. This gives us a 1-parameter group of diffeomorphisms $\phi_t \colon T^*M \to T^*M$ generated by $\mathbf{X}_H$, the "flow" of $\mathbf{X}_H$. (We might only have a "local flow", a local 1-parameter group of local diffeomorphisms, but for simplicity we will simply speak in terms of flows.)

COROLLARY (LIOUVILLE'S VOLUME THEOREM). The maps ϕ_t of the flow of $\mathbf{X}_H$ are all volume preserving.

PROOF. For

$$F_t(\mathring{q}^1, \dots, \mathring{q}^n, \mathring{p}_1, \dots, \mathring{p}_n) = \phi_t(\mathring{q}^1, \dots, \mathring{q}^n, \mathring{p}_1, \dots, \mathring{p}_n),$$

the x_i are the solutions (q^i, p_i) of (H), and the $x_i{}'$ are the $(\dot{q}^i, \dot{p}_i)$, so

$$\begin{aligned}
\Delta &= \sum_{i=1}^{n} \frac{\partial \dot{q}^i}{\partial q^i} + \sum_{i=1}^{n} \frac{\partial \dot{p}_i}{\partial p_i} \\
&= \sum_{i=1}^{n} \frac{\partial}{\partial q^i}\left(\frac{\partial H}{\partial p_i}\right) + \sum_{i=1}^{n} \frac{\partial}{\partial p_i}\left(-\frac{\partial H}{\partial q^i}\right) = 0. \ ❖
\end{aligned}$$

Although we have stated this result as a theorem on $\mathbb{R}^{2n}$, we can easily restate it on T^*M, in terms of the 2-form ω on T^*M, as on page 512.

Liouville's theorem plays an important role in thermodynamics, where we have to analyze a very large number of particles. If we consider all the particles in some region U_0 at some time t_0, and the region U_t occupied by the same particles at a later time t, Liouville's theorem says that U_t and U_0 have the same volume, and applying this to a small region, we conclude that the density of the particles remains constant. Of course, we should really think in terms of an infinite number of particles in order for the region occupied by a collection of particles to be defined, or else in some probabilistic terms. Fortunately, we don't have to get into that here.

PROBLEMS

550 **1.** (a) In dimension 1, the harmonic oscillator equation $mq'' + \omega^2 q = 0$ describes the motion of a particle under an attractive central force $-\omega^2 q$; the potential $V = \omega^2 q^2/2$, and the Lagrangian is

$$L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - \frac{\omega^2 q^2}{2}.$$

Show (cf. Problem 16-1, applied only to $\dot{q}$, not q!) that the Hamiltonian is

$$H(q, p) = \frac{p^2}{2m} + \frac{\omega^2 q^2}{2}.$$

(b) For a particle in $\mathbb{R}^3$ moving under a force with potential $V(x, y, z)$, the Hamiltonian is

$$H(x, y, z, p_x, p_y, p_z) = \frac{1}{2m}({p_x}^2 + {p_y}^2 + {p_z}^2) + V(x, y, z).$$

(c) For a particle in $\mathbb{R}^2$ moving under a central force with potential function $V(r)$, using coordinates $q^1 = r$, $q^2 = \theta$ where r and θ are the polar coordinates, the Lagrangian is (page 443)

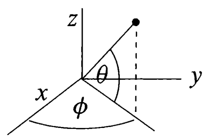
$$L(r, \theta, \dot{r}, \dot{\theta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r).$$

Denoting the corresponding p_1, p_2 by p_r, p_θ, show that the Hamiltonian is

$$H(r, \theta, p_r, p_\theta) = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\phi}^2}{r^2}\right) + V(r).$$

(d) For the same problem in $\mathbb{R}^3$ with coordinates (r, θ, ϕ) defined by

$$x = r\sin\theta\cos\phi$$
$$y = r\sin\theta\cos\phi$$
$$z = r\cos\theta$$



[note that θ is different from that used for the spherical pendulum], we have

$$L(r, \theta, \phi, \dot{r}, \dot{\theta}, \dot{\phi}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2) - V(r),$$

$$H(r, \theta, \phi, p_r, p_\theta, p_\phi) = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2} + \frac{{p_\phi}^2}{r^2\sin^2\theta}\right) + V(r).$$

2. (From Landau and Lifschitz [1].) In terms of the formula on page 530, the partial derivative $\partial S/\partial q^i$ can be written as

$$\frac{\partial S}{\partial q^i} = \left.\frac{d}{du}\right|_{u=0} S(q^1, \dots, q^i + u, \dots, q^n, t) = \left.\frac{d}{du}\right|_{u=0} \int_0^t L\big(c_{(q,t)}{}'(\tau)\big)\, d\tau,$$

where each $c_{(q,t)}$ is a solution of Lagrange's equations. Use the Boundary Term Corollary (page 462) to deduce that $\partial S/\partial q^i(q, t) = p_i(\gamma_q(t))$.

CHAPTER 18
HAMILTON-JACOBI THEORY

man muss immer umkehren
one must always invert
— Jacobi

Jacobi, unlike many of the continental mathematicians, read the *Philosophical Transactions of the Royal Society* regularly, and Hamilton's papers greatly excited him and led him to reconsider the whole subject of dynamics. He looked at all the results from quite a different point of view, which resulted in many further developments.

While Hamilton had regarded the equations (H) on page 529 as a tool for investigating the principal function, Jacobi brought these equations into prominence, anointing them with their now standard name,

Hamilton's canonical equations $$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i},$$

and he then turned his consideration to the equation that is known today as

The Hamilton–Jacobi equation $$\frac{\partial S}{\partial t} + H\left(q^1, \dots, q^n, \frac{\partial S}{\partial q^1}, \dots, \frac{\partial S}{\partial q^n}, t\right) = 0.$$

The complete integral. Rather than trying to solve for S, and then using equation (6) on page 531 to find equations for the $p_i(\gamma_q(t))$, Jacobi approached Hamilton's theory from the point of view of the general theory of first order PDE's, especially in connection with the concept of a "complete integral" for a first order PDE. We will be able to state and prove the main result as soon as we've explained what a complete integral is, though the next two optional sections are included for those who prefer some motivating ideas behind the result.

For a general first order PDE on $\mathbb{R}^n$, with partial derivatives denoted by subscripts, as on page 316,

$$F(x_1, \dots, x_n, u, u_{x_1}, \dots, u_{x_n}) = 0,$$

i.e.,

$$F(x_1, \dots, x_n, u(x_1, \dots, x_n), u_{x_1}(x_1, \dots, x_n), \dots, u_{x_n}(x_1, \dots, x_n)) = 0,$$

we can often "solve" the equation in the sense of finding an n-parameter family of solutions, usually by the method of separation of variables, a method that is often used for higher-order equations also.

For example, in Chapter 8 we considered (pages 314–315) a second order PDE, the 2-dimensional wave equation

$$\frac{\partial^2 u}{\partial t^2}(x,t) = v^2 \frac{\partial^2 u}{\partial x^2}(x,t) \qquad \text{or} \qquad u_{tt} = v^2 u_{xx}$$

by looking for a solution of the form $u(x,t) = X(x)T(t)$, finding that we must have $T''(t)/T(t) = v^2 X''(x)/X(x)$, so that the two sides must be a constant.

For the heat equation for a function u on $\mathbb{R}$,

$$u_{xx} = u_t$$

we obtain similarly $X''(x)/X(x) = T'(t)/T(t) = K$ and thus the solutions

$$u(x,t) = \begin{cases} a \sin\big(\sqrt{-K}(x-b)\big)e^{Kt} & K < 0 \\ a \sinh\big(\sqrt{K}(x-b)\big)e^{Kt} & K > 0. \end{cases}$$

In the case of a first order PDE, which is what interests us now, we usually look for a solution expressed as a sum. As a very simple example, consider the equation

$$u_x{}^2 + u_y{}^2 = 1.$$

Assuming a solution of the form $u(x,y) = \phi(x) + \psi(y)$, we obtain

$$(\phi'(x))^2 = a^2 = 1 - (\psi'(y))^2 \qquad \text{for some constant } a^2,$$

giving us the solutions

$$u(x,y) = ax + \big(\sqrt{1-a^2}\big)y + b$$

(the additive constant b is to be expected, since any solution u gives rise to the solutions $u + b$). Similarly, in the equation

$$V_x{}^2 + V_y{}^2 + V_z{}^2 = 1/v^2$$

on page 524, for constant v we have the solutions

$$V(x,y,z) = \frac{1}{v^2}\big[ax + by + \big(\sqrt{1-a^2-b^2}\big)z + c\big].$$

For a general first order PDE on $\mathbb{R}^n$

$$(1) \qquad 0 = F(x_1, \ldots, x_n, u, u_{x_1}, \ldots, u_{x_n}) = 0$$

we define a **complete integral** to be an n-parameter family of "independent" solutions, that is, a function $\phi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ such that each

$$u(x_1, \ldots x_n) = \phi(x_1, \ldots, x_n, a_1, \ldots a_n)$$

is a solution of (1), and such that ϕ satisfies

$$0 \neq \det\left(\frac{\partial^2\phi}{\partial a_j\,\partial x_k}\right). \tag{2}$$

Condition (2) is actually only assumed to be true in some region about a given point (x, y, a, b), as our entire discussion is local, though for simplicity we will continue to write our equations blithely ignoring this fact.

You can now skip right to page 547 if you want to avoid the computationally involved, yet revealing, motivation for Jacobi's fundamental theorem.

(**Optional**) **Envelopes of solutions.**[1] From the specialized set of solutions given by a complete integral we can obtain a much larger class of solutions by means of envelopes; for simplicity, we will illustrate the construction for the case $n = 2$. Consider the 1-parameter family of solutions

$$u(x, y) = \phi(x, y, a, w(a)),$$

for some given function w. To find an envelope for this family the first step is to "differentiate this equation with respect to a" (compare page 518), to get

$$0 = \phi_a(x, y, a, w(a)) + \phi_b(x, y, a, w(a)) \cdot w'(a).$$

Because of (2), the implicit function theorem implies that if this equation holds for some x_0, y_0, a_0, then in a neighborhood we can solve for a as a function of x and y, i.e., there is a function $A\colon \mathbb{R}^2 \to \mathbb{R}$ such that the equation holds when we replace a by $A(x, y)$, so that

$$\begin{aligned} 0 = \phi_a\big(x, y, A(x, y), w(A(x, y)\big) \\ + \phi_b\big(x, y, A(x, y), w(A(x, y))\big) \cdot w'(A(x, y)). \end{aligned} \tag{3}$$

We then substitute this solution back into $u(x, y) = \phi(x, y, a, w(a))$ to get

$$\Phi(x, y) = \phi\big(x, y, A(x, y), w(A(x, y))\big),$$

and we claim that Φ is also a solution of (1).

---

[1] Although the optional sections are self-contained, they are easier to understand if one already knows the basic facts about first order PDE's, which can be found in the PDE Primer starting on page 667.

To prove this, we note that

$$\Phi_x = \phi_x + [\phi_a + \phi_b w'(A)]a_x$$
$$\Phi_y = \phi_y + [\phi_a + \phi_b w'(A)]a_y,$$

where, in the usual way, the arguments of functions are ruthlessly suppressed. But equation (3) says that the term in brackets vanishes, leaving us with $\Phi_x = \phi_x$ and $\Phi_y = \phi_y$. Since all $u(x, y) = \phi(x, y, a, b)$ satisfy (1), so that

$$0 = F\big(x, y, \phi(x, y, a, b), \phi_x(x, y, a, b), \phi_y(x, y, a, b)\big)$$

for all a and b, this holds in particular for $a = A(x, y)$ and $b = w(A(x, y))$, which just says that Φ also satisfies the equation.[1]

Although we will not go into the details here, it turns out that by considering more than one envelope we can even arrange for a way to obtain the general solution for the first order PDE from a complete integral, with the additional steps involving only "algebraic" manipulations (including solving for implicitly defined functions), rather than ones involving derivatives.

This entire analysis generalizes to the case of a complete integral for a first order PDE on $\mathbb{R}^n$. In this general case, we consider an $(n-1)$-parameter family of solutions

$$u(x_1, \ldots, x_n) = \phi(x_1, \ldots, x_n, a, w_2(a), \ldots, w_n(a))$$

and form the envelope by solving the $n - 1$ equations

$$0 = \phi_a(x, a, w_2(a), \ldots, w_n(a)) + \sum_{i=2}^{n} \phi_{a_i}(x, a, w_2(a), \ldots, w_n(a)) w_i{}'(a)$$

for a function $A\colon \mathbb{R}^n \to \mathbb{R}$ such that the equations hold when we replace a by $A(x_1, \ldots, x_n)$.

(**Optional**) **Inverting the process; contact curves.** Though all of this might be considered of some theoretical interest, it certainly doesn't seem very promising as a way to solve mechanics problems! Even if we have a complete integral for the Hamilton–Jacobi equation, we would have to find envelopes, and solve algebraic relations between them to get a general solution of the equation, before we could use equation (6) on page 531 to find equations for the $p_i(\gamma_q(t))$.

[1] It should be pointed out that this construction of the envelope of a 1-parameter family of a complete integral for a first order PDE on $\mathbb{R}^2$ does not correspond exactly to what we did on page 518 for the 1-dimensional Clairaut equation and on page 519 for the 2-dimensional equation; in those cases we found an envelope of the *entire* family of solutions for the complete integral, giving us the "singular solution" of the Clairaut equation, a topic we need not pursue here.

The necessary PDE background for Jacobi's quite different approach involves the "contact curves" along which the graph of Φ intersects the graph of any one of the functions $(x, y) \mapsto \phi(x, y, a, b)$ in our 1-parameter family.[1] This amounts to saying that we are looking at points of the envelope where $A(x, y)$, and hence $w'(A(x, y))$, are constant. Since A satisfies (3) on page 542 this means that the intersection curve can be parameterized as $x(\sigma), y(\sigma)$, where

$$\begin{aligned} \phi_a(x(\sigma), y(\sigma), a, b) &= C_a\sigma \\ \phi_b(x(\sigma), y(\sigma), a, b) &= C_b\sigma \end{aligned} \tag{4}$$

for certain constants C_a and C_b.

Differentiating equations (4) with respect to σ gives

$$\begin{aligned} \phi_{ax}\frac{dx}{d\sigma} + \phi_{ay}\frac{dy}{d\sigma} &= C_a \\ \phi_{bx}\frac{dx}{d\sigma} + \phi_{by}\frac{dy}{d\sigma} &= C_b. \end{aligned} \tag{5}$$

On the other hand, since ϕ is a complete integral we have

$$0 = F(x, y, \phi(x, y, a, b), \phi_x(x, y, a, b), \phi_y(x, y, a, b))$$

for all a and b, so

$$\begin{aligned} 0 = \frac{\partial F}{\partial a} &= F_u\phi_a + F_p\phi_{xa} + F_q\phi_{ya} \\ 0 = \frac{\partial F}{\partial b} &= F_u\phi_b + F_p\phi_{xb} + F_q\phi_{yb}. \end{aligned}$$

Using (4), and dividing by σF_u, we obtain

$$\begin{aligned} \phi_{xa}(-F_p/\sigma F_u) + \phi_{ya}(-F_q/\sigma F_u) &= C_a \\ \phi_{xb}(-F_p/\sigma F_u) + \phi_{yb}(-F_q/\sigma F_u) &= C_b. \end{aligned} \tag{6}$$

Comparing (5) and (6), and noting that by (2) on page 542 the determinant of these two equations is non-zero, we conclude that

$$\frac{dx}{d\sigma} = -\frac{F_p}{\sigma F_u}, \qquad \frac{dy}{d\sigma} = -\frac{F_q}{\sigma F_u}.$$

[1] We will be obtaining, by direct calculations, results that are obtained as part of the basic theory of solutions for first order PDE's, as in the PDE Primer.

By changing the arbitrary parameterization σ, we can then assume that

(7)
$$\begin{aligned} \frac{dx}{d\sigma} &= F_p \\ \frac{dy}{d\sigma} &= F_q. \end{aligned}$$

In addition, by differentiating (1) on page 541 we get

(8x)
$$0 = F_x + F_u p + F_p \frac{\partial p}{\partial x} + F_q \frac{\partial q}{\partial x}$$

and a similar equation (8y) for y. Now equation (7) shows that

$$\begin{aligned} F_p \frac{\partial p}{\partial x} + F_q \frac{\partial q}{\partial x} &= \frac{\partial p}{\partial x}\frac{dx}{d\sigma} + \frac{\partial q}{\partial x}\frac{dy}{d\sigma} \\ &= \frac{\partial p}{\partial x}\frac{dx}{d\sigma} + \frac{\partial p}{\partial y}\frac{dy}{d\sigma} \qquad \text{since } \frac{\partial q}{\partial x} = \frac{\partial p}{\partial y} \\ &= \frac{dp}{d\sigma}. \end{aligned}$$

Therefore (8x) and the corresponding (8y) yield

(9)
$$\begin{aligned} \frac{dp}{d\sigma} &= -(F_x + F_u p_x) \\ \frac{dq}{d\sigma} &= -(F_y + F_u p_y). \end{aligned}$$

We also have

$$\frac{du}{d\sigma} = u_x \frac{dx}{d\sigma} + u_y \frac{dy}{d\sigma} = pF_p + qF_q \qquad \text{by (7).}$$

Combining this with (7) and (9), we have altogether

(C)
$$\begin{cases} \dfrac{dx}{d\sigma} = F_p, & \dfrac{dy}{d\sigma} = F_q \\ \dfrac{dp}{d\sigma} = -(F_x + F_u p_x), & \dfrac{dq}{d\sigma} = -(F_y + F_u p_y) \\ \dfrac{du}{d\sigma} = pF_p + qF_q. \end{cases}$$

For the general case of a first order PDE

$$0 = F(x_1, \dots, x_n, u, u_{x_1}, \dots, u_{x_n}) = F(x_1, \dots, x_n, u, p_1, \dots, p_n)$$

we find, completely analogously, that the contact curves give us the equations

$$\begin{cases} \text{(C1)} \qquad \dfrac{dx_i}{d\sigma} = F_{p_i} \\[2ex] \text{(C2)} \qquad \dfrac{dp_i}{d\sigma} = -(F_{x_i} + F_u p_i) \\[2ex] \text{(C3)} \qquad \dfrac{du}{d\sigma} = \displaystyle\sum_{i=1}^{n} p_i F_{p_i}. \end{cases}$$

Now consider, in particular, the Hamilton–Jacobi equation, written in the usual notation used for first order PDE's as

$$(*) \qquad \frac{\partial u}{\partial t} + H(x_1, \dots, x_n, p_1, \dots, p_n, t) = 0, \qquad p_i = \frac{\partial u}{\partial x_i}.$$

We have an equation in $n+1$ variables, $x_1, \dots, x_n, t$ and it will be convenient to temporarily write t as x_{n+1}, so that $\partial u/\partial t$ is just p_{n+1}, and we have the equation

$$\begin{aligned} 0 &= F(x_1, \dots, x_n, x_{n+1}, u, p_1, \dots, p_n, p_{n+1}) \\ &= p_{n+1} + H(x_1, \dots, x_n, p_1, \dots, p_n, x_{n+1}). \end{aligned}$$

Note that while the partial derivatives of u appear in the equation, u itself does not, so $F_u = 0$.

In (C1), the equation with $i = n+1$ is just

$$\frac{dx_{n+1}}{d\sigma} = \frac{\partial F}{\partial p_{n+1}} = 1$$

so that we might as well take the parameter σ to be $x_{n+1} = t$. Since $F_u = 0$, we find that the equations (C) become

$$\begin{cases} \text{(C1)} \qquad \dfrac{dx_i}{dt} = \dfrac{\partial H}{\partial p_i} \\[2ex] \text{(C2)} \qquad \dfrac{dp_i}{dt} = -\dfrac{\partial H}{\partial x_i} \\[2ex] \text{(C3)} \qquad \dfrac{du}{dt} = \displaystyle\sum_{i=1}^{n} p_i \frac{\partial H}{\partial p_i} + \frac{\partial u}{\partial t} = \sum_{i=1}^{n} p_i \frac{\partial H}{\partial p_i} - H. \end{cases}$$

Note that (C1) and (C2) are just the canonical equations for H, while (C3) says that du/dt is the Legendre transform of H (compare equation (3) on page 530).

JACOBI'S THEOREM. Now suppose we have found a complete integral of the Hamilton–Jacobi equation, written in the usual notation for first order PDE's, as

$$(*) \qquad \frac{\partial u}{\partial t} + H(x_1, \ldots, x_n, p_1, \ldots, p_n, t) = 0, \qquad p_i = \frac{\partial u}{\partial x_i}.$$

Jacobi's idea was not to solve for u (the principal function S), but *to solve the corresponding canonical equations.*

[This idea is so attractive and looks so promising because, as explained in the optional sections, the canonical equations are precisely the equations for the contact curves that we get when we produce new solutions by taking envelopes, which suggests that in trying to solve the canonical equations directly we might be able to short-circuit the whole process of taking envelopes.]

The equation $(*)$ is special, in that u itself does not appear in the equation, only the partial derivatives of u, which requires a bit of jiggering with the definitions. Clearly if u is a solution of $(*)$, then so is $u + a$ for any constant a. So a complete integral amounts to a function $\phi\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ such that

$$\text{(a)} \qquad u(x_1, \ldots, x_n, t) = \phi(x_1, \ldots, x_n, a_1, \ldots, a_n, t) + a$$

is a solution of $(*)$ for any constants $a_1, \ldots, a_n$ (and a), and condition (2) on page 542 now says that at all points of the region of interest

$$\text{(b)} \qquad 0 \neq \det\left(\frac{\partial^2 \phi}{\partial a_j \partial x_k}\right) \qquad j, k = 1, \ldots, n.$$

With these matters settled, we now consider the set of equations

$$\frac{\partial}{\partial a_j}\phi(x_1, \ldots, x_n, a_1, \ldots, a_n, t) = b_j, \qquad j = 1, \ldots, n$$

for any constants $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, where, with the usual abuse of notation, the symbol $\partial/\partial a_j$ really means the partial derivative of ϕ with respect to its $(n+j)^{\text{th}}$ argument. Because of condition (b), the implicit function theorem implies that if this equation is satisfied for some $(\mathring{x}_1, \ldots, \mathring{x}_n, t_0)$, then in some neighborhood we can "solve implicitly for the x_i as functions of t" that is, there are functions $X_1, \ldots, X_n$ for which

$$\text{(c)} \qquad \frac{\partial}{\partial a_j}\phi\big(X_1(t), \ldots, X_n(t), a_1, \ldots, a_n, t\big) = b_j, \qquad 1 \leq j \leq n.$$

We claim that the X_i automatically satisfy the equations (with arguments on the right omitted, as usual)

$$\frac{dX_i}{dt} = \frac{\partial H}{\partial p_i}.$$

To prove this we first take the partial derivative of (c) with respect to t, to get

(d)
$$\sum_{k=1}^{n} \frac{\partial^2\phi}{\partial a_j\,\partial x_k}\frac{dX_k}{dt} + \frac{\partial^2\phi}{\partial a_j\,\partial t} = 0$$

[$\partial/\partial x_k$ simply signifies the partial derivative with respect to the k^{th} argument].

Next, we note that since (a) is a solution of $(*)$ we have

(e)
$$\frac{\partial\phi}{\partial t} + H\left(X_1,\dots,X_n,\frac{\partial\phi}{\partial x_1},\dots,\frac{\partial\phi}{\partial x_n},t\right) = 0,$$

and taking the derivative with respect to a_j gives

(f)
$$\frac{\partial^2\phi}{\partial a_j\,\partial t} + \sum_{k=1}^{n} \frac{\partial H}{\partial p_k}\frac{\partial^2\phi}{\partial a_j\,\partial x_k} = 0$$

[$\partial/\partial p_k$ signifies the partial derivative with respect to the $(n+k)^{\text{th}}$ argument].

Comparing (d) and (f), we find that

$$\sum_{k=1}^{n} \frac{\partial^2\phi}{\partial a_j\,\partial x_k}\left(\frac{dX_k}{dt} - \frac{\partial H}{\partial p_k}\right) = 0, \qquad 1 \le j \le n,$$

which by condition (b) implies that we do indeed have

(H_1)
$$\frac{dX_k}{dt} = \frac{\partial H}{\partial p_k}.$$

Moreover, if we now set

$$p_j(t) = \frac{\partial}{\partial x^j}(\phi(X_1(t),\dots,X_n(t),a_1,\dots,a_n,t),$$

then differentiating with respect to t gives

(g)
$$\frac{dp_j}{dt} = \sum_{k=1}^{n} \frac{\partial^2\phi}{\partial x_j\,\partial x_k}\frac{dX_k}{dt} + \frac{\partial^2\phi}{\partial x_j\,\partial t},$$

while differentiating (e) with respect to x_j gives

(h)
$$0 = \frac{\partial^2\phi}{\partial x_j\,\partial t} + \frac{\partial H}{\partial x_j} + \sum_{k=1}^{n} \frac{\partial H}{\partial p_k}\frac{\partial^2\phi}{\partial x_k\,\partial x_j}.$$

Subtracting (h) from (g) and using (H_1) just proved, we get

(H_2)
$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i}.$$

Note: Now that we've taken care of all the details, we can afford to be sloppy and write the functions $X_i(t)$ in (c) simply as $x_i(t)$.

The discussion on pages 542–546 provides motivational background, but the straightforward calculations on the previous page give Jacobi's complete proof, in his *Vorlesungen über Dynamik*, of this beautiful theorem, which you will hardly ever find mentioned in a modern mechanics book:

JACOBI'S THEOREM. Given a complete integral ϕ of the Hamilton–Jacobi equation, the solutions of the corresponding canonical equations for H can be found by choosing constants a_j and b_j, solving the equations

$$\frac{\partial\phi}{\partial a_j}\big(x_1(t),\ldots,x_n(t),a_1,\ldots,a_n,t\big) = b_j$$

for x_i, and then setting

$$p_j(t) = \frac{\partial\phi}{\partial x_j}\big(x_1(t),\ldots,x_n(t),a_1,\ldots,a_n,t\big).$$

It might seem that the usefulness of Jacobi's theorem is somewhat limited by the fact that the functions $x_i(t)$ are described only implicitly, by the n equations

$$\frac{\partial\phi}{\partial a_j}\big(x_1(t),\ldots,x_n(t),a_1,\ldots,a_n,t\big) = b_j.$$

Remember, however, that a complete integral of ϕ is usually found, as in the examples below, by separation of variables, so that ϕ is actually of the form

$$\Phi_1(x_1,a_1,\ldots,a_n) + \cdots + \Phi_n(x_n,a_1,\ldots,a_n),$$

which usually makes things a lot simpler.

Jacobi's theorem and mechanics. We will now examine how Jacobi's theorem can be used to study the equations of mechanics, which modern mechanics books usually postpone until the material of the next chapter has been presented.

We switch back to q^i for our basic variables, with the solutions to the canonical equations for H corresponding to the solutions to Lagrange's equations for the Lagrangian L, and to streamline notation we now use S for ϕ, so that Jacobi's theorem states that for a complete integral S of the Hamilton–Jacobi equation, solving the equations

$$\frac{\partial S}{\partial a_j}(q^1(t),\ldots,q^n(t),a_1,\ldots,a_n,t) = b_j$$

for the q^i gives the solutions of Lagrange's equations for L, with the conjugate momenta p_i given by

$$p_j = \frac{\partial S}{\partial q^j}(q^1(t),\ldots,q^n(t),a_1,\ldots,a_n,t).$$

We don't even care what the solution of the Hamilton–Jacobi equation is, or, for that matter, what the equation is actually an equation for!

It also might seem strange that we have replaced the problem of solving a system of n second order equations, or of $2n$ first order differential equations, with the problem of solving a partial differential equation, normally considered to be a harder problem. This will be explained to some extent in Chapter 19, but in any case, for most familiar problems of mechanics a complete integral can be found easily. In fact, in some cases finding a complete integral may be easier, or more straightforward, than solving the corresponding system of ordinary differential equations. In Addendum A we mention some problems that were solved precisely by this method, but for now we simply illustrate how the method works, by considering a few familiar simple cases, where, as might be expected, this method may actually be considerably more cumbersome than the elementary computations.

The Hamiltonians for all these examples have been derived in Problem 17-1.

♦ **Harmonic Oscillator.** Taking $m = 1$ for simplicity, the Hamiltonian is

$$H(q, p) = \tfrac{1}{2}(p^2 + \omega^2 q^2),$$

so the Hamilton–Jacobi equation is a partial differential equation in two variables q and t,

(a) $$\frac{\partial S}{\partial t} + \frac{1}{2}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{1}{2}\omega^2 q^2 = 0.$$

Now t is "cyclic"—the variable t itself doesn't appear in the equation, only derivatives with respect to t—which suggests that as a first step in finding a complete integral we separate out t by looking for a function S of the form

$$S(q, t) = W(q) + f(t).$$

This will be a solution if

$$f'(t) + H\left(q, \frac{\partial W}{\partial q}\right) = 0 \iff f'(t) = -\alpha = H(q, \partial W/\partial q)$$

for some constant α, so that we get the equations

(a1) $$S(q, \alpha, t) = W(q, \alpha) - \alpha t$$

(a2) $$H\left(q, \frac{\partial W(q, \alpha)}{\partial q}\right) = \alpha.$$

We write $S(q, \alpha, t)$ and $W(q, \alpha)$ even though α is a constant, because we are trying to find a general integral S for (a), that is, a solution depending on a parameter α. Since H doesn't depend on t, we will have $\alpha = E$ for the solution we ultimately obtain (page 529).

Equation (a2) is

$$\frac{1}{2}\left(\frac{\partial W(q,\alpha)}{\partial q}\right)^2 + \frac{1}{2}\omega^2 q^2 = \alpha,$$

giving

$$W(q,\alpha) = \omega \int \sqrt{\frac{2\alpha}{\omega^2} - q^2}\, dq, \qquad S(q,\alpha,t) = \omega \int \sqrt{\frac{2\alpha}{\omega^2} - q^2}\, dq \;-\; \alpha t.$$

There's no point doing the integration right now, since we want to define $q(t)$ implicitly by the formula

$$\frac{\partial S(q(t),\alpha,t)}{\partial \alpha} = b$$

for a constant b (here, of course, it is definitely necessary to recognize the dependence of S on α!). So we want

(b) $$b = \frac{\partial S(q(t),\alpha,t)}{\partial \alpha} = \frac{1}{\omega}\int \frac{dq}{\sqrt{\dfrac{2\alpha}{\omega^2} - (q(t))^2}} - t,$$

giving

$$t + b = -\frac{1}{\omega}\arccos\left(\frac{\omega}{\sqrt{2\alpha}}\cdot q(t)\right)$$

or, finally setting $\alpha = E$,

(b′) $$q(t) = \frac{\sqrt{2E}}{\omega}\cos\omega(t+b).$$

The constants E and b can be determined from the initial values, say the values q_0, p_0 of q, p at $t = 0$ (but only once the problem has been solved). Evaluating (b′) at $t = 0$ gives $0 = -E + \frac{1}{2}{p_0}^2 + \frac{1}{2}\omega^2 q_0^2$ and substituting this value of E into (b′) gives an equation for $\cos\omega b$, allowing b to be determined. In the particular case where $t = 0$ is a maximum or minimum point, $p_0 = 0$, we simply have $E = \omega^2 {q_0}^2/2$, and (b′) gives $q_0 = q_0\cos\omega(t+b)$, so we can just take $b = 0$.

Although this hardly looks like a more convenient way of analyzing the harmonic oscillator, it has the virtue of following a systematic method that doesn't depend on any previous knowledge, or any guessing, about the solution. The next example, in more than one dimension, makes a somewhat better case.

♦ **Central force in polar coordinates.** In this example, we will now separate out not only t, but also one of the two space variables. We have

$$L(r,\theta,\dot r,\dot\theta) = \frac{1}{2}m(\dot r^2 + r^2\dot\theta^2) - V(r),$$

$$H(r,\theta,p_r,p_\theta) = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2}\right) + V(r).$$

Once again separating out t, we seek a solution of the Hamilton–Jacobi equation of the form

(a1) $$S(r,\theta,\alpha,t) = W(r,\theta,\alpha) - \alpha t$$

(a2) $$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial\theta}\right)^2\right] + V(r) = \alpha.$$

Now θ, like t, is cyclic, leading us to look for a solution of (a2) of the form

$$W(r,\theta,\alpha,\alpha_\theta) = R(r) + \alpha_\theta\cdot\theta$$

for a constant α_θ with

$$\alpha_\theta = \frac{\partial W}{\partial\theta} = \frac{\partial S}{\partial\theta} = p_\theta \; (= mr^2\dot\theta).$$

Then

$$\frac{1}{2m}\left[\big(R'(r)\big)^2 + \frac{1}{r^2}{\alpha_\theta}^2\right] + V(r) = \alpha,$$

hence

$$W(r,\theta,\alpha,\alpha_\theta) = \int\sqrt{2m(\alpha - V(r)) - \frac{{\alpha_\theta}^2}{r^2}}\,dr + \alpha_\theta\theta.$$

So for two constants b_1, b_2 we want

(b1) $$b_1 = \frac{\partial W}{\partial\alpha} = \int\frac{dr}{\sqrt{\dfrac{2}{m}(\alpha - V(r)) - \dfrac{{\alpha_\theta}^2}{r^2}}} - t$$

(b2) $$b_2 = \frac{\partial W}{\partial\alpha_\theta} = -\int\frac{\alpha_\theta\,dr}{r^2\sqrt{\dfrac{2}{m}(\alpha - V(r)) - \dfrac{{\alpha_\theta}^2}{r^2}}} + \theta.$$

Since $\alpha = E$, equation (b1), with $\alpha_\theta = h$, is equivalent to (A) on page 121, giving r as a function of t, and (b2) is equivalent to the equation on the top of page 122, determining the shape of the orbit, with $b_2 = \theta(0)$.

♦ **Central force in spherical polar coordinates.** We now consider the central force problem in 3-dimensions, mainly as an illustration that a complete integral may sometimes be found even when more than one variable is not cyclic. Using the coordinates on page 539, we have

$$L(r,\theta,\phi,\dot r,\dot\theta,\dot\phi) = \frac{1}{2}m(\dot r^2 + r^2\dot\theta^2 + r^2\cos^2\theta\,\dot\phi^2) - V(r)$$

$$H(r,\theta,\phi,p_r,p_\theta,p_\phi) = \frac{1}{2m}\left({p_r}^2 + \frac{{p_\theta}^2}{r^2} + \frac{{p_\phi}^2}{r^2\cos^2\theta}\right) + V(r),$$

leading to

(a1) $$S(r,\theta,\phi,\alpha,t) = W(r,\theta,\phi,\alpha) - \alpha t$$

(a2) $$\frac{1}{2m}\left[\left(\frac{\partial W}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial W}{\partial \theta}\right)^2 + \frac{1}{r^2\cos^2\theta}\left(\frac{\partial W}{\partial \phi}\right)^2\right] + V(r) = \alpha.$$

The variable ϕ is cyclic, but r and θ are not. Nevertheless, we still look for a solution of the form

$$W(r,\theta,\phi) = R(r) + \Theta(\theta) + \Phi(\phi).$$

Since ϕ is cyclic, we have, as in the previous example, $\Phi(\phi) = \alpha_\phi \cdot \phi$, for the constant α_ϕ, which is just the constant conjugate momentum p_ϕ. We then have

$$r^2\left[R'^2 + 2m\big(V(r)-\alpha\big)\right] = -\left[\Theta'^2 + \frac{{\alpha_\phi}^2}{\cos^2\theta}\right].$$

Since the left side depends only on r and right side only on θ, we set them both equal to a constant $-{\alpha_\theta}^2$, so that we have

(a3) $$\Theta'^2 + \frac{{\alpha_\phi}^2}{\cos^2\theta} = {\alpha_\theta}^2, \qquad R'^2 + 2m\big(V(r)-\alpha\big) + \frac{{\alpha_\theta}^2}{r^2} = 0,$$

and thus $W = W(r,\theta,\phi,\alpha,\alpha_\theta,\alpha_\phi)$ is given by

$$W = \sqrt{2m}\int\sqrt{\alpha - V(r) - \frac{{\alpha_\theta}^2}{2mr^2}}\,dr + \int\sqrt{{\alpha_\theta}^2 - \frac{{\alpha_\phi}^2}{\cos^2\theta}}\,d\theta + \alpha_\phi\phi.$$

For now, we simply use this example as an illustration of the process of finding complete integrals, leaving the intricacies of the solution to Problem 1.

♦ **Arbitrary force on a particle.** For a particle of mass m under a force with potential $V(x,y,z)$, where (Problem 17-1 (b)) we have

$$H(x,y,z,p_x,p_y,p_z) = \frac{1}{2m}({p_x}^2 + {p_y}^2 + {p_z}^2) + V(x,y,z),$$

we now obtain the equation

$$\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 + \left(\frac{\partial W}{\partial z}\right)^2 = 2m(E-V).$$

We don't try to solve this in general, of course, but we want to point out that when $V = 0$, this has the same form as equation (C) on page 524 for the characteristic function in optics. For $W(x,y,z) = X(x) + Y(y) + Z(z)$ we get $X' = a_x$, $Y' = a_y$, $Z' = a_z$ for constants a_x, a_y, and a_z, and thus $p_x = a_x$, giving $x = a_x t + b_x$ for $x(0) = b_x$ and $x'(0) = a_x$, and similarly for y and z.

Hamilton's characteristic function. In all of these examples, as in most problems of mechanics, the variable t didn't appear in the Hamiltonian, so that we had an equation of the form

$$\text{(a)} \qquad \frac{\partial S}{\partial t} + H\left(q^1,\dots,q^n,\frac{\partial S}{\partial q^1},\dots,\frac{\partial S}{\partial q^n}\right) = 0,$$

and we separated out the variable t to get

$$\text{(a1)} \qquad S(q^1,\dots,q^n,t) = W(q^1,\dots,q^n) - \alpha t$$

$$\text{(a2)} \qquad H\left(q^1,\dots,q^n,\frac{\partial W}{\partial q^1},\dots,\frac{\partial W}{\partial q^n}\right) = \alpha,$$

with α normally being the energy E of our path.

The function W has already appeared in another context. Note that in equation (3) on page 530 the term $H(\gamma_q(\tau))$ is simply E, since we are assuming H does not depend on t, so

$$\begin{aligned} W(q,t) &= S(q,t) + Et \\ &= \int_0^t \sum_{j=1}^n (p_j \circ \gamma_q)(\tau)(q^j \circ \gamma_q)'(\tau) \;-\; H(\gamma_q(\tau))\,d\tau + \int_0^t H(\gamma_q(\tau))\,d\tau \\ &= \int_0^t \sum_{j=1}^n (p_j \circ \gamma_q)(\tau)(q^j \circ \gamma_q)'(\tau)\,d\tau, \end{aligned}$$

which is just the integral of the action (page 464) along the curve. As mentioned on page 527, Hamilton originally used this integral as his "characteristic function" for mechanics, which he later discarded in favor of his principal function S. So (a2), known as the "reduced Hamilton–Jacobi equation", is also referred to as the "Hamilton–Jacobi equation for Hamilton's characteristic function".

HAMILTON-JACOBI THEORY AND THE SCHRÖDINGER WAVE EQUATION. While Jacobi's presentation of Hamilton's work had the beneficial effect of simplifying many of Hamilton's constructions and introducing Hamiltonian mechanics to a much wider audience, it also had the unfortunate side-effect that the connection with Hamilton's earlier optical investigations was almost completely neglected.[1]

This connection was only brought back into prominence with the advent of the Schrödinger wave equation. Building on de Broglie's notion that a moving particle with momentum p and energy E has associated with it a "wave" with wave length λ and frequency ν satisfying

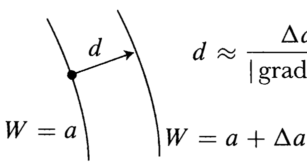
$$\lambda = h/p, \quad \nu = E/h \qquad \text{for Planck's constant } h,$$

Schrödinger sought a wave-like equation to explain the emission spectra of atoms, which have the quantum mechanical aspect of taking on only discrete values. After publishing his basic papers in German physics journals in 1926, at the end of that year he published an English review article, Schrödinger [1], outlining how one can be led to his equation by exploring the correspondence between Hamilton's optical and mechanical theories.

Schrödinger begins with the simple example of a particle of mass m acted on by a force with potential function V, as on page 554, where the equation for Hamilton's characteristic function W can be written as

$$|\operatorname{grad} W| = |(W_x, W_y, W_z)| = \sqrt{\left(\frac{\partial W}{\partial x}\right)^2 + \left(\frac{\partial W}{\partial y}\right)^2 + \left(\frac{\partial W}{\partial z}\right)^2} = \sqrt{2m(E-V)}.$$

Since grad W is always perpendicular to the sets $W =$ constant, for small Δa the perpendicular distance d between a point of $W = a$ and the set $W = a + \Delta a$

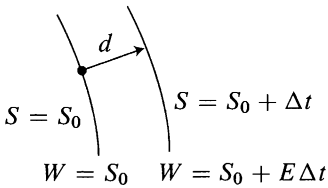


$$d \approx \frac{\Delta a}{|\operatorname{grad} W|}$$

is approximately $\Delta a/|\operatorname{grad} W|$ evaluated at that point.

[1] Leading to the historical irony that Bruns [1] in 1895 reinvented Hamilton's characteristic function for optics, which he called the *eikonal*, from the Greek εικῶν = image, the term now standard in optics, and apparently unaware of Hamilton's optical work, though acquainted with Hamiltonian mechanics, he remarked that "The eikonal concept now plays an entirely similar role as the Hamiltonian viewpoint in mechanics, though, to be sure, in the far narrower domain of geometrical optics."

We can also consider the surfaces $S = \text{constant}$, though these constants are out of step with the constants for the surfaces $W = \text{constant}$. At time $t = 0$ the surface $S = S_0$ is the same as the surface $W = S_0$, while the surface $S = S_0 + \Delta t$ is the same as the surface $W = S_0 + E\Delta t$. So, if we think of the surfaces $S = t$



$$d \approx \frac{E\Delta t}{|\operatorname{grad} W|}$$

as parameterized by time t, after a short time Δt, a point on the surface $S = S_0$ will have moved a distance of approximately $E\Delta t/|\operatorname{grad} W|$, with a velocity close to

$$(1) \qquad v = \frac{E\Delta t}{|\operatorname{grad} W|\Delta t} = \frac{E}{\sqrt{2m(E - V)}},$$

which plays the role of the "velocity" v in our wave equations on pages 315–317 and in Addendum 15B, except that it is no longer a constant.

[Since

$$\frac{\partial W}{\partial x} = \frac{\partial S}{\partial x} = p_x, \qquad \frac{\partial W}{\partial y} = \frac{\partial S}{\partial y} = p_y, \qquad \frac{\partial W}{\partial z} = \frac{\partial S}{\partial p} = p_z,$$

the vector grad W is just the momentum vector $\mathbf{p}$ of the particle, so this velocity v varies inversely with the velocity of the particle, $|\operatorname{grad} W|$; as on page 526 (with S playing the role the optical characteristic function V), when the wave is moving fast, with v large, the "wave fronts" $S = \text{constant}$ are close together, so that $\mathbf{p}$ is small.]

Schrödinger concludes the first section of his paper by noting Hamilton's correlation of Huygens' construction with Fermat's principle, and begins the second section by saying:

> Nothing of what has hitherto been said is in any way new. All this was very much better known to Hamilton himself than it is in our day to a good many physicists. Indeed, the theory of the propagation of light in a non-homogeneous medium, which Hamilton had developed about ten years earlier, became, by the striking analogy which occurred to him, the starting-point for his famous theories in pure mechanics. Notwithstanding the great popularity reached by the latter, the way which had led to them was nearly forgotten.

This optical-mechanical correspondence, however, really applies only to geometrical optics, which is merely an approximation to the wave theory of optics, and Schrödinger compares this to the fact that

> ... ordinary mechanics is really not applicable to mechanical systems of the very small, viz. of atomic dimensions. Taking into account this fact, which impresses its stamp upon all modern physical reasoning, is one not greatly tempted to investigate whether the non-applicability of ordinary mechanics to micro-mechanical problems is perhaps of exactly the same kind as the non-applicability of geometrical optics to the phenomena of diffraction or interference and may, perhaps, be overcome in an exactly similar way?

Answering his own question in the positive, he notes that

> ... Well known methods of wave-theory, somewhat generalized, lend themselves readily. The conceptions, roughly sketched in the preceding are fully justified by the success which has attended their development.

Just as the equation for the vibrating string with fixed ends (pages 314–315) gives us solutions $A(x)\sin(\omega_n t+\phi)$ for suitable constants ω_n, Schrödinger looks for a wave-function ψ of the form

$$\begin{aligned}\psi &= A(x,y,z)\sin(S/K)\\ &= A(x,y,z)\sin\big(-Et/K + W(x,y,z)/K\big),\end{aligned}$$

for some constant K. Since S has the dimensions ET of action, the constant K must also have these dimensions, which is the siren call of Planck's constant h. In fact, since the frequency of our wave is

$$\nu = E/2\pi K,$$

we can get de Broglie's relationship $\nu = E/h$ by choosing $K = h/2\pi$.

Now we want to look at the wave equation

$$\Delta\psi - \psi''/v^2 = 0,$$

and substitute for v the expression given by equation (1). Since ψ depends only on the space coordinates and on the frequency E/h, we want the same to be true for ψ''. This means that the dependence of ψ on time should involve only the factor $e^{\pm 2\pi i t E/h}$, so that

$$\psi'' = -4\pi^2 E^2\psi/h^2,$$

leading finally to the *Schrödinger wave equation*:

$$\boxed{\Delta\psi + 8\pi^2 m(E - V)\psi / h^2 = 0}$$

This is not, of course, a derivation, but something closer to a divination, fitting quite comfortably into the tradition of the investigators into the theory of light who preceded him. And it worked:

> Putting for instance
>
> $$V = -e^2/4,$$
>
> (e = electronic charge, $r = (x^2 + r^2 + z^2)^{\frac{1}{2}}$, we get for the simplified hydrogen atom or one body problem:
>
> $$\Delta\psi + \pi^2 m(E + e^2/r)\psi / h^2 = 0.$$
>
> Now this equation for a great part of the possible values of the energy or frequency constant E, proves to offer no solution at all which is continuous, finite and single-valued throughout the whole space [and approaching 0 at ∞. The set of negative values that do,]
>
> $$E = -2\pi^2 me^4 / h^2 n^2 \qquad (n = 1, 2, 3, 4 \cdots)$$
>
> ... corresponds exactly to Bohr's stationary energy levels of the elliptic orbits.

Which is all we will have to say on this matter, for a long, long time.

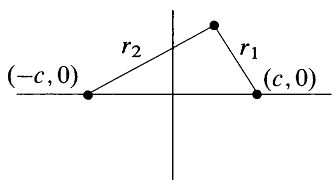
ADDENDUM 18A

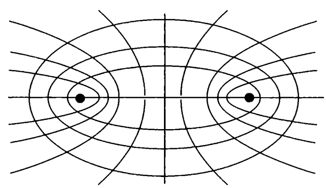
MOTION IN THE FIELD OF TWO FIXED MASSES

GEODESICS ON ELLIPSOIDS

We consider the problem, posed by Euler in 1760, of a body moving in the gravitational field provided by two fixed masses. That would seem to be only of theoretical interest, since the two bodies certainly won't remain fixed, so that it is more like a poor-man's version of the restricted three-body problem considered in Addendum 10A, where the two bodies remain at a fixed distance from each other. On the other hand, it can be applied, for example, to the problem of a negatively charged body moving in the field of two fixed positively charged bodies. The solution even has a use in modern-day problems involving gravitation, as we will mention a little later.

Jacobi's method for solving this problem depends on a suitable standard coordinate system called elliptic coordinates. We will consider only motion in a plane, with the two bodies located at $(c,0)$ and $(-c,0)$, and for simplicity assume the two bodies have the same mass m, with the third body simply having mass 1. If r_1 and r_2 are the distances from a point in the plane to the points





$(c,0)$ and $(-c,0)$, we define

$$\lambda = \tfrac{1}{2}(r_1 + r_2), \qquad \mu = \tfrac{1}{2}(r_1 - r_2),$$

with the curves $\lambda =$ constant being ellipses having the fixed points as foci, and the curves $\mu =$ constant being one branch of a hyperbola with the same foci.

A not inconsiderable amount of grunt work is first required to obtain the basic facts about these coordinates. From

$$r_1{}^2 = (x-c)^2 + y^2, \quad r_2{}^2 = (x+c)^2 + y^2$$

we get

(a) $$x^2 + y^2 + c^2 = \tfrac{1}{2}(r_2{}^2 + r_1{}^2) = \lambda^2 + \mu^2$$

(b) $$cx = \tfrac{1}{4}(r_2{}^2 - r_1{}^2) = -\lambda\mu.$$

Substituting $\mu = -cx/\lambda$ into (a) then gives

(c) $$c^2y^2 = (\lambda^2 - c^2)(c^2 - \mu^2).$$

If the path of the body has coordinates $x(t), y(t)$ and corresponding $\lambda(t), \mu(t)$, we then have

$$\frac{y'}{y} = \frac{\lambda\lambda'}{\lambda^2 - c^2} - \frac{\mu\mu'}{c^2 - \mu^2},$$

$$c^2 y'^2 = (\lambda^2 - c^2)(c^2 - \mu^2)\left(\frac{\lambda\lambda'}{\lambda^2 - c^2} - \frac{\mu\mu'}{c^2 - \mu^2}\right)^2,$$

leading finally to

$$\begin{aligned} T &= \tfrac{1}{2}(x'^2 + y'^2) \\ &= \frac{1}{2c^2}\left[(\lambda\mu' + \lambda'\mu)^2 + (\lambda^2 - c^2)(c^2 - \mu^2)\left(\frac{\lambda\lambda'}{\lambda^2 - c^2} - \frac{\mu\mu'}{c^2 - \mu^2}\right)^2\right] \\ &= \tfrac{1}{2}(\lambda^2 - \mu^2)\left(\frac{\lambda'^2}{\lambda^2 - c^2} + \frac{\mu'^2}{c^2 - \mu^2}\right), \end{aligned}$$

while for the potential function V we have (taking the constant of gravitation as 1 for convenience)

$$V = -\left(\frac{m}{r_1} + \frac{m}{r_2}\right) = -\left(\frac{m}{\lambda + \mu} + \frac{m}{\lambda - \mu}\right) = -\frac{2m}{\lambda^2 - \mu^2}.$$

Thus the Hamiltonian is

$$H = \tfrac{1}{2}p_\lambda{}^2\frac{\lambda^2 - c^2}{\lambda^2 - \mu^2} + \tfrac{1}{2}p_\mu{}^2\frac{c^2 - \mu^2}{\lambda^2 - \mu^2} - \frac{2m\lambda}{\lambda^2 - \mu^2},$$

and the reduced Hamilton–Jacobi equation is

$$\left(\frac{\partial W}{\partial \lambda}\right)^2(\lambda^2 - c^2) + \left(\frac{\partial W}{\partial \mu}\right)^2(c^2 - \mu^2) = 2(\lambda^2 - \mu^2)\alpha + 4m\lambda,$$

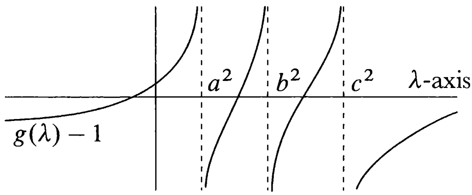
which is easy to separate, writing $W(\lambda, \mu) = \Lambda(\lambda) + M(\mu)$. The resulting equations give explicit expressions for the motion of the body in terms of elliptic integrals.

Considerable discussion of the orbits, where the two fixed bodies also need not have the same mass, is given in Pars [1; §17.10]. This problem is also treated in Boccaletti and Pucacco [1; §5.6], with §5.7 explaining how the problem can be applied to satellites in the gravitational field of the oblate-spheroid shaped earth, which is approximated by two fixed bodies whose masses and distances from the center are imaginary numbers!

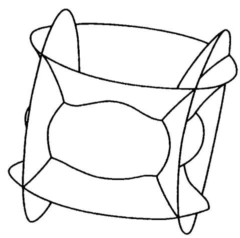
Jacobi also defined elliptical coordinates for 3-dimensional space; they are the three roots λ_1, λ_2, and λ_3 of the equation

$$(*)\quad g(\lambda) = \frac{x^2}{a^2-\lambda} + \frac{y^2}{b^2-\lambda} + \frac{z^2}{c^2-\lambda} = 1, \qquad \text{for given } 0 < a^2 < b^2 < c^2.$$

For $\lambda < a^2$ we obtain ellipsoids, for $a^2 < \lambda < b^2$ hyperboloids of one sheet, and for $b^2 < \lambda < c^2$ hyperboloids of two sheets. For any (x, y, z) with $x, y, z \neq 0$, the function $g(\lambda) - 1$ clearly must be 0 for at least one $\lambda_1 < a^2$, one λ_2 with



$a^2 < \lambda_2 < b^2$, and one λ_3 with $b^2 < \lambda_3 < c^2$. There are only 3 roots, since $g(\lambda) - 1 = 0$ is equivalent to a cubic equation in λ. Thus one surface from each family passes through each such point (x, y, z). At a point (x, y, z) on the



surface $g(\lambda_i) = 1$, the normal vector has the direction

$$\frac{1}{2}(D_1 g(\lambda_i), D_2 g(\lambda_i), D_3 g(\lambda_i)) = \left(\frac{x}{a^2-\lambda_i}, \frac{y}{b^2-\lambda_i}, \frac{z}{c^2-\lambda_i}\right).$$

At a point (x, y, z) on the two surfaces $g(\lambda_i) = 1$ and $g(\lambda_j) = 1$, the inner product of the two normal vectors is therefore

$$\frac{x^2}{(a^2-\lambda_i)(a^2-\lambda_j)} + \frac{y^2}{(b^2-\lambda_i)(b^2-\lambda_j)} + \frac{z^2}{(c^2-\lambda_i)(c^2-\lambda_j)}$$

which can be written as

$$\frac{g(\lambda_i) - g(\lambda_j)}{\lambda_j - \lambda_i} = 0,$$

so our system is orthogonal.

Jacobi used these coordinates for the purely geometric problem of geodesics on ellipsoids. A description of his results may be found in Arnold [2; §47]; the equations themselves are derived in Fasano and Marmi [1; §11.2].

ADDENDUM 18B

HUYGENS' CONSTRUCTION FOR HYPERBOLIC EQUATIONS

We consider second order PDE's on $\mathbb{R}^n \times \mathbb{R}$ of the form

$$(*) \qquad \sum_{i,j=1}^{n} a_{ij}(x) u_{x_i x_j}(x,t) = u_{tt}(x,t),$$

where the symmetric matrix $A = (a_{ij})$ is everywhere positive definite, so that at each point it can be reduced to the wave equation in $\mathbb{R}^n$. For $n = 3$, equation $(*)$ is the general wave equation in 3-space for a possibly non-homogeneous, non-isotropic medium, as discussed on pages 494-495, in the case where our indicatrices are simply ellipsoids.

The basic theory of such PDE's leads to a result that truly deserves the name of *Huygens' construction*, though it is also sometimes called Huygens' principle. We will only state the result here, but even the statement has been postponed until now because it relies on ideas that are illustrated to some extent in the PDE Primer at the end of this book, which was first mentioned in this chapter.

A surface in $\mathbb{R}^3$, or more generally an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$, is called a **characteristic surface** when there are two different solutions of $(*)$ that are tangent along the surface, or equivalently, if initial conditions along the surface, together with initial conditions for the first derivative along the surface in the direction of a normal vector field, do not uniquely determine the solution in the region to which the normal vector field points. They play a role for second order PDE's analogous to that of the characteristic curves for first order PDE's.

Physically, characteristic surfaces are of interest because they represent "wave fronts": if we consider a solution u of the wave equation representing the result of a disturbance beginning at a point (or more generally a closed set) at time $t = 0$, then at time t the solution will be 0 outside of some surface W_t, the wave front at time t, so W_t will be a characteristic surface, since the 0 solution and the solution u will agree along W_t.

Now suppose that for each point $x \in W_t$ we consider the solution of the same wave equation for a disturbance starting at x at time $t = 0$, and look at its wave front w_x at time Δt. Then the envelope of all w_x has two components, and it is a theorem that the one that lies in the region where $u = 0$ is the wave front $W_{t+\Delta t}$. This situation is again analogous to various considerations for first order PDE's discussed in the PDE Primer.

What about the other component of the envelope, the one within the region that the wave has already reached? This question presents a problem only if we allow the purely mathematical content of this result to be confused with some presumed physical mechanism. Courant and Hilbert [1; Chap. VI, §1.9] points out that though we find "what seems at first sight to be a paradox", it is easily put to rest by the observation that "A characteristic surface can, *but need not*, contain discontinuities of the solution u, and the envelope construction can also lead, without contradicting our theory, to surfaces on which the wave is not discontinuous at the time t." [Italics mine.]

This sort of resolution was not available in the original, very mechanistic, theories of light. After Huygens' description of secondary waves (page 491), he adds that it would seem that the particles of the ether should all be of equal size, "because otherwise there ought to be some reflexion of movement backwards when it passes from a smaller particle to a larger one".

When Fresnel introduced Huygens' hypothesis of secondary waves into his transversal wave theory of light in order to account for diffraction, Huygens' mechanistic theory itself wasn't applicable, and Fresnel had to introduce an elaborate ingenious argument, together with just the right *ad hoc* assumptions, to show that the interfering secondary waves would exactly cancel out in the backwards direction.

Once the theory of light became part of Maxwell's theory of electromagnetic waves, the idea of secondary waves became not only unnecessary, but basically absurd: electromagnetic waves are emitted by accelerating charges, not by other electromagnetic waves. Then, of course, the problem became to explain why Fresnel's calculations did work, even though they were based on a fairy-tale construction. This was accomplished by Kirchhoff, who showed that Fresnel's construction could be regarded as an approximate from of an integral theorem.

Detailed explanations of Fresnel's constructions, and of Kirchhoff's integral theorem, may be found in Born and Wolf [1; §§8.1–8.3].

PROBLEM

1. Consider the central force problem in spherical polar coordinates, with W as given on page 553.

(a) From the formula for L we have

$$p_\phi = mr^2 \cos^2\theta\,\dot{\phi}.$$

Calculate that this is the vertical component of the angular momentum $\mathbf{L}$.

Then use (a3) to show that

$$\alpha_\theta{}^2 = p_\theta{}^2 + \frac{p_\phi{}^2}{\cos^2\theta} = m^2r^4(\dot{\theta}^2 + \cos^2\theta\,\dot{\phi}^2),$$

and conclude that α_θ is the length L of the angular momentum $\mathbf{L}$.

We will henceforth set

$$\begin{aligned} \alpha &= E && \text{the constant energy of the final solution,}\\ \alpha_\theta &= L, && \\ \alpha_\phi &= L_3 && \text{as a convenient abbreviation.} \end{aligned}$$

(b) For three constants b_1, b_2, b_3, we have the equations

(b1) $$b_1 = \frac{\partial W}{\partial E} = \sqrt{\frac{m}{2}} \int \frac{dr}{\sqrt{E - V(r) - \dfrac{L^2}{2mr^2}}} - t,$$

(b2) $$b_2 = \frac{\partial W}{\partial L} = \frac{1}{\sqrt{2m}} \int \frac{-L\,dr/r^2}{\sqrt{E - V(r) - \dfrac{L^2}{2mr^2}}} + \int \frac{L\,d\theta}{\sqrt{L^2 - \dfrac{L_3{}^2}{\cos^2\theta}}},$$

(b3) $$b_3 = \frac{\partial W}{\partial L_3} = \int \frac{-L_3\,d\theta/\sin^2\theta}{\sqrt{L^2 - \dfrac{L_3{}^2}{\cos^2\theta}}} + \phi.$$

(c) Writing (b3) as

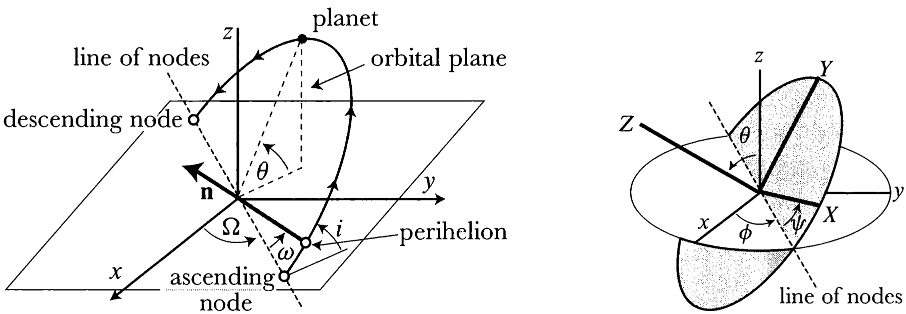
$$\phi - b_3 = \int \frac{\sec^2\theta \, d\theta}{\sqrt{\left(\dfrac{L^2}{{L_3}^2} - 1\right) - \tan^2\theta}},$$

and noting that $\sec^2\theta = d(\tan\theta)$, conclude that

(i) $$\tan\theta = \sqrt{\left(\frac{L^2}{{L_3}^2} - 1\right)} \sin(\phi - b_3),$$

which is a plane orbit, since it is linear relation between the direction cosines $(\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta)$.

To connect this with the standard astronomical terminology, consider the following diagram (with the similar diagram for the Euler angles on the right).



The horizontal plane on the left is the "plane of the ecliptic", the plane in which the earth's orbit lies, with the x-axis chosen so that it points toward the vernal equinox (which is used to denote both a direction, and a time). The angle i is the "inclination", while Ω, a.k.a. ϕ, is the "longitude of the ascending node", and ω, a.k.a. ψ, is the "argument of the perihelion". We've also added the normal $\mathbf{n}$ to the orbital plane, which is not a standard element of the astronomical picture.

Show that

$$\mathbf{n} = (\sin i \sin\Omega, -\sin i \cos\Omega, \cos i),$$

and letting

$$\mathbf{x} = r(\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta)$$

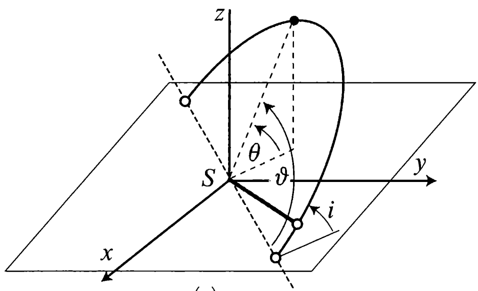
(all letters in both of these equations are functions of time t), show that $\langle \mathbf{n}, \mathbf{x} \rangle = 0$ is equivalent to

(ii) $$\tan\theta = \tan i \sin(\phi - \Omega);$$

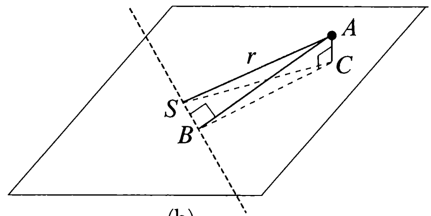
comparison of (i) and (ii) then shows that

(iii) $$b_3 = \Omega, \qquad \cos i = \frac{L_3}{L}.$$

(d) In (a) of the figure below, we show the angle ϑ from the line of nodes to the planet, measured in the plane of the orbit. Part (b) shows the planet at



(a)



(b)

position A, at distance r from the sun S. Note that $\angle ABC = i$. Determine AC and AB, and conclude that

(iv) $$\sin\theta = \sin i \sin\vartheta.$$

Using (iii) and (iv), write the second integral on the right of (b2) as

$$\int \frac{L\,d\theta}{\sqrt{L^2 - L_3{}^2\sec^2\theta}} = \int \frac{\cos\theta\,d\theta}{\sqrt{\cos^2\theta - \cos^2 i}},$$

and then use the substitution (iv) to show that this integral is simply ϑ, so that equation (b2) becomes

$$\vartheta - b_2 = \frac{L}{\sqrt{2m}} \int \frac{dr}{r^2\sqrt{E - V(r) - \dfrac{L^2}{2mr^2}}},$$

which is the first equation at the top of page 122 for the polar coordinates (r, ϑ) in the orbital plane.

For $V = -mK/r$, we have the solution on pages 123–124, from which we see that $b_2 = \omega$, the value of ϑ for the perihelion. Equation (b1) then shows that b_1 is the time t_0 at which the planet is at the perihelion.

CHAPTER 19

CANONICAL TRANSFORMATIONS

> ... The advantages of the Hamiltonian formulation lie not in its use as a calculational tool, but rather in the deeper insight it affords into the formal structure of mechanics. ... we are led to newer, more abstract ways of presenting the physical content of mechanics.
>
> — Goldstein, *Classical Mechanics*

This disclaimer, appearing at the beginning of the chapter on canonical transformations in Goldstein's standard text *Classical Mechanics*, serves as a warning to physicists that a barrage of mathematics is about to ensue, long before any physics makes an appearance. But for our readers, an apology is presumably not required for beginning with an extended mathematical presentation. We will take the time to consider several different approaches to the study of canonical transformations, and their interconnections.

Canonical transformations. Our oft-used device of choosing new coordinates to simplify problems can also be expressed in terms of a "transformation" or mapping: for example, for the map $\alpha(r,\theta) = (r\cos\theta, r\sin\theta)$ of a portion of $M = \mathbb{R}^2$ to itself, the Lagrangian $L \circ \alpha_*\colon TM \to TM$ may lead to easier equations than the original Lagrangian L.

Similarly, we might hope that a diffeomorphism $f\colon T^*M \to T^*M$ [not necessarily of the form g^* for any g] can change Hamilton's equations for a given Hamiltonian H into a set of equations for a simpler Hamiltonian K. So we first want to know which transformations always take equations in Hamiltonian form into new equations that are also in Hamiltonian form. Actually, most classical investigations implicitly asked when equations for H become equations specifically for $H \circ f^{-1}$ (for the general formulation, see the bottom of page 571).

We work locally, and think of each cotangent space $M_p{}^*$ simply as $\mathbb{R}^n$. To determine which transformations have the desired property, it is natural to introduce the $2n \times 2n$ matrix

$$\mathbb{J} = \mathbb{J}_n = \left(\begin{array}{c|c} 0 & I_n \\ \hline -I_n & 0 \end{array}\right) \qquad \text{where } I_n \text{ denotes the } n \times n \text{ identity matrix,}$$

since the Hamiltonian equations for a curve c in T^*M,

$$\begin{aligned} \dot{q}^i(c'(t)) &= \frac{\partial H}{\partial p_i}(c(t)) = H_{p_i}(c(t)) \\ \dot{p}_j(c'(t)) &= -\frac{\partial H}{\partial q^j}(c(t)) = -H_{q^j}(c(t)), \end{aligned}$$

can then be written as

(a) $$\dot{c}^{\mathbf{t}} = \mathbb{J}(DH)^{\mathbf{t}}$$

where

$$\begin{aligned} {}^{\mathbf{t}} \text{ denotes } & \text{ the transpose of a matrix,} \\ \dot{c} \text{ denotes } & ((q^1 \circ c)', \ldots, (q^n \circ c)', (p_1 \circ c)', \ldots, (p_n \circ c)'), \\ (DH) \text{ denotes } & (H_{q^1} \circ c, \ldots, H_{q^n} \circ c, H_{p_1} \circ c, \ldots, H_{p_n} \circ c). \end{aligned}$$

Since $\mathbb{J}^2 = -I_{2n}$, this can also be written as

(b) $$-\mathbb{J}\dot{c}^{\mathbf{t}} = (DH)^{\mathbf{t}}.$$

For a map $f: T^*M \to T^*M$, the curve $\gamma = f \circ c$ satisfies

(c) $$\dot{\gamma}^{\mathbf{t}} = (Df)\dot{c}^{\mathbf{t}},$$

where (Df) denotes the Jacobian matrix of f with respect to the coordinates $q^1, \ldots, q^n, p_1, \ldots, p_n$. And for nonsingular Df, the map $K = H \circ f^{-1}$ satisfies

(d) $$\begin{aligned} (DK)^{\mathbf{t}} &= (Df^{-1})^{\mathbf{t}}(DH)^{\mathbf{t}} \\ &= -(Df^{-1})^{\mathbf{t}}\mathbb{J}\dot{c}^{\mathbf{t}} && \text{by (b)} \\ &= -(Df^{-1})^{\mathbf{t}}\mathbb{J}(Df^{-1})\dot{\gamma}^{\mathbf{t}} && \text{by (c).} \end{aligned}$$

Since $\mathbb{J}^2 = -I$ implies that $\mathbb{J}^{-1} = -\mathbb{J}$, this can be written as

(a′) $$\begin{aligned} \dot{\gamma}^{\mathbf{t}} &= -(Df)^{\mathbf{t}}\mathbb{J}^{-1}(Df)(DK)^{\mathbf{t}} \\ &= (Df)^{\mathbf{t}}\mathbb{J}(Df)(DK)^{\mathbf{t}}, \end{aligned}$$

and the equations (a′) will have the same form as (a) whenever (Df) satisfies

(A) $$(Df)^{\mathbf{t}}\mathbb{J}(Df) = \mathbb{J}.$$

To interpret (A), recall that *congruence* of $m \times m$ matrices, $B = C^{\mathbf{t}}AC$, is usually introduced in connection with the quadratic form associated with a symmetric matrix A. It may equally be applied to a skew-symmetric $A = (a_{ij})$, to which we associate the skew-symmetric map $\mathcal{A}: \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ defined by

$\mathcal{A}(\mathbf{e}_i, \mathbf{e}_j) = a_{ij}$, which can also be written as

$$\mathcal{A} = \tfrac{1}{2} \sum_{i,j=1}^{m} a_{ij} \cdot \mathbf{e}_i{}^* \wedge \mathbf{e}_j{}^*.$$

Remark. Note, for later use, that if $\mathbf{v}$ and $\mathbf{w}$ are vectors ($1 \times n$ matrices), then $\mathcal{A}(\mathbf{w}, \mathbf{v}) =$ (the single entry of) $\mathbf{w}A\mathbf{v}^{\mathbf{t}}$.

As in the symmetric case, for a linear $T\colon \mathbb{R}^m \to \mathbb{R}^m$ with $T(\mathbf{e}_i) = \sum_{k=1}^{n} c_{ki}\mathbf{e}_k$ we have

$$\mathcal{A}(T(\mathbf{e}_i), T(\mathbf{e}_j)) = \sum_{k,l=1}^{n} c_{ki}c_{lj}\mathcal{A}(\mathbf{e}_k, \mathbf{e}_l) = \sum_{k,l=1}^{n} c_{ki}c_{lj}a_{kl},$$

which is the (i, j) entry of the matrix $C^{\mathbf{t}}AC$, so that $C^{\mathbf{t}}AC$ *is the matrix that corresponds to* $\mathcal{A}$ *in the basis* $\{T(\mathbf{e}_i)\}$. Writing the standard basis of $\mathbb{R}^{2n}$ as $\mathbf{e}_1, \dots, \mathbf{e}_n, \bar{\mathbf{e}}_1, \dots, \bar{\mathbf{e}}_n$, the map corresponding to $\mathbb{J}$ is

$$\tfrac{1}{2}\sum_{i=1}^{n} \mathbf{e}_i{}^* \wedge \bar{\mathbf{e}}_i{}^*.$$

Translating this in terms of the dq^i and dp_j on $M_{p}{}^*$, equation (A) thus tells us that $\sum_i dq^i \wedge dp_i = \sum_i f^*(dq^i) \wedge f^*(dp_i)$, so that

$$\text{(A}'\text{)} \qquad f^*\omega = \omega,$$

which is precisely the modern definition of a "canonical transformation".

DEFINITION. A diffeomorphism $f\colon T^*M \to T^*M$ is called a **canonical transformation** if the map $f^*\colon T^*(T^*M) \to T^*(T^*M)$ satisfies

$$f^*\omega = \omega.$$

Note that f^{-1} is also a canonical transformation. Since f^* also preserves the n-fold product $\omega \wedge \cdots \wedge \omega$ considered on page 512, f is orientation preserving, and in terms of coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$ around $f(p)$, which we will often denote by (q, p) for brevity, and $(q \circ f, p \circ f)$ around p, we have $\det f^* = 1$ [equation (A) on the previous page already implies that $(\det f^*)^2 = 1$].

Having reached our elegant definition, we adopt the standard mathematical ploy of ditching the motivating considerations, and replacing them with a theorem. Recall, from page 533, that the solutions to Hamilton's canonical equations for the Hamiltonian H are the same as the solutions of the vector field $\mathbf{X}_H$ defined by

$$\mathbf{X}_H \lrcorner\, \omega = -dH.$$

Such vector fields are called **Hamiltonian vector fields**.

LEMMA. Let $f: M \to N$ be a diffeomorphism, X a vector field on M, and λ a k-form on M. Then

$$(f^{-1})^*(X \lrcorner \lambda) = f_*X \lrcorner (f^{-1})^*\lambda.$$

PROOF. A straightforward unraveling of definitions. ❖

1. THEOREM. The diffeomorphism $f: T^*M \to T^*M$ is canonical if and only if for all H

$$f_*(\mathbf{X}_H) = \mathbf{X}_{H \circ f^{-1}} = \mathbf{X}_{f^*H}.$$

In particular, if f is canonical, then f_* always takes Hamiltonian vector fields into Hamiltonian vector fields.

PROOF. The definition $\mathbf{X}_H \lrcorner \boldsymbol{\omega} = -dH$ and the Lemma give, respectively,

$$\begin{aligned}(f^{-1})^*(\mathbf{X}_H \lrcorner \boldsymbol{\omega}) &= -(f^{-1})^*(dH) = -d(H \circ f^{-1})\\ (f^{-1})^*(\mathbf{X}_H \lrcorner \boldsymbol{\omega}) &= f_*(\mathbf{X}_H) \lrcorner (f^{-1})^*\boldsymbol{\omega},\end{aligned}$$

so we have

$$-d(H \circ f^{-1}) = f_*(\mathbf{X}_H) \lrcorner (f^{-1})^*\boldsymbol{\omega}.$$

If f is canonical, then f^{-1} is also canonical, $(f^{-1})^*\boldsymbol{\omega} = \boldsymbol{\omega}$, so we have

$$-d(H \circ f^{-1}) = f_*(\mathbf{X}_H) \lrcorner \boldsymbol{\omega} \implies f_*(X_H) = X_{H \circ f^{-1}}.$$

Conversely, if $f_*(X_H) = X_{H \circ f^{-1}}$ for all H, then we have $-d(H \circ f^{-1}) = \mathbf{X}_{H \circ f^{-1}} \lrcorner (f^{-1})^*\boldsymbol{\omega}$ and thus $-dK = \mathbf{X}_K \lrcorner (f^{-1})^*\boldsymbol{\omega}$ for all K, while also $-dK = \mathbf{X}_K \lrcorner \boldsymbol{\omega}$ for all K, implying that $(f^{-1})^*\boldsymbol{\omega} = \boldsymbol{\omega}$. ❖

At this point, it might be nice to have a few concrete examples of canonical transformations $f: T^*M \to T^*M$. Canonical transformations are often defined in terms of the corresponding change of variables: given coordinates (q, p) on the domain, we give a formula for the corresponding coordinates

$$\begin{aligned} Q^i &= q^i \circ f \\ P_j &= p_j \circ f \end{aligned} \qquad \text{with} \qquad \begin{aligned} dQ^i &= f^*(dq^i) \\ dP_j &= f^*(dp_j), \end{aligned}$$

on the range. For example, in dimension 1, we can consider

$$Q = q + p\tau + \tfrac{1}{2}g\tau^2, \qquad P = p + g\tau$$

for constants g and τ. Since

$$dQ \wedge dP = f^*(dq + \tau\, dp) \wedge f^*dp = f^*(dq \wedge dp),$$

this is a canonical transformation. Often, we don't even bother writing the f^*.

For example, if we define, for constants ω and τ,

$$Q = q\cos\omega\tau + \tfrac{1}{\omega}p\sin\omega\tau, \qquad P = -\omega q\sin\omega\tau + p\cos\omega\tau,$$

then

$$\begin{aligned} dQ \wedge dP &= \left(\cos\omega\tau\, dq + \tfrac{1}{\omega}\sin\omega\tau\, dp\right) \wedge (-\omega\sin\omega\tau\, dq + \cos\omega\tau\, dp) \\ &= \cos^2\omega\tau\, dq \wedge dp + \sin^2\omega\tau\, dq \wedge dp = dq \wedge dp. \end{aligned}$$

These examples might call to mind the equations for an object traveling in a parabolic arc and for the harmonic oscillator, leading you to anticipate a result from a later section. At the moment, however, the most important point about these examples is that the q^i and the p_j are simply treated as independent coordinates, as we noted on page 534. For Lagrange's equations, we can choose the transformation $q^i \mapsto Q^i$ arbitrarily, but the total transformation $(q^i, \dot{q}^i) \mapsto (Q^i, \dot{Q}^i)$ from TM to TM is then completed determined if we want to end up with second order equations. But in the case of Hamilton's equations, we can allow transformations from T^*M to T^*M that mix the q^i and p_j, though now we want to consider only those transformations that are canonical. The equal footing of the q^i and p_j is strikingly illustrated by the canonical transformation

$$Q^k = p_k, \quad P_k = -q^k$$

for any k, which simply interchanges q^k and p_k, with a sign change. In general, any coordinates α^i, β_j for which we have

$$\boldsymbol{\omega} = \sum_{i=1}^{n} d\beta_i \wedge d\alpha^i$$

are called **canonical coordinates**, with the β_i **canonically conjugate** to the α^i.

For one more example of a canonical transformation, which will appear in Problem 3, we instead give formulas for $q = Q \circ f^{-1}$ and $p = P \circ f^{-1}$:

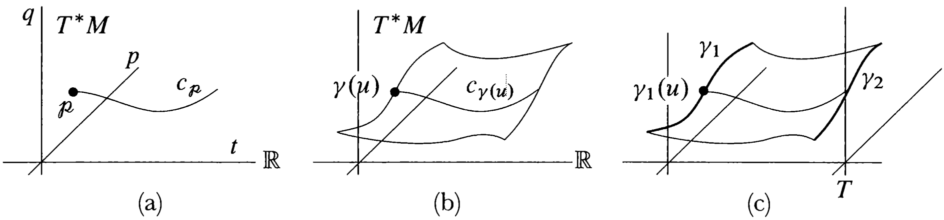
$$q = \sqrt{\frac{2P}{\omega}}\,\sin Q, \qquad p = \sqrt{2\omega P}\,\cos Q.$$

Simple calculations again prove that they define a canonical transformation.

As already pointed out, Theorem 1 actually only identifies transformations taking equations for the Hamiltonian H into equations for the Hamiltonian $H \circ f^{-1}$; the possibility that such equations are transformed into equations for other Hamiltonians K, depending on H, is seldom addressed. But, in fact, if f is a "generalized canonical transformation", $f^*\boldsymbol{\omega} = a\boldsymbol{\omega}$ for a constant a, then this is true for $K = (1/a)H \circ f^{-1}$, which can easily be seen from both our matrix proof and the proof of Theorem 1, and simply amounts to reparameterizing curves $c(t)$ as $c(at)$. The not-very-informative proof that these are the only other possibilities is given in Addendum B.

Hamiltonian flows and integral invariants. In contrast to the previous section, where we studied the relationship between an individual transformation $f\colon T^*M \to T^*M$ and all Hamiltonian vector fields, we will now consider a specific Hamiltonian vector field X, and look at the 1-parameter group of diffeomorphisms $\phi_t\colon T^*M \to T^*M$ generated by X, as on page 538.

If we have a point $\mathscr{p} \in T^*M$, we can consider (a) the curve $c_{\mathscr{p}}$ in $T^*M \times \mathbb{R}$ given by $t \mapsto (\phi_t(\mathscr{p}), t)$, showing the action of the ϕ_t on $\mathscr{p}$. And if we have a



curve $\gamma\colon [0,1]$ in T^*M, given by $u \mapsto \gamma(u)$, then (b) we can join all the images of the $c_{\gamma(u)}$ to obtain a surface, which is parameterized by the map

$$A(u,t) = (\phi_t(\gamma(u)), t).$$

In particular (c), if we start with a curve γ_1 in $T^*M = T^*M \times \{0\}$, and follow all the c curves to time T, we end up with a curve of the form $\{(\gamma_2(u), T)\}$ for some curve γ_2 in T^*M.

This picture is the basic set-up for the "integral invariants of Poincaré", and everything that we need to know about these invariants, with proofs complete, will eventually be given in a few lines. But, emulating the presentation of the previous section, we will slowly sneak up on the slick method. To begin, we will reach back in time to the venerable classic Whittaker [1] for a version that is, in its own way, pretty slick.

First approach. Suppose our Hamiltonian H comes from a Lagrangian L. Then the various curves c are just the image under $\mathbf{F}DL$ of curves $\mathscr{c}$ in TM satisfying the corresponding Lagrange equations, and A is just the image under $\mathbf{F}DL$ of a variation $\mathscr{A}$ of the curves $\mathscr{c}$. Since all the curves $\mathscr{c}$ in the variation are extremals for L, we can apply the Boundary Term Corollary on page 462 at each u, to obtain

$$\text{(A)} \qquad \frac{dJ(\bar{\mathscr{A}}(u))}{du} = \sum_{i=1}^{n} \frac{\partial q^i(\mathscr{A}(u,t))}{\partial u} \cdot \frac{\partial L(\bar{\mathscr{A}}(u)'(t),t)}{\partial \dot{q}^i}\Bigg|_{t=0}^{t=T},$$

where $J(\bar{\mathscr{A}}(u))$ is the integral of L over $\bar{\mathscr{A}}(u) = t \mapsto \mathscr{A}(u,t)$ on $[0,T]$.

I_1. THEOREM. If γ_1 is a smooth closed curve, then

$$\int_{\gamma_1} \sum_{i=1}^{n} p_i \, dq^i = \int_{\gamma_2} \sum_{i=1}^{n} p_i \, dq^i .$$

PROOF. We integrate (A) from 0 to 1. On the left we get 0, since we are integrating a derivative along a closed curve. Remembering that $\partial L/\partial \dot{q}^i = p_i$, we see that on the right the terms for $t = 0$ and $t = T$ give, respectively,

$$\int_{\gamma_1} \sum_{i=1}^{n} p \, dq^i \qquad \text{and} \qquad \int_{\gamma_2} \sum_{i=1}^{n} p \, dq^i . \;❖$$

The form $\boldsymbol{\theta} = \sum_{i=1}^{n} p_i \, dq^i$ is called a 1-dimensional *relative integral invariant of Poincaré.* The "relative" refers to the fact that its integral over a curve is an invariant under the flow of the Hamiltonian only when the curve is closed. On the other hand, if we consider a disc with boundary curve γ_1, or more generally, a 2-chain D_1 with $\partial D_1 = \gamma_1$, and the corresponding D_2 with $\partial D_2 = \gamma_2$, then each

$$\int_{D_i} \boldsymbol{\omega} = \int_{D_i} d\boldsymbol{\theta} = \int_{\partial D_i} \boldsymbol{\theta} = \int_{\gamma_i} \boldsymbol{\theta},$$

so that $\displaystyle\int_{D_1} \boldsymbol{\omega} = \int_{D_2} \boldsymbol{\omega}$. Hence $\boldsymbol{\omega}$ is a 2-dimensional *integral invariant of Poincaré.*

In connection with Theorem I_1, Whittaker, like many another author, points out: "This is essentially the same as the hydrodynamical theorem that the circulation in any circuit moving with the fluid does not alter with the time." Arnold [2; Chap. 9] provides a set of variations on this theme, and Kozlov [1] orchestrates a full-scale symphony.

Second approach. Although this use of the Boundary Term Corollary is intriguing, there is a more natural proof that will also give a more general result.

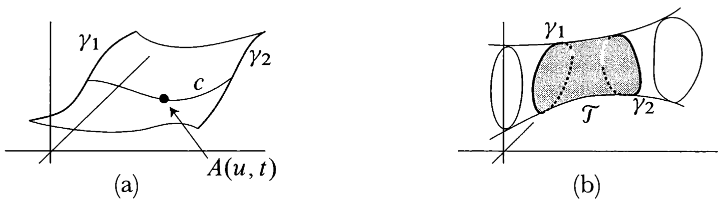
Recall that we used the equations at the bottom of page 531 to obtain the Hamilton–Jacobi equation. In condensed notation, these equations might be written

$$dS = \sum_{i=1}^{n} p_i \, dq^i - H \, dt, \tag{B}$$

which looks like an analogue of the equation on page 526

$$dV = \sum_{i=1}^{3} p_i \, dx^i ,$$

except that (B) is meaningless as stated, since dS is a 1-form on $M \times \mathbb{R}$ while the right side is a 1-form on $T^*M \times \mathbb{R}$. However, given a variation A as on page 572, we can define S on the image surface (a) by letting $S(A(u,t))$ be the integral of L on $[0,t]$ along the curve $\mathcal{c}$ for which the corresponding c contains $A(u,t)$,



and then (B) holds on the image surface. We can then apply (B) to the "tube" $\mathcal{T}$ that we get (b) by finding solutions of Hamilton's equations starting from a closed curve, and cutting it off along curves γ_1 and γ_2.

I_2. THEOREM. For any closed curves γ_1 and γ_2 surrounding a tube made up of solutions to Hamilton's equations, we have

$$\int_{\gamma_1} \sum_{i=1}^{n} p_i\, dq^i - H\, dt = \int_{\gamma_2} \sum_{i=1}^{n} p_i\, dq^i - H\, dt.$$

PROOF. Letting $\mathcal{H} = \sum_{i=1}^{n} p_i\, dq^i - H\, dt = dS$, we have

$$\int_{\gamma_2} \mathcal{H} - \int_{\gamma_1} \mathcal{H} = \int_{\gamma_2 - \gamma_1} \mathcal{H} = \int_{\partial\mathcal{T}} \mathcal{H} = \int_{\mathcal{T}} d\mathcal{H} = \int_{\mathcal{T}} d(dS) = 0. \quad ❖$$

Theorem I_1 is a special case, since we then have t constant on γ_1 and γ_2, so that the $H\,dt$ term is 0. The 1-form $\sum_{i=1}^{n} p_i\, dq^i - H\, dt$ is called the *integral invariant of Poincaré–Cartan* (without bothering to add the "relative").

Third approach. Although our proof of Theorem I_2 shows how the result arises from the relationship between Hamilton's equations and S, we would like to have a proof that uses Hamilton's equations directly, in line with our program of focusing more attention on the Hamiltonian equations themselves. (Moreover, our current proof only works when H arises from a regular Lagrangian, not for general H.)

PROOF 2 OF THEOREM I_2. We have

$$\int_{\gamma_1 - \gamma_2} \sum_{i=1}^{n} p_i\, dq^i - H\, dt = \int_{\partial\mathcal{T}} \sum_{i=1}^{n} p_i\, dq^i - H\, dt = \int_{\mathcal{T}} d\Big(\sum_{i=1}^{n} p_i\, dq^i - H\, dt\Big)$$

$$= \int_{\mathcal{T}} \sum_{i=1}^{n} dp_i \wedge dq^i - \frac{\partial H}{\partial p_i} dp_i \wedge dt - \frac{\partial H}{\partial q^i} dq^i \wedge dt = \int_{\mathcal{T}} \mathcal{A}, \quad \text{say.}$$

It will be convenient to regard $T^*M \times \mathbb{R}$ as $\mathbb{R}^{2n+1}$. At any $\mathscr{p} \in M$, consider the $(2n+1)\times(2n+1)$ skew-symmetric matrix $A = (a_{ij})$ corresponding to $\mathcal{A}(\mathscr{p})$, as on pages 568–569. From the form of $\mathcal{A}$, we see that the matrix A is

$$A = \begin{array}{cc} \begin{array}{cc} \quad 2n \quad & \quad 1 \quad \end{array} & \\ \left(\begin{array}{c|c} \mathbb{J} & DH^{\mathbf{t}} \\ \hline -DH & 0 \end{array}\right) & \begin{array}{c} 2n \\ 1 \end{array} \end{array}$$

The Remark on page 569 shows that for tangent vectors $\mathbf{v}$ and $\mathbf{w}$ at $\mathscr{p}$ we have

$$(*) \qquad \mathcal{A}(\mathscr{p})(\mathbf{w}, \mathbf{v}) = \text{(the single entry of) } \mathbf{w}A\mathbf{v}^{\mathbf{t}}.$$

Now the vector

$$\mathbf{v} = (H_{p_1}, \dots, H_{p_n}, -H_{q^1}, \dots, -H_{q^n}, 1)$$

is easily seen to be a tangent vector along the flow of the vector field X corresponding to the Hamiltonian H, so at each $\mathscr{p} \in \mathcal{T}$ it lies in the tangent space of $\mathcal{T}$. But we easily compute that

$$A\mathbf{v}^{\mathbf{t}} = 0,$$

so that by $(*)$ we have

$$\mathcal{A}(\mathscr{p})(\mathbf{w}, \mathbf{v}) = 0 \qquad \text{for all } \mathbf{w} \text{ at } \mathscr{p}.$$

Hence $\mathcal{A}(\mathscr{p})(\mathbf{w}, \mathbf{v}) = 0$ for linearly independent $\mathbf{w}$ and $\mathbf{v}$. On the 2-dimensional tangent space $\mathcal{T}_{\mathscr{p}}$ we therefore have $\mathcal{A}(\mathscr{p}) = 0$, for each $\mathscr{p}$ in $\mathcal{T}$. ❖

We have now shown in three different ways that on T^*M the 1-form $\boldsymbol{\theta} = \sum_{i=1}^n p_i\, dq^i$ is a relative integral invariant. Hence, $\boldsymbol{\omega}$ is an integral invariant: its integral over any disc is the same as its integral over the image of that disc under any ϕ_t of the flow generated by a Hamiltonian vector field. An approximation argument could then be used to prove that $\boldsymbol{\omega}$ itself must be preserved, so that

THEOREM. Each ϕ_t is a canonical transformation.

In short, the flow of any Hamiltonian vector field gives a 1-parameter family of canonical transformations.

Thus, at the end of all these clever manipulations with integral invariants, a bit more work would give a result that doesn't mention them at all. Of course, it would certainly be much more convenient if we could go in the opposite direction: Knowing that each ϕ_t is canonical would tell us immediately, without approximation arguments, that $\boldsymbol{\omega}$ is an integral invariant, and thus also that $\boldsymbol{\theta}$ is a relative integral invariant.

With a wave of the wand, here comes the proof.

Hamiltonian flows and canonical transformations. The claim that all ϕ_t for a Hamiltonian vector field $\mathbf{X}_H$ are canonical transformations amounts to saying that the Lie derivative $L_{\mathbf{X}_H}$ satisfies $L_{\mathbf{X}_H}\omega = 0$. We will want to use the formulas

$$d(L_X\lambda) = L_X(d\lambda),$$
$$L_X(\lambda \wedge \mu) = L_X\lambda \wedge \mu + \lambda \wedge L_X\mu,$$

and the Cartan formula, or as it is dubbed in Marsden and Ratiu [1],

$$L_X\lambda = d(X \lrcorner \lambda) + X \lrcorner d\lambda \qquad \text{Cartan's Magic Formula,}$$

and also take this opportunity to introduce, for later use, the less familiar

$$[X, Y] \lrcorner \lambda = L_X(Y \lrcorner \lambda) - Y \lrcorner (L_X\lambda) \qquad \text{Cartan's Bracket Formula,}$$

all of which are reviewed in Problem 1.

2. THEOREM. The flow of any Hamiltonian vector field consists of canonical transformations.

PROOF 1 (Hogwarts version). For $\mathbf{X}_H$ satisfying $\mathbf{X}_H \lrcorner \omega = -dH$ we have

$$\begin{aligned} L_{\mathbf{X}_H}\omega &= d(\mathbf{X}_H \lrcorner \omega) + \mathbf{X}_H \lrcorner d\omega \\ &= d(-dH) + \mathbf{X}_H \lrcorner 0 = 0. \ ❖ \end{aligned}$$

PROOF 2 (Muggles version). We have

$$\begin{aligned} L_{\mathbf{X}_H}\left(\sum_{i=1}^n dp_i \wedge dq^i\right) &= \sum_{i=1}^n L_{\mathbf{X}_H} dp_i \wedge dq^i + \sum_{i=1}^n dp_i \wedge L_{\mathbf{X}_H} dq^i \\ &= \sum_{i=1}^n d(\mathbf{X}_H(p_i)) \wedge dq^i + dp_1 \wedge d(\mathbf{X}_H(q^i)) \\ &= \sum_{i=1}^n d\left(-\frac{\partial H}{\partial q^i}\right) \wedge dq^i + \sum_{i=1}^n dp_i \wedge d\left(\frac{\partial H}{\partial p_i}\right). \end{aligned}$$

The first sum is

$$\sum_{i,j=1}^n \left(-\frac{\partial^2 H}{\partial q^j \partial q^i} dq^j \wedge dq^i - \frac{\partial^2 H}{\partial p_j \partial q^i} dp_j \wedge dq^i\right)$$

$$= \qquad 0 \qquad - \sum_{i,j=1}^n \frac{\partial^2 H}{\partial p_j \partial q^i} dp_j \wedge dq^i,$$

and we find similarly that the second sum is the negative of this. ❖

Note: Theorem 2 immediately implies Liouville's volume theorem.

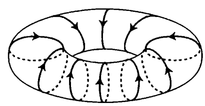
3. THEOREM. (Converse of Theorem 2). If M is simply-connected (or more generally, if $H^1(M;\mathbb{R}) = 0$), and all ϕ_t of the flow of X are canonical, then X is a Hamiltonian vector field.

PROOF. If all ϕ_t are canonical, then $L_X\omega = 0$, so

$$0 = L_X\omega = X \lrcorner\, d\omega + d(X \lrcorner\, \omega) = d(X \lrcorner\, \omega).$$

Thus $X \lrcorner\, \omega$ is closed, and the hypothesis on M implies that it is therefore exact, so that we have $X \lrcorner\, \omega = -dH$ for some function H. ❖

Without the extra hypothesis on M, we still have that X is **locally Hamiltonian** (i.e., Hamiltonian in a neighborhood of each point). For a simple example of a locally Hamiltonian vector field X that is not globally Hamiltonian, consider the vector field on a torus with integral curves shown below. As we follow any of the



integral curves, H would have to increase, so it couldn't be well-defined on the closed integral curve. Or, simply note that H would have to have a maximum point on the torus, and there X would have to be 0; this argument works even when the integral curves wind around the torus forever, at an irrational slope.

Generating functions. To specify a canonical transformation $f\colon T^*M \to T^*M$, we would seem to need $2n$ functions of $2n$ variables. On the other hand, the fact that f is canonical gives $2n$ relations between these functions, so we might expect to need only one.

In fact, if we have a canonical $f\colon T^*M \to T^*M$ and a canonical coordinate system (q,p), then for the functions $Q = q \circ f$ and $P = p \circ f$ we have

$$\sum_{i=1}^{n} dp_i \wedge dq^i = \sum_{i=1}^{n} dP_i \wedge dQ^i \implies d\Bigl(\sum_{i=1}^{n} p_i\, dq^i - \sum_{i=1}^{n} P_i\, dQ^i\Bigr) = 0,$$

and hence locally there is a function $\mathcal{S}\colon T^*M \to \mathbb{R}$ with

$$\text{(A)} \qquad \sum_{i=1}^{n} p_i\, dq^i - \sum_{i=1}^{n} P_i\, dQ^i = d\mathcal{S}.$$

Sometimes $\mathcal{S}$, which is only determined up to a constant, is called a generating function of f, but the classical notion of a generating function treats $\mathcal{S}$ in a particularly clever way, in order to disentangle the (q,p) and (Q,P) coordinates. Our considerations are local, so we regard f as a map $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ (shrinking down the region that $\mathbb{R}^n \times \mathbb{R}^n$ represents as needed for any additional assumptions to hold).

DANGEROUS CURVE

Suppose that for our canonical transformation f we have the following condition on the Jacobian matrix,

$$\text{(J}_1\text{)}\qquad 0 \neq \det \frac{\partial Q(q,p)}{\partial p} = \det \frac{\partial(q, Q(q,p))}{\partial(q,p)},$$

so that knowing q for a point in T^*M and $Q(q,p)$ for its image under f determines p uniquely [as trivial examples, in dimension 1 this is true for $f(q,p) = (q, p+q)$, but not for the identity function $f(q,p) = (q,p)$]. This condition implies that the map

$$\chi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n \quad \text{defined by} \quad \chi(q,p) = (q, Q(q,p))$$

is a diffeomorphism, and thus (q, Q) is a coordinate system on $\mathbb{R}^n \times \mathbb{R}^n$ (no one is saying that it is canonical!).

When (J_1) holds, so that (q, Q) is a coordinate system, equation (A) on page 577 can simply be summed up by the equations

$$(1)\qquad \frac{\partial \mathcal{S}}{\partial q^i} = p_i, \quad \frac{\partial \mathcal{S}}{\partial Q^i} = -P_i \quad \textit{in the } (q, Q) \textit{ coordinate system on } \mathbb{R}^n \times \mathbb{R}^n.$$

In order to relate these equations to the standard (q,p) coordinate system, we define the "type 1 generating function" $S_1\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$(\text{F}_1)\qquad S_1(q, Q(q,p)) = \mathcal{S}(q,p), \quad \text{or, more precisely,} \quad S_1 = \mathcal{S} \circ \chi^{-1}.$$

Then the formal definition of $\partial/\partial q^i$ and $\partial/\partial Q^i$ translates equations (1) into

$$D_i S_1(q, Q(q,p)) = p_i, \qquad D_{n+i} S_1(q, Q(q,p)) = -P_i,$$

where D_j denotes the partial derivative with respect to the j^{th} variable. This can also be written in conventional notation as

$$(\text{F}_1\text{-a})\qquad \frac{\partial S_1}{\partial q^i}(q, Q(q,p)) = p_i \qquad \text{or briefly} \qquad \frac{\partial S_1}{\partial q^i}(q, Q) = p_i$$

$$(\text{F}_1\text{-b})\qquad \frac{\partial S_1}{\partial Q^i}(q, Q(q,p)) = -P_i \quad \text{or briefly} \qquad \frac{\partial S_1}{\partial Q^i}(q, Q) = -P_i$$

where $\partial/\partial q^i$ ***and*** $\partial/\partial Q^i$ ***are now just being used as convenient abbreviations for*** D_i ***and*** D_{n+i}. It will be clear when they are being used this way because the argument will be $(q, Q(q,p))$, or more elliptically (q, Q).

In some contexts we might simply denote $\partial/\partial Q^i$ by $\partial/\partial p_i$, as when we observe that $(\text{F}_1\text{-a})$ implies

$$(2)\qquad \delta_i^j = \frac{\partial}{\partial p_j}\left(\frac{\partial S_1}{\partial q^i}(q, Q(q,p))\right) = \sum_{k=1}^n \frac{\partial^2 S_1}{\partial q^i \partial p_k}(q, Q(q,p)) \frac{\partial Q^k}{\partial p_j}(q,p),$$

so $\left(\partial^2 S_1/\partial q \partial p\right)(q, Q(q,p))$ and $\left(\partial Q/\partial p\right)(q,p)$ are inverses of each other, and in particular $\left(\partial^2 S_1/\partial q \partial p\right)$ is non-singular at the points of interest to us.

The name "generating function" refers to the fact that we can go in the other direction, and use a function $S : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying

$$\det \frac{\partial^2 S}{\partial q \partial p} \neq 0$$

to define, or generate, a canonical transformation. In fact, this condition implies, by the implicit function theorem, that there are functions $Q^i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying the equivalent of (F_1-a), namely

$$\frac{\partial S}{\partial q^i}(q, Q(q,p)) = p_i. \tag{3}$$

The calculation (2), now applied to (3) instead of (F_1-a), shows that we have

$$\det \frac{\partial Q(q,p)}{\partial p} \neq 0. \tag{4}$$

This means, once again, that we can also use (q, Q) as a coordinate system, and then (3) can be written as

$$\text{(a)} \qquad \frac{\partial S}{\partial q^i} = p_i \qquad \textit{in the } (q, Q) \textit{ coordinate system.}$$

We now define P_i by the analogue of (F_1-b),

$$P_i = -\frac{\partial S}{\partial p_i}(q, Q(q,p)),$$

so that we also have

$$\text{(b)} \qquad \frac{\partial S}{\partial Q^i} = -P_i \qquad \textit{in the } (q, Q) \textit{ coordinate system.}$$

Using (a) and (b) we now compute, *in the* (q, Q) *coordinate system*, that

$$\begin{aligned} \sum_{i=1}^n p_i\, dq^i - P_i\, dQ^i &= \sum_{i=1}^n \left(\frac{\partial S}{\partial q^i}\, dq^i + \frac{\partial S}{\partial Q^i}\, dQ^i \right) \\ &= dS, \end{aligned}$$

and taking d of this equation shows that the function $f = (Q, P)$ is canonical.

In addition, the argument on the previous page showed that (4) $\equiv$ (J_1) implies that f has a generating function S_1 satisfying equations (F_1-a) and (F_1-b), which are the same as (a) and (b) for S, so that $S = S_1 + \text{constant}$, and we can conclude that our given S is in fact a type 1 generating function for the transformation $f = (Q, P)$ that we have just defined.

Though condition (J1) can easily fail to be true, we might have instead

$$\text{(J2)} \qquad 0 \neq \det \frac{\partial P(q,p)}{\partial p} = \det \frac{\partial(q, P(q,p))}{\partial(q,p)}$$

(which does hold for the identity map). In this case, q and $P(q,p)$ determine p, so (q, P) is a coordinate system, and we have the diffeomorphism

$$\psi(q,p) = (q, P(q,p)).$$

In order to find the partial derivatives $\partial \mathcal{S}/\partial q^i$ and $\partial \mathcal{S}/\partial P_i$, we use the relation $d(P_i Q^i) = P_i\, dQ^i + Q^i\, dP_i$ to write equation (A) on page 577 as

$$d\left(\mathcal{S} + \sum_{i=1}^{n} P_i Q^i\right) = \sum_{i=1}^{n} p_i\, dq^i + \sum_{i=1}^{n} Q^i\, dP_i.$$

If we write the left side as $d\bar{\mathcal{S}}$, then we have

$$\frac{\partial \bar{\mathcal{S}}}{\partial q^i} = p_i, \qquad \frac{\partial \bar{\mathcal{S}}}{\partial P^i} = Q^i,$$

and defining the "type 2 generating function" S_2 by

$$\text{(F2)} \quad S_2(q, P(q,p)) = \mathcal{S}(q,p) + \sum_{i=1}^{n} P_i Q^i(q,p), \quad \text{that is,} \quad S_2 = \bar{\mathcal{S}} \circ \psi^{-1},$$

we end up with

$$\text{(F2-a)} \qquad \frac{\partial S_2}{\partial q^i}(q, P(q,p)) = p_i$$

$$\text{(F2-b)} \qquad \frac{\partial S_2}{\partial P_i}(q, P(q,p)) = Q^i.$$

From (F2-a) we also obtain the analogue of equation (2) on page 578,

$$(2') \quad \delta_i^j = \frac{\partial}{\partial p_j}\left(\frac{\partial S_2}{\partial q^i}(q, P(q,p))\right) = \sum_{k=1}^{n} \frac{\partial^2 S_2}{\partial q^i \partial p_k}(q, P(q,p)) \frac{\partial P_k}{\partial p_j}(q,p).$$

Just as with type 1 generating functions, the condition $\det\bigl(\partial^2 S/\partial q \partial p\bigr) \neq 0$ can also be used to generate a canonical transformation by means of a type 2 generating function: First, we choose functions $P_i\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfying the equivalent of (F2-a). We next note that applying the calculation (2′) to S shows that $\bigl(\partial P/\partial p\bigr)$ is nonsingular, so that we can use (q, P) as a coordinate system. And we then define the Q^i by means of equation (F2-b). We leave it to the reader to check that we now have $\sum_{i=1}^n p_i\, dq^i - P_i\, dQ^i = d\bigl(S - \sum_{i,j=1}^n P_i Q^i(q,p)\bigr)$.

Other types of generating functions need not be considered until later.

Time-dependent canonical transformations. Before examining the use of generating functions, we want to extend our results not only to Hamiltonians $H\colon T^*M \times \mathbb{R} \to \mathbb{R}$ that depend on time, but also to the more general situation where the canonical transformations themselves may depend on time. That is, we want to consider maps $g\colon T^*M \times \mathbb{R} \to T^*M \times \mathbb{R}$ of the form

$$g(\mathcal{p},t) = (f(\mathcal{p},t),t), \qquad \text{for } \mathcal{p} \in T^*M,$$

with each $\mathcal{p} \mapsto f(\mathcal{p},t)$ being canonical (e.g., the 1-parameter family of canonical transformations generated by a Hamiltonian vector field). The big difference in this more general case is that the new Hamiltonian will not simply be $H \circ g^{-1}$.

Among the different possible approaches to this question, we choose one that requires no additional abstractions, though this will not relieve us from having to contend with somewhat fussy notation. Given any $f\colon T^*M \times \mathbb{R} \to T^*M \times \mathbb{R}$, for each t we will let ${}^tf\colon T^*M \to T^*M$ be the function

$${}^tf(\mathcal{p}) = f(\mathcal{p},t),$$

so that we have a 1-parameter collection of maps from T^*M to T^*M and saying that $\mathcal{p} \mapsto f(\mathcal{p},t)$ is canonical is equivalent to saying that all ${}^tf$ are canonical.

Locally, g will be now given by a collection of $2n+1$ coordinate functions

$$Q^1(\mathcal{p},t), \dots, Q^n(\mathcal{p},t),\ P_1(\mathcal{p},t), \dots, P_n(\mathcal{p},t), \text{ and } t,$$

and for each fixed t we have corresponding functions ${}^tQ^1, \dots, {}^tQ^n, {}^tP_1, \dots, {}^tP_n$, which are the coordinate functions for ${}^tf$.

[To go along with this notation, we will allow q^i, p_j on T^*M to be confused with coordinate systems with the same names on $T^*M \times \mathbb{R}$, so that we can write

$$q^i(\mathcal{p},t) \text{ as well as } q^i(\mathcal{p}) \qquad \text{and} \qquad p_j(\mathcal{p},t) \text{ as well as } p_j(\mathcal{p});$$

then for all t, the functions ${}^tq^i, {}^tp_j$ are the same coordinates q^i, p_j on T^*M.]

The fact that all ${}^tf$ are canonical means that for each t

$$\sum_{i=1}^n d({}^tp_i) \wedge d({}^tq^i) = \sum_{i=1}^n d({}^tP_i) \wedge d({}^tQ^i)$$

so that

$$d\left(\sum_{i=1}^n {}^tp_i \wedge d({}^tq^i) - \sum_{i=1}^n {}^tP_i \wedge d({}^tQ^i)\right) = 0,$$

and thus locally

$$(1) \qquad \sum_{i=1}^n {}^tp_i\, d({}^tq^i) - \sum_{i=1}^n {}^tP_i\, d({}^tQ^i) = d\mathcal{s}_t$$

for some function $\mathcal{s}_t\colon T^*M \to \mathbb{R}$. We can put all the $\mathcal{s}_t$ together into a function

$$\mathcal{S}\colon T^*M \times \mathbb{R} \to \mathbb{R} \qquad \text{defined by } \mathcal{S}(\mathcal{p},t) = \mathcal{s}_t(\mathcal{p})$$

(so that, officially, $\mathcal{s}_t = {}^t\mathcal{S}$).

After all this notational kanoodling, we next want to note that if D_{n+1} denotes the partial derivative with respect to the $(n+1)^{\text{st}}$ variable, then we can write

$$dQ^i = d({}^tQ^i) + D_{n+1}(Q^i)\,dt,$$

where this equation means that when we write $dQ^i(p,t)$ in terms of $dq^i(p,t)$, $dp_j(p,t)$, and $dt(p,t)$, the coefficients in the dq^i, dp_j part come from $d({}^tQ^i)$, while the coefficient of dt is $D_{n+1}(Q^i)(p,t)$. By abuse of notation we will simply write

$$dQ^i = d({}^tQ^i) + \frac{\partial Q^i}{\partial t}\,dt.$$

It follows that

$$(2)\qquad \sum_{i=1}^n P_i\,dQ^i = \sum_{i=1}^n P_i\,d({}^tQ^i) + \sum_{i=1}^n P_i\frac{\partial Q^i}{\partial t}\,dt.$$

We likewise have

$$d\mathcal{S} = d({}^t\mathcal{S}) + \frac{\partial \mathcal{S}}{\partial t}\,dt = d\mathfrak{s}_t + \frac{\partial \mathcal{S}}{\partial t}\,dt,$$

so (1) can be written as

$$\sum_{i=1}^n p_i\,dq^i - \sum_{i=1}^n P_i\,d({}^tQ^i) = d\mathcal{S} - \frac{\partial \mathcal{S}}{\partial t}\,dt,$$

where we keep the ${}^tQ^i$ to remind us that $d({}^tQ^i)$ is computed separately for each fixed t. Substituting (2) into this equation we then obtain, at long last,

$$(\tilde{\mathrm{A}})\qquad \sum_{i=1}^n p_i\,dq^i \;-\; \sum_{i=1}^n P_i\,dQ^i = d\mathcal{S} - \left[\sum_{i=1}^n P_i\frac{\partial Q_i}{\partial t} \;+\; \frac{\partial \mathcal{S}}{\partial t}\right]dt.$$

4. THEOREM. Let $g\colon T^*M\times\mathbb{R}\to T^*M\times\mathbb{R}$ be a canonical transformation, and suppose that the curve $c:\mathbb{R}\to T^*M\times\mathbb{R}$ satisfies the canonical equations for a Hamiltonian $H\colon T^*M\times\mathbb{R}\to\mathbb{R}$. Then the curve $g\circ c$ satisfies the canonical equations for the Hamiltonian $K\colon T^*M\times\mathbb{R}\to\mathbb{R}$ given by

$$(\mathrm{B})\qquad K\circ g = H \;+\; \left[\sum_{i=1}^n P_i\frac{\partial Q^i}{\partial t} \;+\; \frac{\partial \mathcal{S}}{\partial t}\right].$$

PROOF. The "right" proof (or perhaps we should say the canonical proof), would be to generalize Theorem 1 appropriately to $T^*M \times \mathbb{R}$. But empathizing with physicists, we will resort to a much easier elementary approach.

Setting $\boldsymbol{\theta} = \sum_{i=1}^{n} p_i \, dq^i$, we find that if K is defined by (B), then equation ($\widetilde{\text{A}}$) can be written as

$$\boldsymbol{\theta} - g^*\boldsymbol{\theta} = d\mathcal{S} - (K \circ g - H)\, dt$$

and thus as

$$\boldsymbol{\theta} - H\, dt = g^*(\boldsymbol{\theta} - K\, dt) + d\mathcal{S}.$$

For a curve $c\colon [a,b] \to T^*M \times \mathbb{R}$, this shows that

$$\int_c \boldsymbol{\theta} - H\, dt \; - \int_{g\circ c} \boldsymbol{\theta} - K\, dt \; = \text{ the constant } \mathcal{S}(c(b), b) - \mathcal{S}(c(a), a).$$

So the stationary curves for $\boldsymbol{\theta} - K\, dt$ are the images under g of the stationary curves for $\boldsymbol{\theta} - H\, dt$. But the extended Hamilton's principle (page 534) says that the stationary curves for $\boldsymbol{\theta} - H\, dt$ are the curves satisfying Hamilton's equations for H, while those for $\boldsymbol{\theta} - K\, dt$ are similarly the curves satisfying Hamilton's equations for K. ❖

Our discussion of generating functions is easily adapted to the generalized case. Again working locally, we consider $g : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$. We now suppose that g has the property that knowing q, Q, and t determines p; the condition for this will be

$$(\widetilde{\text{J}}_1) \qquad 0 \neq \det \frac{\partial Q(q,p,t)}{\partial p} = \det \frac{\partial(q, Q(q,p,t), t)}{\partial(q,p,t)}.$$

We define the "type 1 generating function" $S_1 : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ by

$$(\widetilde{\text{F}}_1) \qquad S_1(q, Q(q,p,t), t) = \mathcal{S}(q,p,t),$$

and find that

$$(\widetilde{\text{F}}_1\text{-a}) \qquad \frac{\partial S_1}{\partial q^i}(q, Q(q,p,t), t) = p_i$$

$$(\widetilde{\text{F}}_1\text{-b}) \qquad \frac{\partial S_1}{\partial Q^i}(q, Q(q,p,t), t) = -P_i.$$

Applying ($\widetilde{\text{F}}_1$-b) to equation (B) on the previous page gives

$$K \circ g = H - \sum_{i=1}^{n} \frac{\partial S_1}{\partial Q^i} \frac{\partial Q^i}{\partial t} + \frac{\partial \mathcal{S}}{\partial t}.$$

But taking $\partial/\partial t$ of $(\tilde{\mathrm{F}}_1)$ gives

$$\frac{\partial \mathcal{S}}{\partial t} = \frac{\partial S_1}{\partial t} + \sum_{i=1}^{n} \frac{\partial S_1}{\partial Q^i}\frac{\partial Q^i}{\partial t},$$

and we thus obtain

$(\tilde{\mathrm{F}}_1\text{-c})$
$$\boxed{K \circ g = H + \frac{\partial S_1}{\partial t}.}$$

If instead of $(\tilde{\mathrm{J}}_1)$ we have

$(\tilde{\mathrm{J}}_2)$
$$0 \neq \det \frac{\partial P(q,p,t)}{\partial p} = \det \frac{\partial(q, P(q,p,t),t)}{\partial(q,p)},$$

and we define the "type 2 generating function" S_2 by

$(\tilde{\mathrm{F}}_2)$
$$S_2(q, P(q,p,t),t) = \mathcal{S}(q,p,t) + \sum_{i=1}^{n} P_i Q^i(q,p,t),$$

then as before we obtain

$(\tilde{\mathrm{F}}_2\text{-a})$
$$\frac{\partial S_2}{\partial q^i}(q, P(q,p,t),t) = p_i$$

$(\tilde{\mathrm{F}}_2\text{-b})$
$$\frac{\partial S_2}{\partial P_i}(q, P(q,p,t),t) = Q^i.$$

Moreover, taking $\partial/\partial t$ of $\tilde{\mathrm{F}}_2$ gives

$$\frac{\partial S_2}{\partial t} + \sum_{i=1}^{n} \frac{\partial S_2}{\partial P_i}\frac{\partial P_i}{\partial t} = \frac{\partial \mathcal{S}}{\partial t} + \frac{\partial}{\partial t}\sum_{i=1}^{n} P_i Q^i,$$

so

$$\begin{aligned}
K \circ g &= H + \sum_{i=1}^{n} P_i \frac{\partial Q^i}{\partial t} + \frac{\partial \mathcal{S}}{\partial t} \\
&= H + \sum_{i=1}^{n} P_i \frac{\partial Q^i}{\partial t} + \frac{\partial S_2}{\partial t} + \sum_{i=1}^{n} \frac{\partial S_2}{\partial P_i}\frac{\partial P_i}{\partial t} - \frac{\partial}{\partial t}\sum_{i=1}^{n} P_i Q^i \\
&= H + \frac{\partial S_2}{\partial t} + \sum_{i=1}^{n} \frac{\partial S_2}{\partial P_i}\frac{\partial P_i}{\partial t} - \sum_{i=1}^{n} \frac{\partial P_i}{\partial t} Q^i,
\end{aligned}$$

and $(\tilde{\mathrm{F}}_2\text{-b})$ then gives

$(\tilde{\mathrm{F}}_2\text{-c})$
$$\boxed{K \circ g = H + \frac{\partial S_2}{\partial t}.}$$

In both cases, the reverse direction discussed on page 579 also generalizes in a straightforward way.

Using generating functions to simplify Hamilton's equations. After all this to-do, we recall that we first considered canonical transformations as a way of transforming Hamilton's equations for the Hamiltonian H into a new set having a Hamiltonian K that is easier to solve. Generating functions promise to make this simpler, because the use of generating functions reduces the problem to one involving a single function of $2n$ variables (you can probably see where this is heading). The ideal situation would be to have $K = 0$, since the canonical equations for the Q^i and P_i would now give $\dot{Q}^i = 0 = \dot{P}_i$, with solutions $P_i = a_i$, $Q^i = b_i$ for some constants a_i and b_i.

Consider first a type 2 generating function, which we will simply call S, with

$$(\widetilde{\mathrm{F}}_2\text{-a}) \qquad \frac{\partial S(q, P, t)}{\partial q^i} = p_i$$

$$(\widetilde{\mathrm{F}}_2\text{-b}) \qquad \frac{\partial S(q, P, t)}{\partial P_i} = Q^i$$

$$(\widetilde{\mathrm{F}}_2\text{-c}) \qquad K \circ g = H + \frac{\partial S(q, P, t)}{\partial t}.$$

In order to have $K = 0$, the generating function S must satisfy

$$\frac{\partial S(q, P, t)}{\partial t} + H(q^1, \dots, q^n, p_1, \dots, p_n, t) = 0;$$

substituting from $(\widetilde{\mathrm{F}}_2\text{-a})$, we obtain (ah-hah!) the Hamilton–Jacobi equation,

$$\frac{\partial S(q, P, t)}{\partial t} + H\left(q^1, \dots, q^n, \frac{\partial S(q, P, t)}{\partial q^1}, \dots, \frac{\partial S(q, P, t)}{\partial q^n}, t\right) = 0.$$

More precisely, what we have here is the Hamilton–Jacobi equation with n parameters, $P_1, \dots, P_n$, which we want to be the constants $a_1, \dots, a_n$. This means that solving this equation is the same thing as finding a complete integral $\phi(q^1, \dots, q^n, a_1, \dots, a_n)$ of the standard Hamilton–Jacobi equation. There is, in fact, at least one complete integral, namely $S(q, P, t)$, since $(\widetilde{\mathrm{J}}_2)$ shows that $\det(\partial^2\phi/\partial q^j\,\partial a_k) \neq 0$, though of course that doesn't necessarily mean that we'll be able to find one by separation of variables.

If we do succeed in finding a complete integral ϕ, then the desired equations $Q^i = b_i$ become, by $(\widetilde{\mathrm{F}}_2\text{-b})$,

$$\frac{\partial \phi}{\partial a_i}(q^1, \dots, q^n, a_1, \dots, a_n, t) = b_i$$

and $(\widetilde{\mathrm{F}}_2\text{-a})$ then gives

$$p_i = \frac{\partial \phi}{\partial q^i}(q^1, \dots, q^n, a_1, \dots, a_n, t).$$

Thus, we end up with exactly the same set of equations as those in the statement of Jacobi's theorem, so that in this context, canonical transformations merely provide an additional route to Jacobi's theorem (and a more complicated explanation of how to use the Hamilton–Jacobi equation).

The real usefulness of canonical transformations for theoretical work will appear in the following chapters. For the moment, we simply emphasize that a type 2 generating function gives a canonical transformation $(q,p) \mapsto (Q,P)$ for which **the equations $Q^i = b_i$ give solutions to Hamilton's equations.**

For a type 1 generating function, where we have

$$(\widetilde{\mathrm{F}}_1\text{-a}) \qquad \frac{\partial S(q,Q,t)}{\partial q^i} = p_i$$

$$(\widetilde{\mathrm{F}}_1\text{-b}) \qquad \frac{\partial S(q,Q,t)}{\partial Q^i} = -P_i.$$

$$(\widetilde{\mathrm{F}}_1\text{-c}) \qquad K \circ g = H + \frac{\partial S(q,Q,t)}{\partial t},$$

things just get shuffled around a bit. We still end up with the Hamilton–Jacobi equation,

$$\frac{\partial S(q,Q,t)}{\partial t} + H\left(q^1,\dots,q^n, \frac{\partial S(q,Q,t)}{\partial q^1},\dots,\frac{\partial S(q,Q,t)}{\partial q^n}, t\right) = 0,$$

but now the Q^i are the parameters, which we want to be the constants $b_1,\dots,b_n$ and we consider a complete integral $\phi(q^1,\dots,q^n,b_1,\dots,b_n)$ with (F_1) showing that S is in fact a complete integral.

If we succeed in finding a complete integral explicitly, the desired equations $P_i = a_i$ then become, by $(\widetilde{\mathrm{F}}_1\text{-b})$,

$$\frac{\partial \phi}{\partial b_i}(q^1,\dots,q^n,b_1,\dots,b_n,t) = -a_i$$

and $(\widetilde{\mathrm{F}}_1\text{-a})$ then gives

$$p_i = \frac{\partial \phi}{\partial q^i}(q^1,\dots,q^n,b_1,\dots,b_n,t).$$

Analogous to the situation for a type 2 generating function, if there is type 1 generating function, then there is a canonical transformation $(q,p) \mapsto (Q,P)$ for which **the equations $P_j = a_j$ give solutions to Hamilton's equations.**

Generating functions in the time-independent case. In Chapter 21 an extremely important role will be taken by a modification of this approach for the case of a time-independent Hamiltonian $H\colon T^*M \to \mathbb{R}$.

Instead of obtaining a Hamilton–Jacobi equation in which t does not explicitly appear, and then writing $S(q,\alpha,t) = W(q) - \alpha t$, $H(q, \partial W/\partial q) = \alpha$, leading us to solve the Hamilton–Jacobi equation for the characteristic function $W(q,\alpha,\alpha_2,\dots,\alpha_n)$, a more direct approach is simply to look for W right from the start. For example, the equations ($\widetilde{\text{F}}2$-a,b) for a type 2 generating function S not depending on t, which we will now denote by W, are

$$\frac{\partial W(q,P)}{\partial q^i} = p_i, \qquad \frac{\partial W(q,P)}{\partial P_i} = Q^i,$$

which are simply the equations (F2-a,b) from before, while ($\widetilde{\text{F}}2$-c) gives us the equation $K \circ g = H + 0 = \alpha$.

We obviously can't make $K = 0$ now. Instead, when $\partial W(q,P)/\partial q^i$ is substituted for p_i in $H = \alpha$, we get the Hamilton–Jacobi equation for Hamilton's characteristic function $W(q,P)$,

$$H\left(q, \frac{\partial W(q,P)}{\partial q}\right) = \alpha,$$

which we solve as $W(q,\alpha,\alpha_2,\dots,\alpha_n) = \cdots$, and then $K \circ g = H$ means that when we write K in terms of the coordinates (Q,P), as $K(Q,\alpha,\alpha_2,\dots,\alpha_n)$, we simply have $K(Q,\alpha,\alpha_2,\dots,\alpha_n) = \alpha$. So in this coordinate system Hamilton's equations become (in condensed form)

$$\dot P_j = \frac{\partial K}{\partial Q^j} = 0, \qquad \dot Q^1 = \frac{\partial K}{\partial \alpha} = 1, \quad \dot Q^i = \frac{\partial K}{\partial \alpha_i} = 0 \quad i = 2,\dots,n,$$

with the solutions $P_j = \alpha_j$, and $Q^1 = t + b^1$ but $Q^i = b^i$ for $i = 2,\dots,n$.

[As a fairly trivial example, by comparison with the treatment of the harmonic oscillator on pages 550–551, we now solve the Hamilton-Jacobi equation for W, obtaining, as before, $W(q,\alpha) = \omega \int \sqrt{\frac{2\alpha}{\omega^2} - q^2}\, dq$, and we then get

$$t + b = Q^1 = \frac{\partial W(q,\alpha)}{\partial \alpha} = \frac{1}{\omega}\int \frac{dq}{\sqrt{\frac{2\alpha}{\omega^2} - (q(t))^2}},$$

giving the same result.]

More generally, suppose we simply have a generating function $W(q,P)$ with

$$(*) \qquad H\left(q, \frac{\partial W(q,P)}{\partial q}\right) = K(P)$$

for some function K that does not depend on Q (Problem 3 gives a very simple special example where this happens, but the real use for this case occurs in Chapter 21). Then Hamilton's equations for the new Hamiltonian K will simply become

$$\dot P_j = -\frac{\partial K}{\partial Q^j} = 0 \implies P_j = \alpha_j \text{ for constants } \alpha_j, \qquad \dot Q^i = \frac{\partial K}{\partial \alpha_i},$$

which can be solved, in condensed form, as

$$\begin{aligned} Q^i(t) &= Q^i(0) + t\frac{\partial K}{\partial \alpha_i}(\alpha_1, \dots, \alpha_n) \\ &= Q^i(0) + t\nu_i(\alpha_1, \dots, \alpha_n), \quad \text{say.} \end{aligned}$$

In other words, in the coordinate system (Q, P), every solution γ of Hamilton's equations will have constant $P_j(\gamma(t)) = \alpha_i$, while the $Q^i(\gamma(t))$ will all be linear in t, with the coefficients of t determined by those α_i.

Naturally, all these considerations hold, *mutatis mutandis*, for type 1 generating functions.

Despite the apparent efficiency of this approach, and its importance later on, we will often find it conceptually simpler to work directly with the formulas for time-dependent Hamiltonians first, and then, when necessary, specialize to the case of the time-independent Hamiltonian.

Other types of generating functions. The need to consider both type 1 and type 2 generating functions is easily seen from the 1-dimensional example of the harmonic oscillator, with $q = \cos t$, $p = -\sin t$; although these equations simply parameterize the unit circle, q can't be used as a coordinate system in a neighborhood of $(1, 0)$ or $(-1, 0)$, while p can't be used in a neighborhood of $(0, 1)$ or $(0, -1)$.

Problem 4 gives the standard ways to define a type 3 generating function of the form $S(p, Q)$ when we have the condition

(J3) $$\det \frac{\partial Q(q, p)}{\partial q} \neq 0,$$

and a type 4 generating function of the form $S(p, P)$, when we have

(J4) $$\det \frac{\partial P(q, p)}{\partial q} \neq 0,$$

but it is easily seen that none of the conditions (J1)–(J4) is satisfied for the canonical transformation in dimension 2 given by

$$\begin{aligned} Q^1 &= p_1 & P_1 &= -q^1 \\ Q^2 &= q^2 & P_2 &= p_2. \end{aligned}$$

Another useful device is to look for a generating function for the inverse of a given canonical transformation, which will be used in Chapter 22, but that won't help here, either.

In practice, type 1 and type 2 generating functions are the only ones ever needed for solving mechanics problems that actually arise, but that obviously won't satisfy compulsive pure mathematicians. Fortunately, for the canonical transformation just given we have

$$\det \frac{\partial(Q^1, P_2)}{\partial(p_1, p_2)} \neq 0,$$

and we can define a generating function that is a mixture of types 1 and 2,

$$S(q^1, Q^1, q^2, P_2),$$

leading us to ask whether for every canonical transformation in $2n$ dimensions, there is, at each point, always such a mixture. In other words, we want to know whether, for each point, it is always possible to write $\{1, \ldots, n\}$ as the disjoint union $\{1, \ldots, n\} = \mathcal{I}_Q \underset{d}{\cup} \mathcal{I}_P$, in such a way that

$$0 \neq \underset{2n \times 2n}{\det \frac{\partial(q, Q^\alpha, P_\beta)}{\partial(q, p)}} = \underset{n \times n}{\det \frac{\partial(Q^\alpha, P_\beta)}{\partial p}} \qquad \begin{array}{l} \alpha \in \mathcal{I}_Q, \\ \beta \in \mathcal{I}_P. \end{array}$$

The answer is that we can indeed always choose such $\mathcal{I}_Q, \mathcal{I}_P$ [type 3 and type 4 generating functions are in fact virtually never used], and the proof has nothing to do with physics, or even with T^*M. It follows from purely abstract considerations about skew-symmetric functions on vector spaces. So it is fitting that we leave this question to be answered in the next chapter, where we will finally allow ourselves to indulge in one of the pure mathematician's guilty pleasures.

ADDENDUM 19A

TIME-(IN)DEPENDENT HAMILTONIANS

Given a time-independent Hamiltonian, for which we have the energy integral (E) on page 529, we can obtain an equivalent system on a space of two lower dimensions, although the new Hamiltonian will depend on time. This is a classical construction that can be found in Whittaker [1; §141], originally published in 1904, back in the olden days when the differential of a 1-form was still known as the "bilinear covariant". As we will see later, it is also possible to reverse this process and formally reduce a problem for a time-dependent Hamiltonian to one with a Hamiltonian that does not depend on time.

The sketch of the classical construction that we will be giving here follows Arnold [2; §45B], and amounts to a more geometric version of Whittaker's proof. An alternate proof may be found in Abraham and Marsden [1; pg. 391].

Our considerations are local, and for simplicity we simply assume that we are in $M = \mathbb{R}^n \times \mathbb{R}^n$ with coordinates (q, p). We also consider $M \times \mathbb{R} = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$, with coordinates (q, p, t), and the projection

$$\pi_M : M \times \mathbb{R} \to M, \qquad \pi_M(\not p, t) = \not p.$$

We first note that for any 2-form ψ on a manifold N of odd dimension $2n+1$, at each point $\not p$ there is a tangent vector $\mathbf{v}$ such that

$$\psi(\mathbf{v}, \mathbf{w}) = 0 \qquad \text{for all } \mathbf{w} \text{ at } \not p.$$

Proof: the matrix A of ψ is skew-symmetric, $A^{\mathbf{t}} = -A$, so

$$\det A = \det A^{\mathbf{t}} = \det(-A) = (-1)^{2n+1} \det A = -\det A,$$

so $\det A = 0$, and A has an eigenvector $\mathbf{v}$ with eigenvalue 0.

If ψ is nonsingular, then these vectors $\mathbf{v}$ lie in a 1-dimensional subspace, the *characteristic subspace.* In this case, a curve $c : \mathbb{R} \to N$ with non-zero tangent vectors $c'(t)$ is called a *characteristic curve* if $c'(t)$ is always in the characteristic subspace, noting that such curves are determined only up to reparameterization.

In particular, consider the 1-form

$$\theta - H\,dt = \sum_{i=1}^{n} p_i\,dq^i - H\,dt$$

on $M \times \mathbb{R}$, and the differential of this 1-form, denoted by $\mathcal{A}$ in the proof at the bottom of page 574. This proof showed that the characteristic subspace for $\mathcal{A}$ at a point is precisely the set of vectors at that point that are tangent to the flow of the vector field X corresponding to the Hamiltonian H. And this means that if a characteristic curve γ is reparameterized by t, so that it is of the form $(c(t), t)$ for a curve c in M, then c is a solution of Hamilton's equations for H, and conversely.

Now suppose that $(\mathring{q}, \mathring{p})$ is not a critical point of H, so that $dH(\mathring{q}, \mathring{p})$ is non-singular. Renaming the coordinates, if necessary, we can assume that

$$\frac{\partial H}{\partial p_n}(\mathring{q}, \mathring{p}) \neq 0.$$

If $H(\mathring{q}, \mathring{p}) = h$, then by the implicit function theorem, locally we can find a function f with

$$H\left(q^1, \dots, q^n, p_1, \dots, p_{n-1}, f(q^1, \dots, q^{n-1}, -q^n, p_1, \dots, p_{n-1})\right) = h$$

[the $-q^n$ is intentional]. Letting $(\tilde{q}, \tilde{p})$ stand for $(q^1, \dots, q^{n-1}, p_1, \dots, p_{n-1})$, and renaming q_n as $-\tau$, we can say there is a function $K(\tilde{q}, \tilde{p}, \tau)$ with

$$H\left(\tilde{q}, -\tau, \tilde{p}, K(\tilde{q}, \tilde{p}, \tau)\right) = h.$$

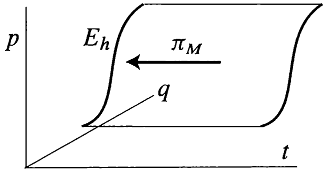
We then have

$$\begin{aligned}\sum_{i=1}^{n} p_i\, dq^i - H\, dt &= \sum_{i=1}^{n-1} p_i\, dq^i - K\, d\tau - H\, dt \\ &= \sum_{i=1}^{n-1} \tilde{p}_i\, d\tilde{q}^i - K\, d\tau - d(Ht) + t\, dH,\end{aligned}$$

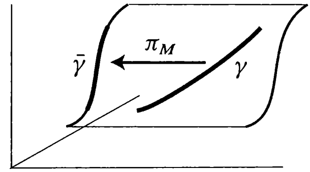
and taking d of this equation we get

$$\begin{aligned}\mathcal{A}_H &= \pi_M{}^*(\mathcal{A}_K) - d(t\,dH) \\ &= \pi_M{}^*(\mathcal{A}_K) \quad \text{on the subspace where } H = h.\end{aligned}$$

In the figure below, (a) shows the set $E_h = \{(q, p) \in M : H(q, p) = h\}$, together with its extension $E_h \times \mathbb{R} \subset M \times \mathbb{R}$. (Unfortunately, since we can



(a)



(b)

at best draw pictures of 3-dimensional objects, the figure is rather misleading, because E_h is usually a submanifold of dimension > 1.)

Now suppose (b) that we have a curve $\gamma(t) = (c(t), t)$ where c is a solution of Hamilton's equations for the Hamiltonian H, and we project γ back down to the curve $\bar{\gamma} = \pi_M \circ \gamma$ in E_h. The original curve γ is a characteristic curve for $\mathcal{A}_H$, so $\bar{\gamma}$ is a characteristic curve for $\mathcal{A}_K$, and thus, when reparameterized by τ, it satisfies Hamilton's equations for the (τ-dependent) Hamiltonian K,

$$\overset{\bullet}{\tilde{q}}{}^i = \frac{\partial K}{\partial \tilde{p}_i}, \qquad \overset{\bullet}{\tilde{p}}_i = -\frac{\partial K}{\partial \tilde{q}^i}, \qquad \bullet = \frac{d}{d\tau}.$$

As one interesting consequence of this reduction, one can now use the extended Hamilton's principle (page 534) to give a rather more direct proof of Maupertuis' form of the Principle of Least Action (page 464), which the reader may work out, or find written out in Arnold [2; §45D]. Note that the additional assumption that $(\mathring{q}, \mathring{p})$ is not a critical point of H with which we have been working is precisely the additional assumption that we found necessary to add at the very end of our argument for the Principle of Least Action, on page 465.

Starting with a Hamiltonian H that depends on time, we can formally reverse this whole construction by letting t be a new space variable. We consider

$$\widetilde{M} = M \times \mathbb{R} \times \mathbb{R} \qquad \text{with coordinates} \qquad (q^1, \dots, q^n, t, p_1, \dots, p_n, p_t),$$

equipped with the symplectic 2-form

$$\widetilde{\omega} = \omega + dp_t \wedge dt \qquad \text{and the Hamiltonian} \qquad \widetilde{H} = H + p_t.$$

For a curve $\tau \mapsto c(\tau) \in \widetilde{M}$, Hamilton's equations for $\widetilde{H}$ are

$$\overset{\bullet}{q}{}^i = \frac{\partial \widetilde{H}}{\partial p_i}, \qquad \overset{\bullet}{p}_i = -\frac{\partial \widetilde{H}}{\partial q^i}, \qquad \overset{\bullet}{t} = 1, \qquad \overset{\bullet}{p}_t = -\frac{\partial \widetilde{H}}{\partial t}, \qquad \bullet = \frac{d}{d\tau}.$$

The first three of these equations show that the solutions for this (τ-independent) system project down on $M \times \mathbb{R}$ to the solutions for the time-dependent system (M, ω, H).

Thus, in theory, it is possible to consider only Hamiltonians that do not depend on time. However, such an approach does not necessarily make things any easier to understand!

It should be mentioned that there are apparently even more general versions of this procedure that sometimes prove to be useful. See Cordani [1; pg. 398].

ADDENDUM 19B

GENERALIZED CANONICAL TRANSFORMATIONS

The fact that generalized canonical transformations also preserve the Hamiltonian structure of equations, and that these are the only such transformations, was first pointed out and proved in Lee [1], where it is derived from a theorem about integral invariants. In our discussion of (relative) integral invariants, it should perhaps have been pointed out that any particular Hamiltonian H will generally have many other such invariants. The point about $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ is that they are "universal" integral invariants for all H, as, of course, are the higher-dimensional $\boldsymbol{\theta}\wedge\boldsymbol{\omega}$, $\boldsymbol{\omega}\wedge\boldsymbol{\omega}$, etc., and Lee proved that the only universal invariants are constant multiples of these, from which the result about generalized canonical transformations is easily deduced (Problem 5). Here we give a direct proof of the result for generalized canonical transformations.

LEMMA. $C^{\mathbf{t}}\mathbb{J}C = \mathbb{J} \iff C\mathbb{J}C^{\mathbf{t}} = \mathbb{J}$.

PROOF. $C^{\mathbf{t}}\mathbb{J}C = \mathbb{J}$ implies $C^{-1}\mathbb{J}^{-1}(C^{\mathbf{t}})^{-1} = \mathbb{J}^{-1}$, and since $\mathbb{J}^{-1} = -\mathbb{J}$, this gives

$$C^{-1}\mathbb{J}(C^{\mathbf{t}})^{-1} = \mathbb{J}.$$

Multiplying by C on the left and $C^{\mathbf{t}}$ on the right then gives $\mathbb{J} = C\mathbb{J}C^{\mathbf{t}}$. ❖

5. THEOREM. If $f\colon T^*M \to T^*M$ takes all Hamiltonian vector fields into Hamiltonian vector fields, then f is a generalized canonical transformation.

PROOF. Given a Hamiltonian K on T^*M, let $H = K \circ f$, so that

$$(DH)^{\mathbf{t}} = (Df)^{\mathbf{t}}(DK)^{\mathbf{t}} \quad \text{(first line of equation (d) on page 568)}$$

and for a curve c in T^*M let $\gamma = f \circ c$, so that

$$\dot{c}^{\mathbf{t}} = (Df^{-1})\dot{\gamma}^{\mathbf{t}} \qquad \text{(equation (c) on page 568).}$$

Then

$$\dot{c}^{\mathbf{t}} - \mathbb{J}(DH)^{\mathbf{t}} = (Df^{-1})\dot{\gamma}^{\mathbf{t}} - \mathbb{J}(Df)^{\mathbf{t}}(DK)^{\mathbf{t}},$$

so we have

$$\begin{aligned}(Df)\big(\dot{c}^{\mathbf{t}} - \mathbb{J}(DH)^{\mathbf{t}}\big) &= \dot{\gamma}^{\mathbf{t}} - (Df)\mathbb{J}(Df)^{\mathbf{t}}(DK)^{\mathbf{t}}\\ &= \dot{\gamma}^{\mathbf{t}} - \mathbb{J}\big(-\mathbb{J}(Df)\mathbb{J}(Df)^{\mathbf{t}}\big)(DK)^{\mathbf{t}}\\ &= \dot{\gamma}^{\mathbf{t}} - \mathbb{J}P(DK)^{\mathbf{t}}, \quad \text{say.}\end{aligned}$$

Since by hypothesis f takes curves satisfying $\dot{c}^{\mathbf{t}} - \mathbb{J}(DH)^{\mathbf{t}} = 0$ into curves γ satisfying Hamilton's equations for some Hamiltonian, it follows that for all K, the product $P(DK)^{\mathbf{t}}$ must be $(D\bar{K})^{\mathbf{t}}$ for some function $\bar{K}$, so that

$$\frac{\partial \bar{K}}{\partial x_i} = \sum_{k=1}^{2n} P_{ik} \frac{\partial K}{\partial x_k}.$$

Using equality of mixed partials of $\bar{K}$ we find that

$$\sum_{k=1}^{2n} \frac{\partial P_{ik}}{\partial x_j} \frac{\partial K}{\partial x_k} + P_{ik} \frac{\partial^2 K}{\partial x_j \partial x_k} = \sum_{k=1}^{2n} \frac{\partial P_{jk}}{\partial x_i} \frac{\partial K}{\partial x_k} + P_{jk} \frac{\partial^2 K}{\partial x_i \partial x_k}.$$

Choosing K for which the second partials are 0, we see that we have separately

$$\sum_{k=1}^{2n} \frac{\partial P_{ik}}{\partial x_j} \frac{\partial K}{\partial x_k} = \sum_{k=1}^{2n} \frac{\partial P_{jk}}{\partial x_i} \frac{\partial K}{\partial x_k} \tag{1}$$

and thus also separately

$$\sum_{k=1}^{2n} P_{ik} \frac{\partial^2 K}{\partial x_j \partial x_k} = \sum_{k=1}^{2n} P_{jk} \frac{\partial^2 K}{\partial x_i \partial x_k}. \tag{2}$$

From (2) we can see that $P_{ij} \neq 0$ only for $i = j$, and moreover that all $P_{ii} = a$ for the same function a. Then from (1) we can see that a is a constant. So $P = aI_{2n}$, or

$$-\mathbb{J}(Df)\mathbb{J}(Df)^{\mathbf{t}} = aI_{2n} \implies (Df)\mathbb{J}(Df)^{\mathbf{t}} = \mathbb{J}aI_{2n} = a\mathbb{J},$$

which, using the Lemma, is the condition for f to be a generalized canonical transformation. ❖

Remarks. This proof comes from Giaquinta and Hildebrant [1; Chap. 9, §3.1]; a proof may also be found in Pars [1; §25.4], while the more general version for time-dependent Hamiltonians is considered in Siegel and Moser [1]; several more recent texts also have proofs. All the proofs, including Lee's, seem to involve similar manipulations.

PROBLEMS

1. Recall that if $\{\phi_t\}$ is the flow of X, we define $L_X\lambda$ for a k-form λ by

$$L_X\lambda = \lim_{h\to 0}\frac{1}{h}[(\phi_h{}^*\lambda - \lambda].$$

It is easy to see that $L_X\lambda$ is also a k-form. If λ is a 0-form, i.e., a function f, then $L_X f = Xf = df(X)$.

(a) Show that

$$\begin{aligned} L_X(\lambda_1+\lambda_2) &= L_X\lambda_1 + L_X\lambda_2 \\ L_X(f\lambda) &= Xf\cdot\lambda f\cdot L_X\lambda \\ L_X(\lambda\wedge\mu) &= L_X\lambda\wedge\mu + \lambda\wedge L_X\mu. \end{aligned}$$

(b) For

$$\begin{aligned} L_X(dx^i) &= \lim_{h\to 0}\frac{1}{h}[(\phi_h{}^*)(dx^i) - dx^i] \\ &= \lim_{h\to 0}\frac{1}{h}\left[\sum_{j=1}^{n}\frac{\partial(x^i\circ\phi_h)}{\partial x^j}\,dx^j - dx^i\right] \end{aligned}$$

the coefficient of dx^j is

$$\lim_{h\to 0}\frac{1}{h}\left[\frac{\partial(x^i\circ\phi_h)}{\partial x^j} - \frac{\partial(x^i\circ\phi_0)}{\partial x^j}\right].$$

Show that this equals

$$\frac{\partial}{\partial x^j}\lim_{h\to 0}\frac{1}{h}[(x^i\circ\phi_h) - (x^i\circ\phi_0)].$$

Hint: Consider the map $A(h,q) = x^i(\phi_h(q))$ from $\mathbb{R}\times M$ to $\mathbb{R}$.

(c) If $X = \sum_{i=1}^{n} a^i\partial/\partial x^i$, then

$$L_X dx^i = \sum_{j=1}^{n}\frac{\partial a^i}{\partial x^j}\,dx^j.$$

(d) Conclude that $L_X(dx^i) = d(L_X x^i)$ and then that in general

$$L_X d\lambda = d(L_X\lambda).$$

(e) For a k-form λ, the definition

$$X \lrcorner\, \lambda(X_2,\ldots,X_k) = \lambda(X,X_2,\ldots,X_k)$$

is naturally extended to mean that for a 1-form λ we have $X\lrcorner\,\omega = \omega(X)$, and we also set $X\lrcorner\, f = 0$ for a function (0-form) f. Check that Cartan's Magic Formula holds for the 0-form f, and for the 1-form dx^i, and conclude that it holds for all k-forms.

(f) Prove Cartan's Bracket Formula similarly.

2. (a) Determine the canonical transformation produced by the type 1 generating function

$$S = \sum_{i=1}^{n} q^i Q^i.$$

(b) Same question for the type 2 generating function

$$S = \sum_{i=1}^{n} q^i P_i.$$

(c) For the type 2 generating function

$$S = \sum_{i=1}^{n} f_i(q,t) P_i,$$

show that $Q^i = f_i(q,t)$, so that we can obtain an arbitrary function $f\colon M \to M$, and that the canonical transformation is simply $g = f^*\colon T^*M \to T^*M$. These are sometimes referred to as "point transformations".
(d) Check that not only does g preserve $\boldsymbol{\omega}$, but in fact g preserves the 1-form

$$\boldsymbol{\theta} = \sum_{i=1}^{n} p_i \, dq^i.$$

Such transformations are sometimes called "homogeneous canonical transformations".
Proving this directly can be confusing, especially if one wants a coordinate-free proof. It also turns out all homogeneous canonical transformations are point transformations. For these matters, see Abraham and Marsden [1; Th. 3.2.12 and Ex. 3.2F] or Marsden and Ratiu [1; §6.3].

3. (a) Show that the type 1 generating function

$$S(q,Q) = \frac{1}{2}\omega q^2 \cot Q$$

gives the canonical transformation on page 571.
(b) Show that this transformation changes the Hamiltonian

$$H(q,p) = \frac{1}{2}(p^2 + \omega^2 q^2)$$

for the harmonic oscillator to

$$\bar{H} = \omega P \implies P = \frac{E}{\omega},$$

so that

$$\dot{Q} = \frac{\partial \bar{H}}{\partial P} = \omega \implies q = \frac{\sqrt{2E}}{\omega} \sin(\omega t + b),$$

equivalent to the solution on page 551 found in Chapter 18.

4. Because of equation (F_1-b) on page 578, the type 2 generating function S_2 may be thought of as the negative of the Legendre transform of the type 1 generating function S_1:

$$\begin{aligned}\mathcal{L}(S_1)(q,P) &= \sum_{i=1}^{n} \frac{\partial S_1}{\partial Q^i}(q,Q)Q^i - S_1(q,Q)\\ &= -\left(\sum_{i=1}^{n} P_i Q^i(q,P) + S_1(q,Q)\right) = -\left(\mathcal{S}(q,p) + \sum_{i=1}^{n} P_i Q^i\right).\end{aligned}$$

(a) Similarly, when we have the condition (J_3), use the Legendre transform

$$\mathcal{L}(S_1)(p,Q) = \sum_{i=1}^{n} \frac{\partial S_1}{\partial q^i} q^i - S_1(q,Q) = \sum_{i=1}^{n} p_i q^i - S_1(q,Q)$$

to define a type 3 generating function

$$S_3(Q(q,p),p) = \mathcal{S}(q,p) + \sum_{i=1}^{n} p_i q^i,$$

and show that

$$\begin{aligned}\frac{\partial S_3(p,Q,t)}{\partial p_i} &= -q^i,\\ \frac{\partial S_3(p,Q,t)}{\partial Q^i} &= -P_i,\\ K \circ g &= H + \frac{\partial S_3}{\partial t}.\end{aligned}$$

(b) When we have the condition (J_4), use a double Legendre transformation to define

$$S_4(p,P(q,p)) = \mathcal{S}(q,p) + \sum_{i=1}^{n} P_i Q^i - \sum_{i=1}^{n} p_i q^i,$$

and show that

$$\begin{aligned}\frac{\partial S_4(p,P,t)}{\partial p_i} &= -q_i,\\ \frac{\partial S_4(p,P,t)}{\partial P_i} &= Q^i,\\ K \circ g &= H + \frac{\partial S_4}{\partial t}.\end{aligned}$$

5. A *universal 2-dimensional integral invariant* is a 2-form ϖ on T^*M with the property that for any 2-chain D we have

$$\int_D \varpi = \int_{\phi_t(D)} \varpi$$

for the flow $\{\phi_t\}$ of any vector field $\mathbf{X}_H$ given by $\mathbf{X}_H \lrcorner\, \omega = -dH$ for some H. Lee's theorem says that any such ϖ is just $c\omega$ for some constant c.

Conclude that if $f \colon T^*M \to T^*M$ has the property that f_* takes any vector field of the form $\mathbf{X}_H$ into one of the form $\mathbf{X}_{\bar{H}}$, then $f^*\omega$ is $c\omega$ for some constant c.

CHAPTER 20

SYMPLECTIC MANIFOLDS

Generalize!
Let no one else's work evade your eyes,
Remember why the good Lord made your eyes,
So don't shade your eyes,
But generalize, generalize, generalize!

— with apologies to Tom Lehrer

Like many another specialized type of manifold, symplectic manifolds are based on a particular structure on vector spaces.

Symplectic vector spaces. A bilinear map $\omega \colon V \times V \to \mathbb{R}$ on a real vector space V is **nondegenerate** if it has the following property:

$$\text{if } \omega(\mathbf{v}, \mathbf{w}) = 0 \text{ for all } \mathbf{w}, \text{ then } \mathbf{v} = 0.$$

When ω is symmetric, with $\omega(v, w)$ often denoted by $\langle v, w \rangle$, we call V together with $\langle\ ,\ \rangle$ an inner product space. When ω is *skew-symmetric*, on the other hand, with $\omega(v, w)$ often denoted by $[v, w]$, we call V together with $[\ ,\ ]$ a **symplectic** vector space,[1] and $[\ ,\ ]$ is often called the **skew-inner product**.

For a basis $\mathbf{e}_1, \ldots, \mathbf{e}_d$ of V, the matrix $A = (a_{ij})$ of ω is defined by

$$a_{ij} = \omega(\mathbf{e}_i, \mathbf{e}_j) \qquad i, j = 1, \ldots, d.$$

Equivalently, if we define $\alpha \colon V \to V^*$ by

$$\alpha(v)(w) = \omega(v, w),$$

then the matrix of α with respect to the basis $\mathbf{e}_i$ of V and the dual basis $\mathbf{e}_i{}^*$ of V^* is $A^{\mathbf{t}}$ (with our convention for writing matrices with the columns representing the image vectors). The **rank** of ω is the rank of α, that is, the dimension of $\alpha(V)$, so ω being nondegenerate is equivalent to saying that the rank of ω is the dimension of V.

[1] The origins of the strange word "symplectic" are discussed in Weinstein [2].

When ω is symmetric or skew-symmetric, we can find a basis for which A is especially simple, even without the nondegeneracy condition. For symmetric ω, we first note that either version of the "polarization identity"

$$\langle v, w\rangle = \tfrac{1}{2}\big[\langle v+w, v+w\rangle - \langle v, v\rangle - \langle w, w\rangle\big],$$
$$\langle v, w\rangle = \tfrac{1}{4}\big[\langle v+w, v+w\rangle - \langle v-w, v-w\rangle\big],$$

shows that if $\langle\ ,\ \rangle$ is not identically 0, then $\langle v, v\rangle$ is not always 0, and hence there is some $\mathbf{e}_1 \in V$ such that $\langle \mathbf{e}_1, \mathbf{e}_1\rangle = \alpha_1 = \pm 1$. For $\mathbf{e}_1{}^\perp = \{v \in V : \langle v, \mathbf{e}_1\rangle = 0\}$ we have

$$(\mathbb{R}\cdot \mathbf{e}_1) \oplus \mathbf{e}_1{}^\perp = V,$$

since the intersection of the two subspaces is clearly $\{0\}$, while for any $v \in V$ we have

$$v - \alpha_1\langle v, \mathbf{e}_1\rangle \mathbf{e}_1 \in \mathbf{e}_1{}^\perp.$$

If $\langle\ ,\ \rangle$ is non-zero on $\mathbf{e}_1{}^\perp$, then we can choose $\mathbf{e}_2 \in \mathbf{e}_1{}^\perp$ with $\langle \mathbf{e}_2, \mathbf{e}_2\rangle = \alpha_2 = \pm 1$. Continuing in this way, we find a basis $\mathbf{e}_1, \dots, \mathbf{e}_d$ for which the matrix of $\langle\ ,\ \rangle$ is

$$\begin{pmatrix} \alpha_1 & & & & & \\ & \ddots & & & \text{\Large 0} & \\ & & \alpha_r & & & \\ & & & 0 & & \\ & \text{\Large 0} & & & \ddots & \\ & & & & & 0 \end{pmatrix} \qquad \alpha_i = \pm 1.$$

When ω is nondegenerate, this is just the Gramm-Schmidt orthogonalization process. Note that in this case, for any given non-zero vector $\mathbf{e}$, some multiple of $\mathbf{e}$ can be chosen as $\mathbf{e}_1$.

In the case of a skew-inner product $\omega = [\ ,\ ]$, the situation is a little more involved. The rank of ω turns out to be $2n$ for some integer n, and we get a basis for which the matrix involves our old friend $\mathbb{J} = \mathbb{J}_n$,

$$\begin{array}{cc} 2n & d-2n \\ \multicolumn{2}{c}{\left(\begin{array}{c|c} \mathbb{J}_n & \text{\Large 0} \\ \hline 0 & 0 \end{array}\right)} \end{array} \qquad \text{i.e.,} \qquad \left(\begin{array}{c|c|c} 0 & I_n & 0 \\ \hline -I_n & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right),$$

but the proof isn't much more complicated. If $\omega \neq 0$, then there are two vectors, which we will call $\mathbf{e}_1$ and $\mathbf{e}_{n+1}$, with $\omega(\mathbf{e}_1, \mathbf{e}_{n+1}) \neq 0$, and by dividing $\mathbf{e}_1$ by a constant we can assume that $\omega(\mathbf{e}_1, \mathbf{e}_{n+1}) = 1$. In the 2-dimensional subspace W

spanned by $\mathbf{e}_1, \mathbf{e}_{n+1}$, the matrix of ω is

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

Now we have $V = W \oplus W^{\perp}$, and we continue, exactly as in the symmetric case. Again, for the nondegenerate case, a multiple of any non-zero vector can be chosen as $\mathbf{e}_1$.

It is straightforward to check that for the dual basis $\mathbf{e}_1{}^*, \dots, \mathbf{e}_d{}^*$, we can write ω as

(a) $$\omega = \sum_{i=1}^{n} \mathbf{e}_i{}^* \wedge \mathbf{e}_{n+i}{}^*, \qquad 2n \le d.$$

Clearly, a skew-symmetric ω can be nondegenerate only if V has even dimension $d = 2n$, and in this case we find, as on page 512, that in $\Lambda^n(V)$, the vector space of alternating functions from the n-fold product $V \times \cdots \times V$ to $\mathbb{R}$, we have

(b) $$\omega \wedge \cdots \wedge \omega = (-1)^N n!\, \mathbf{e}_1{}^* \wedge \cdots \wedge \mathbf{e}_{2n}{}^*$$

for some N. This is a non-zero element of $\Lambda^n(V)$, and thus determines an orientation of V.

If V with $[\ ,\ ]$ is a symplectic vector space of dimension $2n$, a linear transformation $T: V \to V$ is naturally called **symplectic** if

$$[Tv, Tw] = [v, w]$$

for all $v, w \in V$; any basis $\mathbf{e}_1, \dots, \mathbf{e}_{2n}$ for which equation (a) holds is called a **symplectic basis**; and T is symplectic if and only if takes any symplectic basis into a symplectic basis. As on page 569, we conclude that $\det T = 1$. There are obvious extensions to linear transformation from one symplectic vector space to another.

Isotropic subspaces. Given two vectors v and w in a symplectic vector space V, we will call them **skew-orthogonal**, $v \perp w$, if $[v, w] = 0$. A subspace W of a symplectic vector space V is called **isotropic** if $W \perp W$, that is, if $v \perp w$ for all $v, w \in W$, or $[v, w] = 0$ for all $v, w \in W$. In particular, suppose we choose a symplectic basis, which we will suggestively call $(q^1, \dots, q^n, p_1, \dots, p_n)$, and we choose a set of indices $\mathcal{I}_Q$ and a set $\mathcal{I}_P$ so that $\{1, \dots, n\}$ is the disjoint union $\{1, \dots, n\} = \mathcal{I}_Q \underset{d}{\cup} \mathcal{I}_P$. Then (a) shows that $\mathcal{I}$ is isotropic, where

$$\mathcal{I} = \text{subspace spanned by all } q^i \text{ for } i \in \mathcal{I}_Q \text{ and all } p_j \text{ for } j \in \mathcal{I}_P;$$

there are 2^n of these isotropic "coordinate subspaces".

In terms of $(q^1, \dots, q^n, p_1, \dots, p_n)$, we can define an inner product on V by declaring this basis to be orthonormal (this is not invariant, but will still be temporarily useful). Since $[\ ,\ ]: V \times V \to \mathbb{R}$ is a bilinear function on this inner product space, there is, as for any bilinear function, a unique linear transformation $T\colon V \to V$ such that

(c) $$[v, w] = \langle T(v), w\rangle$$

for all $v, w \in V$ (the matrix of T with respect to the chosen basis is $\mathbb{J}$).

It follows from (c) that W is isotropic if and only if $T(W)$ is orthogonal to W with respect to the inner product $\langle\ ,\ \rangle$ just defined. Hence:

LEMMA. The dimension of an isotropic subspace of V is $\le n$.

PROOF. The isotropic subspace W and the subspace $T(W)$ have the same dimension, and they can't be orthogonal if they both have dimension $> n$. ❖

THEOREM. For every n-dimensional isotropic subspace W of V, at least one of the 2^n coordinate isotropic subspaces $\mathcal{I}$ is transverse to W, that is, the intersection $W \cap \mathcal{I} = \{0\}$.

PROOF. Let $\mathcal{Q}$ be the isotropic subspace spanned by $q^1, \dots, q^n$, and let

$$W_1 = W \cap \mathcal{Q}, \qquad \dim W_1 = k \le n.$$

Any k-dimensional subspace of $\mathcal{Q}$, in particular the subspace W_1, is transverse to some $n-k$-dimensional coordinate subspace of $\mathcal{Q}$, say

$$\mathcal{Q}_1 = \text{subspace spanned by } q^{i_1}, \dots, q^{i_{n-k}}, \qquad \begin{aligned} W_1 + \mathcal{Q}_1 &= \mathcal{Q}, \\ W_1 \cap \mathcal{Q}_1 &= \{0\}. \end{aligned}$$

Choosing $j_1, \dots, j_k$ so that $\{i_1, \dots, i_{n-k}, j_1, \dots, j_k\}$ are n distinct numbers in $\{1, \dots, n\}$, let

$$\mathcal{I} = \text{subspace spanned by } q^{i_1}, \dots, q^{i_{n-k}}, p_{j_1}, \dots, p_{j_k}, \qquad \mathcal{Q}_1 = \mathcal{I} \cap \mathcal{Q}.$$

We claim that

$$W \cap \mathcal{I} = \{0\}.$$

To prove this, we note that

$$\begin{aligned} W_1 \subset W \text{ and } W \perp W &\implies W_1 \perp W \\ \mathcal{Q}_1 \subset \mathcal{I} \text{ and } \mathcal{I} \perp \mathcal{I} &\implies \mathcal{Q}_1 \perp \mathcal{I}, \end{aligned}$$

which implies that

$$W_1 + \mathcal{Q}_1 \perp W \cap \mathcal{I} \implies \mathcal{Q} \perp W \cap \mathcal{I}.$$

Since $\mathcal{Q}$ has dimension n, the Lemma implies that any vector skew-orthogonal to $\mathcal{Q}$ must be in $\mathcal{Q}$, so we have

$$W \cap \mathcal{I} \subset \mathcal{Q} \implies W \cap \mathcal{I} = (W \cap \mathcal{Q}) \cap (\mathcal{I} \cap \mathcal{Q}) = W_1 \cap \mathcal{Q}_1 = \{0\}. \text{ ❖}$$

On the manifold T^*M we have the standard 2-form ω, which gives a symplectic structure on each tangent space, and a canonical map is just the same as one that is symplectic on each tangent space (going into the tangent space at a different point).

1. COROLLARY.[1] If $g\colon T^*M \to T^*M$ is canonical, then in a neighborhood of any point $(\mathring{q}, \mathring{p})$ we can write $\{1, \dots, n\} = \mathcal{I}_Q \underset{d}{\cup} \mathcal{I}_P$ in such a way that, in the notation introduced on page 589,

$$(*) \qquad 0 \neq \det \left. \frac{\partial(q, Q^\alpha, P_\beta)}{\partial(q, p)} \right|_{(\mathring{q}, \mathring{p})} \qquad \begin{array}{l} \alpha \in \mathcal{I}_Q, \\ \beta \in \mathcal{I}_P. \end{array}$$

PROOF. Considering T^*M as $\mathbb{R}^{2n}$, in the tangent space of $\mathbb{R}^{2n}$ at $(\mathring{q}, \mathring{p})$, let $\mathcal{Q}$ be the isotropic subspace spanned by $q^1, \dots, q^n$, and let

$$W = g_*(\mathcal{Q}) = g_{*(\mathring{q}, \mathring{p})}(\mathcal{Q}).$$

Since g is canonical, g_* is symplectic, so W is isotropic.

Let $\mathcal{I}$ be the isotropic coordinate space transverse to W given by the theorem, determined by certain q^i and p_j, and let $\mathcal{K}$ be the isotropic coordinate space determined by all the other q^r and p_s, so that the tangent space of $\mathbb{R}^{2n}$ at $(\mathring{p}, \mathring{q})$ is the direct sum $\mathcal{K} \oplus \mathcal{I}$. Finally, let $\pi_{\mathcal{K}}$ be the projection onto $\mathcal{K}$ determined by this direct sum decomposition.

Since $\mathcal{I}$ is transverse to W, it follows that $\pi_{\mathcal{K}}$ is one-to-one on $W = g_*(\mathcal{Q})$, and this is equivalent to $(*)$. ❖

We can easily extend this to a canonical $g\colon T^*M \times \mathbb{R} \to T^*M \times \mathbb{R}$: in a neighborhood of any point $\mathring{q}, \mathring{p}, t_0$ we can write $\{1, \dots, n\} = \mathcal{I}_Q \underset{d}{\cup} \mathcal{I}_P$ in such a way that

$$(*) \qquad 0 \neq \det \left. \frac{\partial(q, Q^\alpha, P_\beta, t)}{\partial(q, p, t)} \right|_{(\mathring{q}, \mathring{p}, t_0)} \qquad \begin{array}{l} \alpha \in \mathcal{I}_Q, \\ \beta \in \mathcal{I}_P. \end{array}$$

Applying this to the discussions on pages 586 and 589, we have

2. COROLLARY. For a Hamiltonian system H there is a canonical transformation $(q, p) \mapsto (Q, P)$ and $\mathcal{I}_Q$ and $\mathcal{I}_P$ as above, for which Q^α = constant are solutions for $\alpha \in \mathcal{I}_Q$ and P_β = constant are solutions for $\beta \in \mathcal{I}_P$.

---

[1] This proof, and the preceding theorem, comes from Arnold [2; §48B]. A less abstractly presented proof, perhaps the original(?), appears in Carathéodory [1; §§96, 103].

Symplectic manifolds. A 2-form ω on M which is nondegenerate, as defined on page 512, is simply one for which each $\omega(p)$ is nondegenerate on the tangent space M_p, thus giving us a symplectic structure on each tangent space. One might think that (M, ω) would be called a "symplectic manifold", but the definition actually adds one more condition:

DEFINITION. (M, ω) is a **symplectic manifold** if ω is nondegenerate and

$$d\omega = 0.$$

The prototypical example of a symplectic manifold is, of course, $(T^*M, \boldsymbol{\omega})$. [Despite the possibility of confusion, we use M both for a symplectic manifold, and in situations like T^*M.] A classical theorem of Darboux shows that locally every symplectic manifold looks like such a manifold. Happily, there is now a delightfully short proof of this theorem, due to Weinstein [1], after Moser [1] (also compare DG, Prob. 8-27), which has the advantage of working for infinite dimensional manifolds, although we won't be considering such manifolds here.

THEOREM (DARBOUX). Let ω be a nondegenerate 2-form on a manifold M of dimension $2n$. Then $d\omega = 0$ if and only if around each point $p \in M$ there is a **symplectic coordinate system**, that is, a coordinate system $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ in which

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq^i.$$

[Thus $d\omega = 0$ for a nondegenerate skew-symmetric ω corresponds to $R = 0$ for a nondegenerate symmetric $\langle\ ,\ \rangle$ (a Riemannian metric), since $R = 0$ is the condition that there is a coordinate system for which $\langle\ ,\ \rangle = \sum_{i=1}^{n} dx^i \otimes dx^i$.]

PROOF. If ω has this form, obviously $d\omega = 0$. For the converse, it suffices to find around each point a coordinate system in which ω has constant coefficients, since the procedure on pages 600–601 can then be applied simultaneously to all tangent spaces at the points on which the coordinate system is defined.

Since we are working locally, we assume that $M = \mathbb{R}^n$, and let p be the origin 0. Let ω_1 be the form with constant coefficients that equals $\omega(0)$ at 0, and define

$$\omega_t = t(\omega_1) + (1-t)\omega = \omega + t(\omega_1 - \omega) \qquad 0 \le t \le 1.$$

Then each $\omega_t(0) = \omega(0)$ is nondegenerate, so there is a neighborhood of 0 on which each ω_t is nondegenerate for all $0 \le t \le 1$. Shrinking this down to a ball, we can write the closed form $\omega_1 - \omega = d\alpha$ for some 1-form α, and we can assume $\alpha(0) = 0$.

For each t, let X^t be the smooth vector field with $X^t(0) = 0$ and

$$X^t \lrcorner\, \omega_t = -\alpha.$$

Since $X^t(0) = 0$, by restricting the ball further if necessary, we can assume that the 1-parameter family of diffeomorphisms f_t generated by X^t are all defined on $[0, 1]$. (This 1-parameter family is not usually a 1-parameter group.)

We then have

$$\begin{aligned}\frac{d}{dt}(f_t{}^*\omega_t) &= f_t{}^*(L_{\mathbf{X}_t}\omega_t) + f_t{}^*\Big(\frac{d}{dt}\omega_t\Big) && \text{(cf. Problem 1)}\\ &= f_t{}^*(d(X^t \lrcorner\, \omega_t) + 0)) + f_t{}^*(\omega_1 - \omega) && \text{by the Cartan formula}\\ &= f_t{}^*(-d\alpha + \omega_1 - \omega) = 0.\end{aligned}$$

Therefore $f_1{}^*\omega_1 = f_0{}^*\omega_0 = \omega$, so f_1 is the coordinate system that transforms ω to the constant form ω_1. ❖

Aside from making abstract symplectic manifolds analogous to the examples $(T^*M, \boldsymbol{\omega})$ that spurred our interest in symplectic structures in the first place, $d\omega = 0$ is also the condition that will make Poisson brackets work on abstract symplectic manifolds.

Poisson brackets. The definition of Poisson brackets was originally made for $T^*M \times \mathbb{R}$ (actually, $\mathbb{R}^{2n+1}$), and that is where our discussion will begin also. Given a Hamiltonian H on $T^*M \times \mathbb{R}$, in order for a function $f : T^*M \times \mathbb{R} \to \mathbb{R}$ to be an "integral of the motion", that is, for f to have a constant value on each solution $c : \mathbb{R} \to T^*M$ of Hamilton's equations, we need for each solution c that

$$\begin{aligned}0 &= \frac{d}{dt} f(c(t), t)\\ &= \frac{\partial f}{\partial t}(c(t), t) + \sum_{i=1}^{n} \frac{\partial f}{\partial q^i}(c(t), t) \cdot (q^i \circ c)'(t) + \frac{\partial f}{\partial p_i}(c(t), t) \cdot (p_i \circ c)'(t)\\ &= \frac{\partial f}{\partial t}(c(t), t) + \sum_{i=1}^{n} \frac{\partial f}{\partial q^i}(c(t)) \cdot \frac{\partial H}{\partial p_i}(c(t)) - \frac{\partial f}{\partial p_i}(c(t)) \cdot \frac{\partial H}{\partial q^i}(c(t)).\end{aligned}$$

Since there are solutions through every point, we thus need

$$0 = \frac{\partial f}{\partial t} + \sum_{i=1}^{n} \frac{\partial f}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q^i} \qquad \text{on } T^*M.$$

For any two functions $f, g : T^*M \times \mathbb{R} \to \mathbb{R}$, the **Poisson bracket** $\{f, g\}$ was defined classically as

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}.$$

Thus, f is an integral for H if and only if

$$\frac{\partial f}{\partial t} + \{f, H\} = 0.$$

In particular, in the case where H and f do not depend on t, we simply have the condition

$$\{f, H\} = 0.$$

It will be convenient to restrict our attention mainly to this case, although the general case will arise a couple of times.

Since this classical definition depends on the choice of the canonical coordinates (q, p), a proof was needed that it is actually well-defined—for a canonical transformation, given by (Q, P) as on page 570, we have $\{f, g\}_{[q,p]} = \{f, g\}_{[Q,P]}$, with the subscripts denoting the coordinate systems used for the computations. But it is easy to give an invariant definition right from the start. Remembering that the invariant definition

$$\mathbf{X}_g \lrcorner\, \boldsymbol{\omega} = -dg$$

can be written, as on page 533, as

$$\mathbf{X}_g = \sum_{i=1}^{n} \frac{\partial g}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial g}{\partial q^i}\frac{\partial}{\partial p_i},$$

we see that

$$\{f, g\} = \mathbf{X}_g(f).$$

Moreover, since $\mathbf{X}_f \lrcorner\, \boldsymbol{\omega} = -df$, we have

$$\boldsymbol{\omega}(\mathbf{X}_g, \mathbf{X}_f) = -df(\mathbf{X}_g) = -\{f, g\},$$

or simply

$$\{f, g\} = \boldsymbol{\omega}(\mathbf{X}_f, \mathbf{X}_g).$$

Whichever route we take to defining the Poisson bracket, it is clear that $\{\ ,\ \}$ is skew-symmetric and bilinear over $\mathbb{R}$. One immediate consequence of skew-symmetry is the

HAMILTONIAN FORM OF NOETHER'S THEOREM. Given functions $f, g\colon T^*M \to \mathbb{R}$, suppose that the function f is constant along the 1-parameter group of canonical transformations determined by the vector field $\mathbf{X}_g$. Then g is an integral of motion for Hamilton's equations with Hamiltonian f.

PROOF. The hypothesis says that $0 = \mathbf{X}_g(f)$, so that $0 = \{f, g\}$. Hence $0 = \{g, f\} = \mathbf{X}_f(g)$. ❖

The next thing we want to note is that Poisson brackets provide an easy test for a map to be canonical:

THEOREM. $\phi\colon T^*M \to T^*M$ is canonical if and only if ϕ preserves $\{\ ,\ \}$, that is,

$$\phi^*\{f, g\} = \{\phi^* f, \phi^* g\} \quad \text{i.e.,} \quad \{f, g\} \circ \phi^{-1} = \{f \circ \phi^{-1}, g \circ \phi^{-1}\}.$$

PROOF. An unraveling of definitions shows that for a vector field X on M we have

$$X(f) \circ \phi^{-1} = \phi_* X(f \circ \phi^{-1}).$$

Therefore

$$\{f, g\} \circ \phi^{-1} = \mathbf{X}_g(f) \circ \phi^{-1} = \phi_* \mathbf{X}_g(f \circ \phi^{-1}),$$

while

$$\{f \circ \phi^{-1}, g \circ \phi^{-1}\} = \mathbf{X}_{g \circ \phi^{-1}}(f \circ \phi^{-1}).$$

Thus $\{f, g\} \circ \phi^{-1} = \{f \circ \phi^{-1}, g \circ \phi^{-1}\}$ if and only if

$$\mathbf{X}_{g \circ \phi^{-1}} = \phi_* \mathbf{X}_g$$

for all g, which by Theorem 1 on page 570 is true if and only if ϕ is canonical. ❖

It is easy to check that for any canonical coordinate system (q, p), we have the "fundamental Poisson brackets"

$$\{q^i, q^j\} = 0, \qquad \{p_i, p_j\} = 0, \qquad \{q^i, p_j\} = \delta^i_j,$$

and conversely, these relationships implies that (q, p) is canonical (Problem 2). Consequently, a map $(q, p) \mapsto (Q, P)$ is canonical if and only if

$$\{Q^i, Q^j\} = 0, \qquad \{P_i, P_j\} = 0, \qquad \{Q^i, P_j\} = \delta^i_j.$$

The most important property of $\{\ ,\ \}$ shows that it makes the C^∞ functions on M into a Lie algebra:

THE JACOBI IDENTITY. On T^*M we have

$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0.$$

This can be checked by a straightforward, though somewhat involved, calculation, which is often simplified by the following device: When expanded out, this expression is a sum of terms consisting of the product of two first derivatives and a second derivative. The terms containing second derivatives of f will be

$$\{h, \{f, g\}\} + \{g, \{h, f\}\} = (\mathbf{X}_g \mathbf{X}_h - \mathbf{X}_h \mathbf{X}_g)(f) = [\mathbf{X}_g, \mathbf{X}_h](f).$$

But the formula for the Lie bracket of two vector fields involves only first derivatives, so no second derivatives of f appear, and of course the same is true for the second derivatives of g and h.

Jacobi concocted his identity for a very specific reason: it immediately implies a result that Poisson had proved many years earlier by rather complicated computations.

COROLLARY (POISSON'S THEOREM). Consider a Hamiltonian H on T^*M, and functions $f, g\colon T^*M \to \mathbb{R}$. Suppose $\{f_1, H\} = 0$ and $\{f_2, H\} = 0$, so that f_1 and f_2 are each integrals of motion for Hamilton's equations for H. Then also

$$\{\{f_1, f_2\}, H\} = 0,$$

so that $\{f_1, f_2\}$ is also an integral of motion.

(The proof for the case of $T^*M \times \mathbb{R}$ requires a bit more work, cf. Problem 3.)

Unfortunately, attempts to apply Poisson's theorem may turn out to be disappointing, because the newly derived integrals all too often turn out to be 0, or a combination of previously derived integrals. Problem 5 gives the standard elementary example of a case where a new integral is found.

Poisson brackets *bis*. All these considerations can now be applied to arbitrary symplectic manifolds (M, ω) (and extended analogously to $M \times \mathbb{R}$). In fact, Darboux's theorem shows that around every point we can find a symplectic coordinate system (q, p), for which we have $\omega = \sum_{i=1}^n dp_i \wedge dq^i$, and we can then simply repeat all the calculations made for T^*M.

But of course, we're not going to take that easy way out! Aside from the masochistic pleasure that we will derive from doing everything invariantly, we will be able to isolate just where the condition $d\omega = 0$ plays a crucial role, and we can always hope that we'll obtain some enlightenment along the way.

First of all, for a symplectic manifold (M, ω), we define a diffeomorphism $f\colon M \to M$ to be **symplectic** if $f^*\omega = \omega$. Moreover, given $H\colon M \to \mathbb{R}$, we define the vector field $\mathbf{X}_H$ exactly as before,

$$\mathbf{X}_H \lrcorner\, \omega = -dH,$$

and again call vector fields of this form **Hamiltonian vector fields**. Using the Lemma on page 570, we prove, exactly as before, the corresponding

S1. THEOREM. The map $f\colon M \to M$ is symplectic if and only for all H

$$f_*(\mathbf{X}_H) = \mathbf{X}_{f^*H}.$$

In particular, if f is symplectic, then f_* always takes Hamiltonian vector fields into Hamiltonian vector fields.

In addition, the first proof of Theorem 2 on page 576 goes through exactly as before (using the crucial hypothesis $d\omega = 0$) to prove

S2. THEOREM. The flow of any Hamiltonian vector field on the symplectic manifold (M, ω) consists of symplectic maps.

On any symplectic manifold (M, ω), we can use the same definition as before,

$$\{f, g\} = \omega(\mathbf{X}_f, \mathbf{X}_g) = \mathbf{X}_g(f),$$

and for a symplectic coordinate system we can easily compute that $\{f, g\}$ is given by the formula in the classical definition.

The argument on page 607, with Theorem S1 in place of Theorem 1 on page 570, gives

S3. THEOREM. $\phi\colon M \to M$ is symplectic if and only if for all $f, g\colon M \to \mathbb{R}$,

$$\phi^*\{f, g\} = \{\phi^* f, \phi^* g\} \quad \text{or} \quad \{f, g\} \circ \phi^{-1} = \{f \circ \phi^{-1}, g \circ \phi^{-1}\}.$$

As with T^*M, a coordinate system (q, p) is symplectic if and only if we have the fundamental Poisson brackets

$$\{q^i, q^j\} = 0, \qquad \{p_i, p_j\} = 0, \qquad \{q^i, p_j\} = \delta^i_j,$$

and $(q, p) \mapsto (Q, P)$ is symplectic if and only if

$$\{Q^i, Q^j\} = 0, \qquad \{P_i, P_j\} = 0, \qquad \{Q^i, P_j\} = \delta^i_j.$$

We also have a result that plays an important role right now as well as later on.

S4. THEOREM. On any symplectic manifold,

$$[\mathbf{X}_f, \mathbf{X}_g] = -\mathbf{X}_{\{f,g\}}.$$

PROOF. By the two Cartan formulas on page 576 we have

$$\begin{aligned}
[\mathbf{X}_f, \mathbf{X}_g] \lrcorner\, \omega &= L_{\mathbf{X}_f}(\mathbf{X}_g \lrcorner\, \omega) - \mathbf{X}_g \lrcorner (L_X \omega) \\
&= d\big(\mathbf{X}_f \lrcorner (\mathbf{X}_g \lrcorner\, \omega)\big) + \mathbf{X}_f \lrcorner\, d(\mathbf{X}_g \lrcorner\, \omega) + 0 \\
&= d(\omega(\mathbf{X}_f, \mathbf{X}_g)) + 0 = -\mathbf{X}_{\omega(\mathbf{X}_f, \mathbf{X}_g)} \lrcorner\, \omega,
\end{aligned}$$

so

$$[\mathbf{X}_f, \mathbf{X}_g] = -\mathbf{X}_{\omega(\mathbf{X}_f, \mathbf{X}_g)} = -\mathbf{X}_{\{f,g\}}. \quad ❖$$

1. COROLLARY (THE JACOBI IDENTITY). On any symplectic manifold,

$$\{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0.$$

PROOF. We have

$$\begin{aligned}[\mathbf{X}_f, \mathbf{X}_g](h) &= \mathbf{X}_f(\mathbf{X}_g(h)) - \mathbf{X}_g(\mathbf{X}_f(h)) \\ &= \mathbf{X}_f(\{h, g\}) - \mathbf{X}_g(\{h, f\}) = \{\{h, g\}, f\} - \{\{h, f\}, g\},\end{aligned}$$

while Theorem S4 gives

$$[\mathbf{X}_f, \mathbf{X}_g](h) = -\mathbf{X}_{\{f,g\}}(h) = -\{h, \{f, g\}\}. \text{ ❖}$$

2. COROLLARY (POISSON'S THEOREM). If $\{f, H\} = 0$ and $\{g, H\} = 0$, then

$$\{\{f, g\}, H\} = 0.$$

3. COROLLARY. For $f, g, h\colon M \to \mathbb{R}$ we have

$$\mathbf{X}_h(\{f, g\}) = \{\mathbf{X}_h(f), g\} + \{f, \mathbf{X}_h(g)\}$$

(the "infinitesimal" version of Theorem S3, compare Problem 7).

4. COROLLARY (Another generalization of Noether's theorem). Two vector fields $\mathbf{X}_H$ and $\mathbf{X}_K$ on a connected symplectic manifold M commute, that is, $[\mathbf{X}_H, \mathbf{X}_K] = 0$, if and only if $\{H, K\}$ is constant.

PROOF. By Theorem S4, $[\mathbf{X}_H, \mathbf{X}_K] = 0$ is equivalent to $\mathbf{X}_{\{H,K\}} = 0$. Since

$$\mathbf{X}_{\{H,K\}} \lrcorner\, \omega = -d(\{H, K\}),$$

and ω is nondegenerate, $\mathbf{X}_{\{H,K\}} = 0$ is equivalent to $d(\{H, K\}) = 0$. ❖

PROBLEMS

1. For a 1-parameter family $\{X^t\}$ of vector fields on M let $\{\phi_t\}$ be the corresponding 1-parameter family of diffeomorphisms of M, with $\phi_0 =$ identity, that is generated by $\{X^t\}$, so that for $f\colon M \to \mathbb{R}$ we have

$$(X^t f)(p) = \lim_{h\to 0} \frac{f(\phi_{t+h}(p)) - f(\phi_t(p))}{h}.$$

For a family ω_t of k-forms on M we define the k-form

$$\dot{\omega}_t = \lim_{h\to 0} \frac{\omega_{t+h} - \omega_t}{h}.$$

Show that for $\eta(t) = \phi_t{}^*\omega_t$ we have

$$\dot{\eta}_t = \phi_t{}^*(L_{X^t}\omega_t + \dot{\omega}_t),$$

or more informally,

$$\frac{d}{dt}\phi_t{}^*\omega_t = \phi_t{}^*\left(L_{X^t}\omega_t + \frac{d}{dt}\omega_t\right).$$

2. Let (q, p) be a coordinate system on a symplectic manifold (M, ω) satisfying

$$\{q^i, q^j\} = 0, \qquad \{p_i, p_j\} = 0, \qquad \{q^i, p_j\} = \delta^i_j.$$

(a) Letting A be the matrix of ω in the coordinate system (q, p), show that for $B = A^{-1}$ we have

$$\{q^i, q^j\} = b^{ji}.$$

(b) Similarly,

$$\{q^i, p_j\} = -b^{i+n,j}, \qquad \{p_i, p_j\} = b^{j+n,i+n}.$$

(c) Thus

$$A = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} = \mathbb{J}^{-1},$$

and (q, p) is a symplectic coordinate system.

Note that this is just the symplectic analogue of the fact that in a Riemannian manifold, if a coordinate system $(q^1, \dots, q^n)$ satisfies

$$\left\langle \frac{\partial}{\partial q^i}, \frac{\partial}{\partial q^j} \right\rangle = \delta^{ij},$$

then the metric is given by

$$\langle\ ,\ \rangle = \sum_{i=1}^{n} dq^i \otimes dq^j .$$

3. (a) For $f, g\colon M \times \mathbb{R} \to \mathbb{R}$, we have

$$\frac{\partial}{\partial t}\{f, g\} = \left\{\frac{\partial f}{\partial t}, g\right\} + \left\{f, \frac{\partial g}{\partial t}\right\}.$$

(b) If f and g are integrals of motion, then so is $\{f, g\}$.

4. (a) $\mathbf{X}_{gh} = g\mathbf{X}_h + h\mathbf{X}_g$.
(b) $\{f, gh\} = g\{f, h\} + h\{f, g\}$
(c) More generally, for a function $\phi\colon \mathbb{R}^k \to \mathbb{R}$ we have

$$\{\phi \circ (f_1, \ldots, f_k), g\} = \sum_{\alpha=1}^{k} D_\alpha \phi \circ (f_1, \ldots, f_k) \cdot \{f_\alpha, g\},$$

where D_α is the partial derivative with respect to the α^{th} argument.

5. In $\mathbb{R}^3$, calculate

$$\{q^2 p_3 - q^3 p_2, q^3 p_1 - q^1 p_3\}$$

(this expands out into 16 terms, of which only 2 are non-zero; Problem 4 (b) is useful).

Answer: $q^1 p_2 - q^2 p_1$

Conclude that if angular momentum around the first and second axes are conserved for some Hamiltonian, then angular momentum around the third axis is also conserved.

6. If $\{\phi_t\}$ is the flow of $\mathbf{X}_H$, then

$$\frac{d}{dt} f \circ \phi_t = \{f \circ \phi_t, H\} = \{f, H\} \circ \phi_t$$

(sometimes written as $\dot{f} = \{f, H\}$, and called the equation of motion in Poisson bracket form).

7. (a) If $\{\phi_t\}$ is the flow of $\mathbf{X}_H$, show that

$$\left.\frac{d}{dt}\right|_{t=0} \mathbf{X}_{\phi_t{}^* f} \lrcorner\, \omega = -\left.\frac{d}{dt}\right|_{t=0} d(\phi_t{}^* f) = -d(\mathbf{X}_H f) = \mathbf{X}_{\mathbf{X}_H f} \lrcorner\, \omega,$$

so that we have

$$\left.\frac{d}{dt}\right|_{t=0} \mathbf{X}_{\phi_t{}^* f} = \mathbf{X}_{\mathbf{X}_H f}$$

(it may be useful to compare Problem 19-1 (b) for the justification of some steps).
(b) Taking the derivative at 0 of

$$\phi_t{}^*(\{f, g\}) = \{\phi_t{}^* f, \phi_t{}^* g\},$$

which follows from Theorem S3, prove Corollary 3. One could then use this result to prove the Jacobi identity without using Theorem S4, and Theorem S4 itself from the Jacobi identity.

CHAPTER 21

LIOUVILLE INTEGRABILITY

Liouville's second main theorem about Hamiltonian mechanics characterizes a large class of Hamiltonian systems that are "integrable", meaning that they can be solved completely by quadratures (in terms of indefinite integrals, and solving for implicit functions), and it covers virtually all systems with this property that were known to classical mechanics.

Functions in involution. In Chapter 18 we saw that a complete integral for the Hamilton–Jacobi equation for some Hamiltonian on $T^*M \times \mathbb{R}$ allows us to solve Hamilton's canonical equations by quadratures. Moreover, by Corollary 2 on page 603, there is a canonical transformation $(q, p) \mapsto (Q, P)$ such that n curves $Q^\alpha =$ constant or $P_\beta =$ constant are solutions, with the α's disjoint from the β's. The fundamental Poisson brackets on page 607 or 609 show that in this case, any pair from among this particular set of Q^α and P_β are **in involution**, that is, their Poisson bracket is 0.

From this point of view, rather than viewing a case where the Poisson bracket of two integrals f_1 and f_2 satisfies $\{f_1, f_2\} = 0$ as a disappointing aspect of Poisson's theorem, it can instead be considered as a promising circumstance, suggesting that f_1 and f_2 might be part of a symplectic coordinate system in which Hamilton's equations become trivial to solve.

In fulfillment of this promise, Liouville's theorem shows that the existence of any n integrals in involution with each other guarantees that we can solve Hamilton's equations by quadratures. In fact, given n independent functions $f_1, \dots, f_n$ on $T^*M \times \mathbb{R}$ that are in involution with each other, there is a way to find, by quadratures, a symplectic coordinate system (Q, P) with $P_j = f_j$, and if the f_j are all integrals for a Hamiltonian H on $T^*M \times \mathbb{R}$, the Q^i obtained in this way will also be integrals for H, so that having half the maximum number of independent integrals will automatically determine a complementary half, allowing Hamilton's equations to be solved by quadratures.

More generally, and in more detail, consider a symplectic manifold (M, ω) of dimension $2n$, with symplectic coordinates $q^1, \dots, q^n, p_1, \dots, p_n$ around a point $\wp_0$. Suppose the functions $f_1, \dots, f_n \colon M \times \mathbb{R} \to \mathbb{R}$ are in involution with each other, $\{f_j, f_k\} = 0$, and are independent at $(\wp_0, t_0)$, that is, $df_1, \dots, df_n$ are linearly independent in a neighborhood of $(\wp_0, t_0)$. Then for any constants

$a_1, \ldots, a_n$, the set $N \subset M$, defined by $N = \{p : f_j(p, t_0) = a_j \text{ for all } j\}$ is, in a neighborhood of p_0, an n-dimensional submanifold (if it is not empty).

If we consider the vector fields $X_1, \ldots, X_n$ defined by

$$X_j \lrcorner\, \omega = df_j \qquad \text{(in our usual notation, } X_j = -\mathbf{X}_{f_j}\text{)},$$

then the X_j are linearly independent at any $p \in N$, since ω is nonsingular; moreover, the X_j are all tangent to N, since

$$df_k(X_j) = X_j(f_k) = -\mathbf{X}_{f_j}(f_k) = \{f_j, f_k\} = 0.$$

Since the tangent space of N at p is thus spanned by $X_1, \ldots, X_n$ at p and $\omega(X_j, X_k) = \{f_j, f_k\} = 0$, we have the

ISOTROPY LEMMA. Each tangent space of N is isotropic: ω restricted to the tangent space is zero.

Our assumption concerning independence of $(f_1, \ldots, f_n)$ means that the $n \times 2n$ matrix

$$\frac{\partial f_k}{\partial(q^i, p_j)}$$

has rank n at (p_0, t_0). It follows that for p in a neighborhood of p_0, to which we now restrict our attention, among the q^i, p_j there are n coordinates

$$r_1, \ldots, r_n \quad \text{with} \quad 0 \neq \det \frac{\partial f_k}{\partial r_j}(p, t_0).$$

From the isotropy lemma and the theorem on page 602 we conclude that we can choose the coordinates $r_1, \ldots, r_n$ so that

(a) the indices for the q's and those for the p's are disjoint.

Now the map $(q^i, p_i) \mapsto (p_i, -q^i)$ is canonical, as pointed out on page 571. We can make this switch for any q^i in our collection, and (a) insures that we do not get any duplications of the p_j. Thus, simply by renaming coordinates, we can assume that in fact

$$0 \neq \det \frac{\partial f_k}{\partial p_j}(p_0, t_0).$$

These rather non-classical considerations will serve as preliminaries for a quite classic proof of Liouville's theorem, though we state it more generally in terms of symplectic manifolds. The modern (re)formulation begins immediately afterward, on page 620, and it will be quite interesting to see how various steps in the classic proof correlate with steps in the modern treatment.

THEOREM (LIOUVILLE). Let (M, ω) be a symplectic manifold of dimension $2n$, and let $f_1, \dots, f_n \colon M \times \mathbb{R} \to \mathbb{R}$ be functions that are in involution with each other, and such that $df_1, \dots, df_n$ are linearly independent at $(\not p, t_0)$.

I. Then there are functions $Q^1 \dots, Q^n \colon M \times \mathbb{R} \to \mathbb{R}$ in a neighborhood of $(\not p, t_0)$ such that the coordinate system $(Q^1, \dots, Q^n, f_1, \dots, f_n, t)$ is symplectic, and $(Q^1, \dots, Q^n)$ can be found by quadratures.

II. Moreover, if the f_j are integrals of a Hamiltonian H on $M \times \mathbb{R}$, our method of choosing the Q^i will ensure that they are also integrals.

PROOF. We work locally, in a neighborhood of $(\not p, t_0)$, identifying M with $\mathbb{R}^n \times \mathbb{R}^n$, and, as on page 615, renaming the coordinates (q, p) on M so that we have

$$(*) \qquad 0 \neq \det \frac{\partial f_k}{\partial p_j}(\not p, t_0).$$

I. Condition $(*)$ implies, by the implicit function theorem, that there are functions $\psi_r \colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ with

$$(1) \qquad f_j(q, \psi(q, p, t), t) = p_j, \qquad \psi = (\psi_1, \dots, \psi_n);$$

note that we then also have

$$(2) \qquad \delta_j^k = \frac{\partial p_j}{\partial p_k} = \sum_{l=1}^n \frac{\partial f_j}{\partial p_l}\frac{\partial \psi_l}{\partial p_k} \implies \left(\frac{\partial \psi_l}{\partial p_k}\right) \text{ is nonsingular.}$$

Taking $\partial/\partial q^s$ of equation (1) gives

$$\frac{\partial f_j}{\partial q^s} + \sum_{r=1}^n \frac{\partial f_j}{\partial p_r}\frac{\partial \psi_r}{\partial q^s} = 0,$$

which we can write in terms of matrices as

$$A + BC = 0 \qquad \text{for } A = \left(\frac{\partial f_i}{\partial q^s}\right), \quad B = \left(\frac{\partial f_i}{\partial p_r}\right), \quad C = \left(\frac{\partial \psi_r}{\partial q^s}\right).$$

The hypothesis

$$0 = \{f_j, f_k\} = \sum_{s=1}^n \frac{\partial f_j}{\partial q^s}\frac{\partial f_k}{\partial p_s} - \frac{\partial f_j}{\partial p_s}\frac{\partial f_k}{\partial q^s}$$

can be written $AB^{\mathbf{t}} = BA^{\mathbf{t}}$, hence $B^{-1}A = A^{\mathbf{t}}(B^{\mathbf{t}})^{-1}$, and thus $C = C^{\mathbf{t}}$, i.e.,

$$(3) \qquad \frac{\partial \psi_r}{\partial q^s} = \frac{\partial \psi_s}{\partial q^r} \qquad \text{for all } r, s.$$

[We have thus shown that $\sum_{i=1}^{n} \psi_i \, dq^i$ is closed (as a function only of the q^i), by an argument similar to that used for the isotropy lemma.] Finally, (3) implies that there is a function $S\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ such that

$$\psi_i(q,p,t) = \frac{\partial S}{\partial q^i}(q,p,t), \tag{4}$$

where S is can be written as

$$S(q,p,t) = \int_{\mathring{q},\mathring{p},t_0)}^{(q,p,t)} \psi_1(\chi,p,t)\,d\chi^1 + \cdots + \psi_n(\chi,p,t)\,d\chi^n, \tag{5}$$

the integral being taken along any path from a fixed $(\mathring{q},\mathring{p},t_0)$ to (q,p,t).

We have $\partial^2 S/\partial p_j \partial q^i = \partial/\partial p_j \big(\partial S/\partial q^i\big) = \partial\psi_i/\partial p_j$, and $\big(\partial\psi_i/\partial p_j\big)$ is nonsingular by (2), so

$$\left(\frac{\partial^2 S}{\partial p_j \, \partial q^i}\right) \qquad \text{is nonsingular.}$$

We can thus use S to obtain a type 2 generating function, as on the bottom of page 580: We first find n functions $P_i\colon \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ satisfying

$$p_i = \frac{\partial S}{\partial q^i}(q, P(q,p,t), t), \qquad P = (P_1,\ldots,P_n), \tag{6}$$

and then define the Q^i appropriately [the actual definition isn't important here, just the fact that S is a generating function for (Q,P)].

In view of (4), we can write (6) as

$$p_i = \psi_i(q, P(q,p,t), t), \tag{7}$$

and thus

$$f_j(q,p,t) = f_j\big(q, \psi\big(q, P(q,p,t),t\big),t\big) = P_j(q,p,t) \qquad \text{by (1)}, \tag{8}$$

proving the first part of the theorem.

II. Now suppose, in addition, that the f_i are integrals for a Hamiltonian H on $M \times \mathbb{R}$. Substituting (8) back into (7) gives

$$p_r = \psi_r(q, f(q,p,t), t), \qquad f = (f_1,\ldots,f_n), \tag{9}$$

and taking the partial derivative with respect to t, q^s, and p_s then gives us the following equations [all partial derivatives of ψ_r evaluated at $(q, f(q,p,t),t)$]:

$$(10)\qquad 0 = \frac{\partial\psi_r}{\partial t} + \sum_{i=1}^{n} \frac{\partial\psi_r}{\partial p_i}\frac{\partial f_i}{\partial t},$$

$$(11)\qquad 0 = \frac{\partial\psi_r}{\partial q^s} + \sum_{i=1}^{n} \frac{\partial\psi_r}{\partial p_i}\frac{\partial f_i}{\partial q^s},$$

$$(12)\qquad \delta_r^s = \sum_{i=1}^{n} \frac{\partial\psi_r}{\partial p_i}\frac{\partial f_i}{\partial p_s}.$$

But the f_j are integrals, so (page 606) we have

$$\frac{\partial f_j}{\partial t} + \{f_j, H\} = 0,$$

and substituting into (10) we get

$$(13)\qquad \frac{\partial\psi_r}{\partial t} = \sum_{j=1}^{n} \frac{\partial\psi_r}{\partial p_j} \sum_{s=1}^{n} \left(\frac{\partial f_j}{\partial q^s}\frac{\partial H}{\partial p_s} - \frac{\partial H}{\partial q^s}\frac{\partial f_j}{\partial p_s} \right).$$

Together with (11) and (12), and remembering (3), this gives

$$(14)\qquad \frac{\partial\psi_r}{\partial t} = -\left(\frac{\partial H}{\partial q^r} + \sum_{s=1}^{n} \frac{\partial H}{\partial p_s}\frac{\partial\psi_s}{\partial q^r} \right).$$

Now consider a new Hamiltonian G defined by

$$G(q,p,t) = H(q, \psi(q,p,t), t).$$

Taking (9) into account, we see that

$$\begin{aligned} \frac{\partial G}{\partial q^r}(q, f(q,p,t),t) &= \frac{\partial H}{\partial q^r}(q,p,t) + \sum_{s=1}^{n} \frac{\partial H}{\partial p_s}(q,p,t)\frac{\partial\psi_s}{\partial q^r}(q, f(q,p,t),t) \\ &= -\frac{\partial\psi_r}{\partial t}(q, f(q,p,t),t) \quad \text{by (14)}, \end{aligned}$$

and thus, in general,

$$(15)\qquad \frac{\partial G}{\partial q^r} = -\frac{\partial\psi_r}{\partial t}.$$

But then (5) gives [adding a constant to G if necessary]

$$\frac{\partial S}{\partial t} = -\int \sum_{s=1}^{n} \frac{\partial G}{\partial q^r}\, dq^r = -\int dG = -G \qquad \text{so that} \qquad G + \frac{\partial S}{\partial t} = 0.$$

Equation ($\widetilde{\mathrm{F}}_2$-c) on page 584 then shows that the Hamiltonian K for G in the (Q, P) coordinates is $K = 0$. This, in turn, means that for any solution c of Hamilton's equations for G, we have $Q \circ c = \text{constant}$.

On the other hand, if γ is a solution of Hamilton's equations for the original Hamiltonian H, and we let

$$\begin{aligned} c(t) &= \gamma(q(\gamma(t)), f(q(\gamma(t)), p(\gamma(t)), t), t) \\ &= \gamma(q(\gamma(t)), a_1, \dots, a_n, t) \quad \text{for certain constants } a_i, \end{aligned}$$

then we find that c satisfies Hamilton's equations for G, so

$$Q^i(\gamma(t), t) = Q^i(q(\gamma(t)), a_1, \dots, a_n, t) = Q^i(c(t), t) = \text{constant.} \quad ❖$$

When the Hamiltonian H and the functions $f_1, \dots, f_n$ do not depend on t, the approach on pages 587–588 works out as follows. Equation (10) is now superfluous, and since the f_j are integrals, we have

$$0 = \{f_j, H\} = \sum_{s=1}^{n} \left(\frac{\partial f_j}{\partial q^s} \frac{\partial H}{\partial p_s} - \frac{\partial H}{\partial q^s} \frac{\partial f_j}{\partial p_s} \right),$$

from which we obtain the special case of (13),

$$0 = \sum_{j=1}^{n} \frac{\partial \psi_r}{\partial p_j} \sum_{s=1}^{n} \left(\frac{\partial f_j}{\partial q^s} \frac{\partial H}{\partial p_s} - \frac{\partial H}{\partial q^s} \frac{\partial f_j}{\partial p_s} \right),$$

and thus, using (11), (12) and (3), the special case of (14),

$$0 = -\left(\frac{\partial H}{\partial q^r} + \sum_{s=1}^{n} \frac{\partial H}{\partial p_s} \frac{\partial \psi_s}{\partial q^r} \right).$$

For the Hamiltonian G we then find the special case of (15),

$$\frac{\partial G}{\partial q^r} = 0,$$

so that we have

$$H\left(q, \frac{\partial S(q, P)}{\partial q}\right) = H(q, \psi(q, p)) = K(P),$$

exactly as in the case of the equation $(*)$ on page 587. So in this case we simply conclude that under the canonical transformation obtained in part I of the proof, the solution curves γ have all $P_j(\gamma(t)) = \alpha_j$ for certain constants α_j, while the $Q^i(\gamma(t))$ are of the form

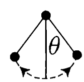
$$Q^i(\gamma(t)) = Q^i(\gamma(0)) + t\nu_i(\alpha_1, \dots, \alpha_n)$$

for certain functions ν_i.

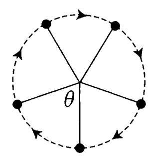
We've noted this special case, and its specific outcome, because nowadays "Liouville's theorem" usually refers to a result involving an additional important feature of mechanics problems, and the original result has become but one strand of a geometric tapestry in which analytic calculations are largely replaced by geometric constructions. This additional feature involves Hamiltonians that are time-independent, which will be reflected in the whole geometric approach.

Conditional periodicity and the invariant tori. One of the very first problems of mechanics that we considered was motion under an inverse square force, with the solutions neatly separating into two classes, the unbounded orbits, parabolas and hyperbolas, and the bounded elliptical orbits, which are periodic.

The pendulum is an even simpler situation where we encounter periodic motion. It should be pointed out that the pendulum doesn't have to oscillate, as in (a); it can also rotate, periodically, about the pivot point (b) if started with a



(a)

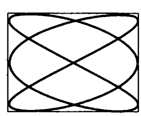


(b)

large enough initial velocity. In addition, the pendulum can simply hang straight down, or, as on page 214, remain vertically above the pivot, when our "string" is a massless thin rigid rod (or we are dealing with a "compound" or "physical" pendulum). As we will point out later on, there are also two non-closed solution paths, currently hiding out, teasing you to discover them.

The equation for harmonic oscillation $x'' + \omega^2 x = 0$ that one gets for springs or simple electronic circuits again gives periodic solutions. However, when we consider the 2-dimensional case (pages 293–295), with the x_1-coordinate and the x_2-coordinate separately satisfying

$$x_1'' + {\omega_1}^2 x_1 = 0,$$
$$x_2'' + {\omega_2}^2 x_2 = 0,$$



we obtain Lissajous figures, which can be periodic, but need not be, even though the components do exhibit periodic motion, since the periods might not be commensurable. This is called "conditionally periodic" motion [only under special conditions does the motion become periodic].

Another example of such conditionally periodic motion is given by the spherical pendulum on pages 290–291. Here the pendulum rotates around the z-axis as ϕ goes from 0 to 2π, while θ oscillates between $\arccos u_1$ and $\arccos u_2$.

Conditional periodicity also occurs for the orbits discussed on pages 128–129 (in fact, it was astronomers who first introduced the term), and the herpolhode of Chapter 9.

For a 3-dimensional example, recall (page 349) that the motion of a heavy top is generally determined by three periodic motions, the rotation about the axis, the nutation, and the precession ϕ, which need not have commensurable periods, so that we again have a conditionally periodic motion.

Finally, page 478 gives an n-dimensional example of conditional periodicity.

When a periodic motion is not an oscillation, physicists often like to describe the motion in terms of "multiple-valued functions". For example, when a pendulum doesn't oscillate, but instead rotates about the pivot point, θ may be thought of as taking on values in $[0, 2\pi)$, followed by values in $[2\pi, 4\pi)$, etc. For the conical pendulum, ϕ is similarly multiple-valued; and of course for the heavy top, the rotation about the axis and the precession ϕ can be regarded as multiple-valued functions (while the nutation is an oscillation).

Mathematicians also employ the lingo of multiple-valued functions, but they have the luxury of alternately describing these same phenomena geometrically.

INVARIANT TORI THEOREM (ARNOLD). Let (M, ω) be a symplectic manifold of dimension $2n$, and let $f_1, f_2, \ldots, f_n \colon M \to \mathbb{R}$ be n functions that are in involution with each other. For some n-tuple of constants $\mathbf{a} = (a_1, \ldots, a_n)$, let

$$M_{\mathbf{a}} = \{x \in M : f_i(x) = a_i\} \quad \text{i.e.,} \quad M_{\mathbf{a}} = \mathbf{f}^{-1}(\mathbf{a}) \ \text{ for } \mathbf{f} = (f_1, \ldots, f_n),$$

and suppose that the df_i are linearly independent at each point of $M_{\mathbf{a}}$ (that is, $\mathbf{f}_*$ has maximal rank n at each point), so that $M_{\mathbf{a}}$ is an n-dimensional submanifold.

Then each (non-empty) compact component $C_{\mathbf{a}}$ of $M_{\mathbf{a}}$ is diffeomorphic to an n-torus $S^1 \times \cdots \times S^1$.

Moreover, if the f_i are in involution with some Hamiltonian $H \colon M \to \mathbb{R}$, then the solutions for Hamilton's equations (the solution curves γ of $\mathbf{X}_H$) take $C_{\mathbf{a}}$ into itself, and $C_{\mathbf{a}}$ has coordinates $(\varphi^1, \ldots, \varphi^n)$ such that these γ satisfy

$$\varphi^i(\gamma(t)) = \varphi^i(\gamma(0)) + \nu_i(\mathbf{a}) \cdot t$$

for constants $\nu_i(\mathbf{a})$, the [**circular**] **frequencies** for $C_{\mathbf{a}}$ (the usual frequencies will then be $\nu_i/2\pi$, and the periods, the reciprocals of the frequencies, are $2\pi/\nu_i$).

Note: The theorem involves circular frequencies because we will choose the φ^i to repeat on intervals of length 2π, rather than length 1. This convention is not universally used, which one has to bear in mind for the various formulas that will occur starting on page 631.

PROOF. As before, we consider the vector fields $X_1, \dots, X_n$ defined by

$$X_j \lrcorner\, \omega = df_j \qquad (X_j = -\mathbf{X}_{f_j}).$$

We have already seen that they are linearly independent everywhere on $M_{\mathbf{a}}$, and everywhere tangent to $M_{\mathbf{a}}$.

In addition, the vector fields X_j commute, that is, $[X_j, X_k] = 0$ for all j, k, since Theorem S4 on page 609 gives

$$[X_j, X_k] = [\mathbf{X}_{f_j}, \mathbf{X}_{f_k}] = -\mathbf{X}_{\{f_j, f_k\}} = 0.$$

Since $C_{\mathbf{a}}$ is compact, the flow $\{\rho_j{}^t\}$ of X_j on $C_{\mathbf{a}}$ is defined for all t, and since the X_j commute, the $\{\rho_j{}^t\}$ and $\{\rho_k{}^s\}$ commute. For $\mathbf{t} = (t_1, \dots, t_n) \in \mathbb{R}^n$ define

$$\rho^{\mathbf{t}} \colon C_{\mathbf{a}} \to C_{\mathbf{a}} \quad \text{by} \quad \rho^{\mathbf{t}} = \rho^{(t_1, \dots, t_n)} = \rho_1{}^{t_1} \circ \dots \circ \rho_n{}^{t_n}.$$

Since the $\rho_j{}^t$ and $\rho_k{}^s$ commute, we have $\rho^{\mathbf{t}+\mathbf{s}} = \rho^{\mathbf{t}} \circ \rho^{\mathbf{s}}$ for all $\mathbf{s}, \mathbf{t} \in \mathbb{R}^n$.

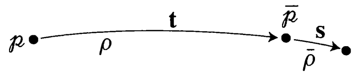
Choosing some fixed point $\mathscr{p} \in C_{\mathbf{a}}$, we can define

$$\rho \colon \mathbb{R}^n \to C_{\mathbf{a}} \qquad \text{by} \qquad \rho(\mathbf{t}) = \rho^{\mathbf{t}}(\mathscr{p}),$$

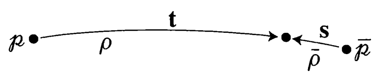
so that $\mathscr{p}$ is moved along the first flow for time t_1, then along the second flow for time t_2, etc. Commutativity of the flows implies that $\rho(\mathbf{t} + \mathbf{s}) = \rho(\mathbf{t}) \circ \rho(\mathbf{s})$.

Since the X_i are linearly independent at $\mathscr{p}$, the implicit function theorem shows that a neighborhood of $0 \in \mathbb{R}^n$ will be taken by ρ onto a neighborhood of $\mathscr{p}$ in $C_{\mathbf{a}}$. Moreover, if $\bar{\mathscr{p}} = \rho(\mathbf{t})$, and $\bar{\rho}$ is the map defined like ρ, but using $\bar{\mathscr{p}}$ instead of $\mathscr{p}$, then we have

$$\rho(\mathbf{t} + \mathbf{s}) = \bar{\rho}(\mathbf{s}),$$



which shows that in fact the image $\rho(\mathbb{R}^n)$ is an open subset of $C_{\mathbf{a}}$. By the same token, it also clearly a closed subset, for if $\bar{\mathscr{p}}$ is in the closure of the image,



then the image of its $\bar{\rho}$ contains a point $\bar{\rho}(\mathbf{s}) = \rho(\mathbf{t})$ and then $\bar{\mathscr{p}} = \rho(\mathbf{t} + \mathbf{s})$. It thus follows that the image of $\rho \colon \mathbb{R}^n \to C_{\mathbf{a}}$ is all of $C_{\mathbf{a}}$.

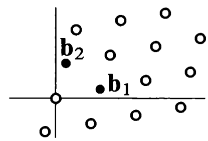
It is easy to see that the set

$$G = \{\mathbf{t} \in \mathbb{R}^n : \rho(\mathbf{t}) = \mathscr{p}\}$$

is a subgroup of $\mathbb{R}^n$, and moreover we get the same subgroup if we start with a different $\mathscr{p}$ (and correspondingly different ρ). Finally, the fact that a neighborhood of $0 \in \mathbb{R}^n$ is taken onto a neighborhood of $\mathscr{p}$ [no matter which $\mathscr{p}$ we choose] shows that G is a *discrete* subgroup of $\mathbb{R}^n$.

An easy exercise (Problem 1) shows that for some $k \le n$ there are linearly independent elements $\mathbf{b}_1, \dots, \mathbf{b}_k$ of $\mathbb{R}^n$ such that G is precisely the set of *integer* linear combinations of the $\mathbf{b}_i$,

$$G = \{m^1\mathbf{b}_1 + \cdots + m^k\mathbf{b}_k : m^1, \ldots, m^k \in \mathbb{Z}\}.$$



Then $C_{\mathbf{a}}$ is isomorphic, as a quotient group, and diffeomorphic, as a quotient manifold, to $\mathbb{R}^n/G$. Since $C_{\mathbf{a}}$ is compact, we must have $k = n$, so that $C_{\mathbf{a}}$ is diffeomorphic to the n-torus $S^1 \times \cdots \times S^1$, with universal covering space $\mathbb{R}^n$ under the map $\rho\colon \mathbb{R}^n \to C_{\mathbf{a}}$.[1]

Every $p \in C_{\mathbf{a}}$ is $\rho(s^1\mathbf{b}_1 + \cdots + s^n\mathbf{b}_n)$ for unique $0 \le s^i < 1$, and letting

$$\varphi^i(p) = 2\pi s^i$$

we obtain "angular coordinates" $\varphi^1, \ldots, \varphi^n$ on $C_{\mathbf{a}}$, which we can extend to multiple-valued coordinates mod 2π on $C_{\mathbf{a}}$ (like the standard use of θ on S^1). We will reserve φ [as opposed to ϕ] for these coordinates.

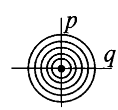
The final statement of the theorem follows from the fact that the f's are constants along the integral curves of $\mathbf{X}_H$. ❖

Remark. The φ's are "dual" to the X's, with $\partial/\partial\varphi^i$ being constant linear combinations of the X_j on $C_{\mathbf{a}}$.

Various names for these tori are: the invariant tori, the Liouville tori, the Arnold tori, or the Liouville-Arnold tori (remarks at the end of the chapter may give some elucidation).

♦ As the simplest, 1-dimensional, example of the theorem, we consider the harmonic oscillator, whose Hamiltonian was found in Problem 17-1 (a). We set $m = 1$ for simplicity, and to begin with, we consider $\omega = 1$, so that we are working with the equation $\ddot{q} + q = 0$, and $H = \frac{1}{2}(q^2 + p^2)$. In this 1-dimensional example, our manifold M is now simply $T^*\mathbb{R} = \mathbb{R}^2$ with coordinates q and p, and we simply choose $f_1 = H$. The various $M_a = C_a$ are now the sets of constant energy a,

$$\{(q, p) : q^2 + p^2 = 2a\}.$$



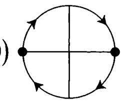
At the origin $(0, 0)$ we have $dH = 0$ and M_0 is just a point, while all other M_a are 1-dimensional tori, the circles of radii $\sqrt{2a}$.

---

[1] More generally, if we don't assume $C_{\mathbf{a}}$ is compact, but assume that the X_i are complete, we can conclude that $C_{\mathbf{a}}$ is diffeomorphic to the product of $\mathbb{R}^{n-k}$ and a k-torus.

The collection of these circles, together with the origin, is called the *phase portrait* for this Hamiltonian. Rather than showing how the position q of the particle varies with time, the phase portrait shows how p varies with q. Since $\mathbf{X}_H(q,p) = (p,-q)$, a typical solution curve, lying on some M_a, is

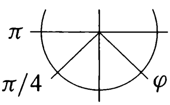
$$\begin{aligned} t &\mapsto (q(t), p(t)) \\ &= \left(\sqrt{2a}\cos t,\ -\sqrt{2a}\sin t\right); \end{aligned}$$



As t goes from 0 to π, with q going from $\sqrt{2a}$ to $-\sqrt{2a}$, the corresponding point (q,p) moves from right to left along the bottom of the circle, where $p \le 0$, while as t goes from π to 2π, with q going from $-\sqrt{2a}$ back to $\sqrt{2a}$, the point (q,p) moves from left to right along the top of the circle, where $p \ge 0$.

In terms of our proof of the theorem, a flow of $\mathbf{X}_H$ in M_a (a trajectory $\gamma\colon \mathbb{R} \to M_a$) makes $\mathbb{R}$ into a covering space of M_a. [Since γ goes clockwise, choosing $\mathbf{b}_1 = 1$ makes the corresponding φ coordinate equal $-\theta$ for the usual polar coordinate θ; choosing $\mathbf{b}_1 = -1$ would give the standard picture.] On M_a the trajectory γ simply has the equation

$$\begin{aligned} \varphi(\gamma(t)) &= \varphi(\gamma(0)) + 1\cdot t, \quad \text{or} \\ \varphi(t) &= \varphi(0) + 1\cdot t \qquad (\text{“condensed”}). \end{aligned}$$



Note that the motion along the image of γ is an immersion; while the particle oscillates back and forth along the line, with velocity 0 at some points, the corresponding curve in the phase portrait simply moves steadily around the circle (but q can't serve as a coordinate system at points where the velocity is 0).

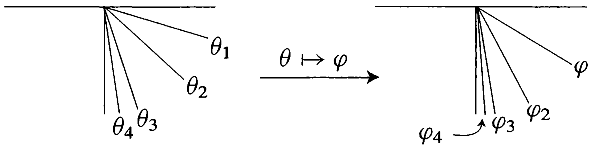
♦ If we consider the more general equation $\ddot{q} + \omega^2 q = 0$, with Hamiltonian $H = \frac{1}{2}(\omega^2 q^2 + p^2)$, we will have ellipses rather than circles, and a typical solution curve is now

$$t \mapsto \left(\frac{\sqrt{2a}}{\omega}\cdot\cos\omega t,\ -\sqrt{2a}\sin\omega t\right),$$

so a complete trajectory is traversed on $[0, 2\pi/\omega]$. This means that in order for φ to be a multiple-valued function mod 2π, we have to let $\varphi = \omega\cdot t$ along a trajectory [in terms of our proof of the theorem, the basis $\mathbf{b}_1$ of G now has length $1/\omega$ rather than 1], and the trajectories γ are thus of the form

$$\begin{aligned} \varphi(t) &= \varphi(0) + \omega\cdot t, \\ \text{i.e.,}\quad \varphi(\gamma(t)) &= \varphi(\gamma(0)) + \omega\cdot t. \end{aligned}$$

Moreover, the line $\varphi = k$ no longer corresponds to the line $\theta = -k$. For each φ, we are at the point $\left(\sqrt{2a}/\omega \cos\varphi, -\sqrt{2a}\sin\varphi\right)$, with slope $-\omega\tan\varphi$, rather than $-\tan\varphi$, so all the slopes are multiplied by ω.



$\varphi = \arctan(\omega\tan\theta)$

♦ The pendulum provides a more illuminating 1-dimensional example. For a pendulum of length l (and a bob of mass 1), the Lagrangian is

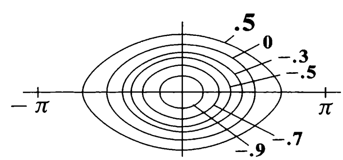
$$L(\theta, \dot\theta) = \tfrac{1}{2} l^2 \dot\theta^2 + gl\cos\theta,$$

where we take the potential to be $V = -gl\cos\theta$, varying from gl at $\theta = \pi$ to $-gl$ at $\theta = 0$.

With θ now denoted by q, the Hamiltonian is

$$H(q,p) = \frac{p^2}{2l^2} - gl\cos q.$$

Consider first the curves $H = E$ for small E; in the figure below, the numbers show the value of E/gl on the curve to which they point. The important difference in this case is that trajectories along the various curves take different times

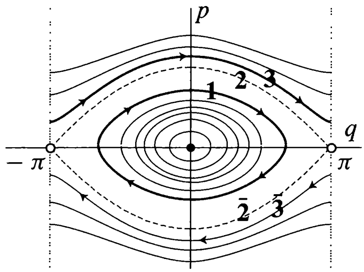


to complete one cycle (the pendulum isn't isochronous), so the frequencies $\nu(E)$ vary from curve to curve, and the set of points $\varphi =$ constant for the different M_a no longer lie along a straight line.

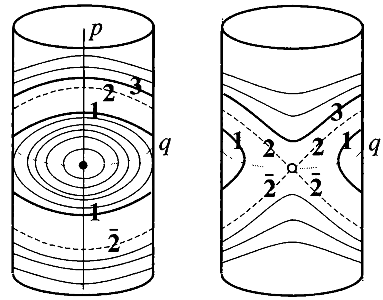
For the complete picture, we need to consider not only the oscillating solutions, but also those that revolve completely around the pivot point. The complete phase space, made up of the curves

$$H = E,$$
$$\text{or} \quad p = \pm\sqrt{2}l\sqrt{E + gl\cos q},$$

is shown below, first as $[-\pi, \pi] \times \mathbb{R}$ and then, more properly, as $S^1 \times \mathbb{R}$. The curves like **1**, where $|E| < gl$, already shown in the figure on the previous page, represent oscillations (or "librations", à la the astronomers). Curves like **3** (also circles), where $|E| > gl$, represent trajectories where the pendulum swings



$[-\pi, \pi] \times \mathbb{R}$

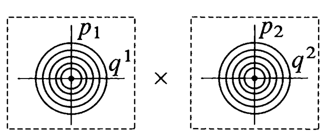


front $S^1 \times \mathbb{R}$ back

completely around ("rotations"); the rotation in the opposite direction is now split into the separate curve $\bar{\mathbf{3}}$. In terms of the universal covering space map $\rho\colon \mathbb{R} \to S^1$, a closed trajectory $\gamma\colon [a, b] \to T^*M$ is a **libration** when there is a covering map $\tilde{\gamma}\colon [a, b] \to \mathbb{R}$, with $\rho \circ \tilde{\gamma} = \gamma$, for which $\tilde{\gamma}$ is also a closed curve, while it is a **rotation** when any covering map $\tilde{\gamma}$ always has $\tilde{\gamma}(b)$ in a different sheet of the covering space from $\tilde{\gamma}(a)$.

In addition to these two families of circles, and the single point $(0, 0)$, the stable equilibrium point where the pendulum hangs straight down, there is the dashed "separatrix", where $E = gl$, between the two families of curves. This looks like a curve that crosses itself, but it is actually three different curves. The first is the single point $\circ$, represented on the left by $(\pi, 0)$ and/or $(-\pi, 0)$, the unstable equilibrium point where the pendulum stays straight up. The other two pieces, **2** and $\bar{\mathbf{2}}$ are the non-compact solution paths that approach this point asymptotically, swinging either clockwise or counterclockwise. With a physical pendulum one might try to obtain (a large portion of) these paths by starting with velocity zero extremely close to the upright point, or try to obtain half the path by starting at the bottom with kinetic energy extremely close to $2gl$, though friction and the required closeness make this pretty futile.

♦ For a particle acted upon by the force $(-\omega_1{}^2 q_1, -\omega_2{}^2 q_2)$, giving independent harmonic oscillations in two directions, T^*M is the product of the phase spaces for two harmonic oscillators, and we can choose $f_i = \frac{1}{2}(\omega_i{}^2 q_i{}^2 + p_i{}^2)$.

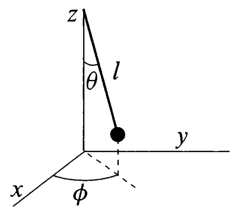


Each 2-dimensional torus corresponds to a pair of circles, one in the phase space of the first factor, one in the phase space of the second factor, and in terms of the corresponding two coordinates φ^1, φ^2, the equation for a trajectory γ will be, in typical condensed notation, the conditionally periodic motion

$$\begin{aligned}\varphi^1(t) &= \varphi^1(0) + \omega_1 \cdot t\\ \varphi^2(t) &= \varphi^2(0) + \omega_2 \cdot t.\end{aligned}$$

Note that although each torus is compact, only the trajectories on this torus that are actually periodic will have compact images, while all others will be dense in that torus.

♦ For a more interesting 2-dimensional example, consider the spherical pendulum of Problem 3-5 and Chapter 8, where we use the coordinates ϕ, θ on $M = S^2 - \{\text{lowest point}\}$ (taking note of the remark on page 479). Choosing the mass $m = 1$ for simplicity, so that the Lagrangian is



$$L = \tfrac{1}{2} l^2(\dot\theta^2 + \sin^2\theta\,\dot\phi^2) + gl\cos\theta,$$

we find that the Hamiltonian is

$$H = \frac{p_\theta{}^2}{2l^2} + \frac{p_\phi{}^2}{2l^2\sin^2\theta} - gl\cos\theta.$$

Along any trajectory we have

$$\text{(a)} \qquad p_\theta = \frac{\partial L}{\partial\dot\theta} = l^2\dot\theta, \qquad p_\phi = \frac{\partial L}{\partial\dot\phi} = l^2\sin^2\theta\,\dot\phi,$$

and p_ϕ is a constant.

We will take $f_1 = H$ and $f_2 = p_\phi$. Since the partial derivatives of f_1 and f_2 with respect to θ, ϕ, p_θ, and p_ϕ are

	$\partial/\partial\theta$	$\partial/\partial\phi$	$\partial/\partial p_\theta$	$\partial/\partial p_\phi$
f_1:	$-\dfrac{p_\phi{}^2\cos\theta}{l^2\sin^3\theta} + gl\sin\theta$	0	$\dfrac{p_\theta}{l}$	$0,$
f_2:	0	0	0	$1,$

the functions f_1 and f_2 can be linearly dependent only at $(\theta, \phi, 0, p_\phi)$ with

$$\frac{p_\phi{}^2\cos\theta}{l^2\sin^3\theta} = gl\sin\theta.$$

Substituting for p_ϕ from equation (a) we find that

$$\dot{\phi} = \sqrt{\frac{g}{l\cos\theta}},$$

and comparing with Problem 1-20, we see that the 1-parameter family of circular motions, at a constant angle θ from the vertical, corresponds to the set of $(f_1, f_2)^{-1}(E_\alpha, \alpha)$ for such linearly dependent f_1, f_2, where E_α is the energy corresponding to angular velocity α.

The inverse images of all other (f_1, f_2) are 2-dimensional, with the compact ones being 2-dimensional tori, where the coordinates φ^1 and φ^2 of a trajectory are periodic, with the motion itself exhibiting conditionally periodic motion.

The family of planar oscillations are excluded from this analysis because our coordinates exclude the lowest point for the pendulum bob; even if they didn't, since $p_\phi = 0$ in these cases, H and p_ϕ would not be linearly independent.

♦ Finally, consider the 3-dimensional example of the heavy top, for which T and V are given on page 445. Now M is SO(3), with the Euler angles (ϕ, θ, ψ) as coordinates. In addition to $f_1 = H$, we have the constants of motion on page 346, namely $f_2 = \langle \mathbf{L}, \mathbf{e}_z \rangle$, the component of the angular momentum $\mathbf{L}$ along the vertical z axis, and $f_3 = \langle \mathbf{L}, \mathbf{e}_Z \rangle$, the component along $\mathbf{e}_Z$.

Checking that f_2 and f_3 are in involution is left as an exercise for the reader. It might also be fun to determine just when f_1, f_2, f_3 are not linearly independent, and correlate these cases, as well as the cases where $C_\mathbf{a}$ is not compact, with the various special types of motions of the top described in Chapter 9.

We now want to reconsider, and essentially reprove, Liouville's theorem in the context of this geometric picture. Note that the invariant tori theorem involves a single torus $C_\mathbf{a}$, and thus, so far we have only been studying individual tori, rather than a neighborhood of one of them. In our figures for the 1-dimensional cases on pages 623 and 625–626, each compact $C_\mathbf{a}$ has a neighborhood made up of compact $C_\mathbf{b}$ for $\mathbf{b}$ near $\mathbf{a}$, and looks like the product of an open interval and $C_\mathbf{a}$. Our first order of business is to prove this generally, in all dimensions.

PRODUCT NEIGHBORHOOD LEMMA. Under the hypotheses for the invariant tori theorem, there is an open disc $D \subset \mathbb{R}^n$, and a neighborhood U of $\mathbf{C_a}$ that is diffeomorphic to the product $D \times C_\mathbf{a}$, under a diffeomorphism $\Phi\colon D \times C_\mathbf{a} \to U$ for which $\mathbf{f} \circ \Phi$ is the projection pr_1 on the first factor, every $\mathrm{pr}_1{}^{-1}(x)$ being equal to some $C_\mathbf{b}$.

$$\begin{tikzcd} D \times C_\mathbf{a} \arrow[rr, "\Phi"] \arrow[dr, "\mathrm{pr}_1"'] & & U \arrow[dl, "\mathbf{f} = (f_1, \dots, f_n)"] \\ & D & \end{tikzcd}$$

PROOF (Simple but messy). We begin with some preliminary observations that are needed to take care of a few details.

1. We claim that there is an open set $V \supset C_{\mathbf{a}}$ such that any $C_{\mathbf{b}} \subset V$ will be compact. In fact, using compactness of $C_{\mathbf{a}}$ and the hypothesis that $\mathbf{f}_*$ has maximal rank n on $C_{\mathbf{a}}$, we see that there is an open set $W \supset C_{\mathbf{a}}$, with compact closure, on which $\mathbf{f}_*$ always has maximal rank. Choose an open V with $C_{\mathbf{a}} \subset V \subset \bar{V} \subset W$.

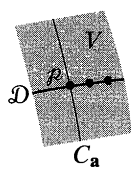
2. Now we claim that if $C_{\mathbf{b}} \subset V$, then $C_{\mathbf{b}}$ must be closed, and thus compact. For otherwise, there would be a point of $\bar{V}$ which is in the closure of $C_{\mathbf{b}}$, though not in $C_{\mathbf{b}}$, itself, contradicting the fact that $\mathbf{f}_*$ has maximal rank at this point.

3. Consequently, if $C_{\mathbf{b}}$ intersects V, but $C_{\mathbf{b}}$ is *not* compact, then $C_{\mathbf{b}}$ must contain points outside of V, and thus by connectedness it must contain a point of the topological boundary bV of V.

4. If there were non-compact $C_{\mathbf{b}}$ arbitrarily close to $C_{\mathbf{a}}$, then there would be a sequence $C_{\mathbf{b}_1}, C_{\mathbf{b}_2}, \ldots$ containing points $p_1, p_2, \ldots$ approaching a point of $C_{\mathbf{a}}$, and corresponding points $q_1, q_2, \ldots$ in bV. These would have a limit point $q \in bV$ for which we would have $\mathbf{f}(q) = \mathbf{a}$, again a contradiction.

Conclusion: Some neighborhood V of the compact $C_{\mathbf{a}}$ *is the union of compact sets of the form* $C_{\mathbf{b}}$.

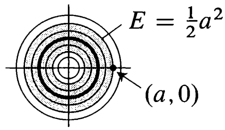
Our construction of the angular coordinates $\varphi^1, \ldots, \varphi^n$ on $C_{\mathbf{a}}$ depended on a choice of a point $p \in C_{\mathbf{a}}$ (the point for which $\varphi^i(p) = 0$ for all i). To choose points $p_{\mathbf{b}}$ in nearby $C_{\mathbf{b}}$, we consider an open set $\mathcal{D} \subset V$ containing p that projects diffeomorphically under $\mathbf{f}$ to an open disc D around $\mathbf{a}$, and simply choose $p_{\mathbf{b}}$ to be the point of $C_{\mathbf{b}}$ that is on $\mathcal{D}$. Choosing D, and thus $\mathcal{D}$,



sufficiently small, each such $C_{\mathbf{b}}$ will be contained completely in V, and the union of these $C_{\mathbf{b}}$ will be a neighborhood V' of $C_{\mathbf{a}}$. Using the $p_{\mathbf{b}}$ as initial points for the φ^i on the $C_{\mathbf{b}}$, we can thus extend the multiple-valued functions φ^i for $C_{\mathbf{a}}$ to functions on V'; for convenience they will also be denoted by φ^i.

Finally, we consider the $2n$ functions $\varphi^1, \ldots, \varphi^n, f_1, \ldots, f_n$ and note that our hypotheses and the compactness of $C_{\mathbf{a}}$ implies that they will be a coordinate system in a neighborhood of $C_{\mathbf{a}}$. So we just shrink D so that $f^{-1}(D)$ is in this neighborhood, and let the union of all the corresponding $C_{\mathbf{b}}$ be U. ❖

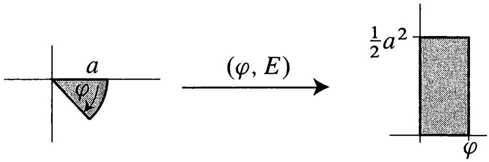
Action-angle variables. In the very simple 1-dimensional case of the harmonic oscillator, for $\omega = 1$, with the single function $f_1 = E$ on $T^*M = \mathbb{R} \times \mathbb{R}$, our φ



$$H = \tfrac{1}{2}(q^2 + p^2)$$

would seem to correspond to the Q^1 given by Liouville's theorem, so it might seem like a good guess that $(\varphi, E)\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ is a symplectic map.

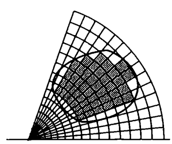
Symplectic is essentially the same as area preserving in the 1-dimensional case, and we note that (φ, E) takes a circular sector having area $\frac{1}{2}\varphi a^2$ into



a rectangle of area $\frac{1}{2}\varphi a^2$, as expected.[1] Moreover, (φ, E) preserves area [up to sign] in general, since the area of an arbitrary region A is given in polar coordinates by

(P) $$\text{area}\, A = \int_A r\, dr \wedge d\varphi = \int_A d\left(\tfrac{1}{2}r^2\right) \wedge d\varphi.$$

This special case of the change of variable formula is often given an elementary proof by approximating a region by subregions for which the formula is obvious from the case of circular sectors. And what this whole argument really

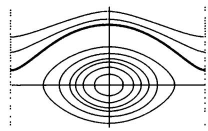


shows is that our map is symplectic not because the value of the second coordinate along the circle of radius a is the energy E along this circle, but because that value happens to be

$$\frac{1}{2\pi} \times (\text{area of the region enclosed by this circle}).$$

[1] Since φ is measured clockwise, one might argue that the area of the circular sector should be counted as $-\frac{1}{2}\varphi a^2$, but we needn't worry about this detail for this intuitive discussion. Being symplectic up to sign is as good as being symplectic, so far as preserving the structure of Hamilton's equations is concerned, and in any case, the direction of φ is actually arbitrarily determined by our choice of 1 or -1 as the generator of the subgroup $G = \mathbb{Z} \subset \mathbb{R}$ in our proof of the invariant tori theorem.

This might suggest that we choose this value for more general cases, like the pendulum. Aside from the fact that we don't really want to mess around with such geometric arguments, this prescription runs into trouble when we consider the $C_{\mathbf{a}}$ for rotations, which don't enclose an area. Fortunately, there is an easy



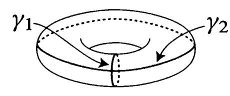
way out, implicit in the calculation of (P) by iterated integrals. When $C_{\mathbf{a}}$ does enclose a region A, so that $C_{\mathbf{a}} = \partial A$, we can use Stokes' theorem to write

$$\int_A dp \wedge dq = \int_A d(p\,dq) = \int_{\partial A} p\,dq = \int_{C_{\mathbf{a}}} p\,dq.$$

So we can define a second coordinate function J on any $C_{\mathbf{a}}$ of U by the formula $J = \dfrac{1}{2\pi}\displaystyle\oint_{C_{\mathbf{a}}} p\,dq$; the traditional circle on the integral sign reminds us that the integral over $C_{\mathbf{a}}$ is computed by integrating over a closed curve γ going once around $C_{\mathbf{a}}$, but it will be eliminated once we start writing integrals over curves.

An even greater advantage of this definition is that it can be used for the case where $C_{\mathbf{a}}$ is an n-dimensional torus T^n. We consider closed curves $\gamma_1, \dots, \gamma_n$ representing generators for the 1-dimensional homology of T^n, and define the numbers J_i for T^n to be

$$J_i = \frac{1}{2\pi}\int_{\gamma_i} \sum_{k=1}^{n} p_k\,dq^k = \frac{1}{2\pi}\int_{\gamma_i} \boldsymbol{\theta},$$



recalling $\boldsymbol{\theta}$ from pages 511–512. This definition does not depend on the particular curve γ_i that we pick: if γ_i and $\gamma_i{}'$ represent the same homology class, so that $\gamma_i - \gamma_i{}' = \partial\sigma$ for some 2-chain σ, then since ω is zero on the tangent space of each $C_{\mathbf{a}}$ by the isotropy lemma on page 615, we have

$$\int_{\gamma_i - \gamma_i{}'} \boldsymbol{\theta} = \int_{\partial\sigma} \boldsymbol{\theta} \overset{\text{Stokes}}{=} \int_{\sigma} d\boldsymbol{\theta} = \int_{\sigma} \omega = 0.$$

Once we have picked $\gamma_1, \dots, \gamma_n$ for $C_{\mathbf{a}}$, the diffeomorphism given by the product neighborhood lemma picks out corresponding curves on each of the $C_{\mathbf{b}} \subset U$. We thus obtain functions $J_1, \dots, J_n \colon U \to \mathbb{R}$ that are constant on each $C_{\mathbf{b}}$. The J_i are called the **action variables**, since they have the dimensions of action (page 464).

While Liouville's theorem gives us symplectic coordinates (Q, f) for which the Q's are integrals, in the case of the more geometric **action-angle** variables

$$(\varphi, J) = (\varphi^1, \dots, \varphi^n, J_1, \dots, J_n,),$$

the J's are actually constants on the invariant tori, but the first set of functions, the φ's, are not integrals. Nevertheless, as in the special case of Liouville's theorem for time-independent Hamiltonians, if (φ, J) is a symplectic coordinate system, then the solutions of Hamilton's equations will still be especially simple: All the φ^i parameter curves, formed by varying one φ^i and keeping the others fixed, are solutions of Hamilton's equations, which implies that H is constant along these curves, so that

$$\frac{\partial H}{\partial \varphi^i} = 0.$$

This means that H does not depend on the φ^i, so that we have

$$\frac{\partial H}{\partial J_j}(p) = \nu_j(J_1(p), \dots, J_n(p))$$

for certain functions ν_j (compare pages 587–588 and 619) and thus Hamilton's equations

$$\dot{\varphi}^i = \frac{\partial H}{\partial J_i}, \qquad \dot{J}_i = -\frac{\partial H}{\partial \varphi^i},$$

will reduce to

$$\dot{\varphi}^i = \nu_i, \qquad \dot{J}_i = 0,$$

with solutions (in condensed notation)

$$J_i = a_i, \quad \varphi^i(t) = b_i + \nu_i(a_1, \dots, a_n) \cdot t \qquad \text{for constants } a_i, b_j.$$

Of course, we really can't expect (φ, J) to be symplectic without some adjustments, since the definition of the φ's depended on various arbitrary choices, not to mention that the f's themselves could be replaced by linear combinations of them. So for our geometric version of Liouville's theorem we just want to prove the existence of certain functions $(\varphi^1, \dots, \varphi^n)$ for which (φ, J) is symplectic and for which the φ's serve as multiple-valued coordinates mod 2π on each of the invariant tori.

Notice, however, that we are no longer working on an arbitrary symplectic manifold, but specifically on the cotangent bundle T^*M, since we use the 1-form $\boldsymbol{\theta} = \sum_{k=1}^n p_k\, dq^k$. Fortunately, this restriction won't be bothersome for the mechanics examples that we will first be looking at, so, saving the general treatment for later, we outline a proof for the case of T^*M; for this proof will also have to assume that (q, J) is a coordinate system [this corresponds to $(*)$ in Liouville's theorem; it wasn't required as an assumption when we were working locally, but now we are working only "semi-locally", in a neighborhood of one of the invariant tori]. The n-tuples $J = (J_1, \dots, J_n)$ will now serve as coordinates for the various tori $C_{\mathbf{b}}$, and we will let $C_{\mathring{J}}$ denote the invariant torus determined by a particular value $\mathring{J}$ of J.

The argument will be similar to that used for Liouville's theorem, with the J's now playing the part of the f's. As in part I of Liouville's theorem, we want to find a type 2 generating function S that gives us the φ's, which means that we want the J's to satisfy

$$\text{(a)} \qquad \frac{\partial S}{\partial q^i}(q, J(q,p)) = p_i,$$

so that we can then define the φ^i by

$$\text{(b)} \qquad \frac{\partial S}{\partial J_i}(q, J(q,p)) = \varphi^i.$$

Since we are assuming that (q, J) is a coordinate system, the implicit function theorem gives us a function $\hat{p}$ such that, for all arguments $\mathring{q}$ and $\mathring{J}$ of q and J, we have

$$J\big(\mathring{q}, \hat{p}(\mathring{q}, \mathring{J})\big) = \mathring{J}.$$

Also, as in the proof of the product neighborhood lemma, we choose initial points $\mathfrak{p}_{\mathring{J}}$ in $C_{\mathring{J}}$ lying along an n-dimensional submanifold transversal to all the tori C_J.

We then define

$$S(\mathring{q}, \mathring{J}) = \int_\gamma \sum_{k=1}^n p_k\, dq^k = \int_\gamma \theta$$

where γ is a curve lying entirely in $C_{\mathring{J}}$, going from $\mathfrak{p}_{\mathring{J}}$ to $\big(\mathring{q}, \hat{p}(\mathring{q}, \mathring{J})\big)$.

Obviously $S(\mathring{q}, \mathring{J})$ is not uniquely defined, since γ can wrap around the torus any number of times, but using the fact that $C_{\mathring{J}}$ is isotropic, we see that S is a well-defined multiple-valued function mod 2π [locally, this corresponds to the use of equation (3) on page 616 to define S on page 617].

Equation (a) is straightforward from the definition of S, and then (b) gives

$$\int_{\gamma_j} d\varphi^i = \int_{\gamma_j} d\left(\frac{\partial S}{\partial J_i}\right) = \frac{\partial}{\partial J_i}\int_{\gamma_j} dS = \frac{\partial}{\partial J_i}(2\pi J_j) = 2\pi\delta_i^j,$$

so that the φ^i are indeed multiple-valued functions mod 2π.

The final step [an analogue of part **II** of Liouville's theorem, but with the J's instead of the f's], is to show that the φ's are actually coordinates on the tori, i.e., that each φ^i parameter curve lies on some torus. For this we basically just reverse the reasoning on page 632: what we need to prove is essentially the same as showing that H is constant along these curves,

$$\frac{\partial H}{\partial \varphi^i} = 0.$$

But this follows automatically: If γ is any solution curve, then $J_i(\gamma(t))$ is constant, since γ lies on an invariant torus, so

$$0 = \frac{dJ_i(\gamma(t))}{dt} = \frac{\partial H}{\partial \varphi^i}(\gamma(t)),$$

and there is some solution curve $\gamma(t)$ through any point of T^*M. **Q.E.D.**

At the end of all this, note that S is defined as an integral, and hence the φ^i and our equations are solvable by quadratures.

As a final remark before looking at two simple examples from mechanics, note that since (φ, J) is a symplectic coordinate system for T^*M, the J_i will be the generalized conjugate momentum to the φ^i for the Lagrangian L corresponding to H,

$$J_i = \frac{\partial L}{\partial \dot{\varphi}^i}.$$

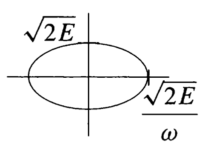
Now L has the dimensions of energy $\mathsf{E} = \mathsf{MV}^2$, as on page 498, while the φ^i are dimensionless, so that the $\dot{\varphi}^i$ will have the dimensions T^{-1}. Consequently, the J_i must have the dimensions ET of action.

♦ For our first example, we consider once again the sorely put-upon harmonic oscillator, with

$$H(q, p) = \tfrac{1}{2}(p^2 + \omega^2 q^2).$$

In this simple case, the curve $H = E$ for a constant E is just the ellipse

$$\frac{\omega^2 q^2}{2E} + \frac{p^2}{2E} = 1,$$



with area equal to $2\pi E/\omega$, so we immediately know that

$$J = \frac{1}{2\pi} \times \frac{2\pi E}{\omega} = \frac{E}{\omega} \qquad \text{or} \qquad H = \omega J \implies \frac{\partial H}{\partial J} = \omega,$$

and the solutions all have circular frequency ω, or ordinary frequency $\omega/2\pi$.

This is obviously too cute by half, but it does illustrate the important general point that knowing the action-angle variables allows us to find the frequencies without having to solve the equations of motion themselves.

If we don't want to take advantage of the special circumstance of this problem, then we must go right back to our calculations in Chapter 18, pages 550–551, where we wrote the solution S of the Hamilton–Jacobi equation as $S(q,\alpha,t) = W(q,\alpha) - \alpha t$ for

$$W(q,\alpha) = \omega \int \sqrt{\frac{2\alpha}{\omega^2} - q^2}\, dq.$$

Note, from the first two equations on page 585 of Chapter 19, that this W is precisely the S in the proof that we sketched on page 633. We can thus write

$$\begin{aligned} J &= \frac{1}{2\pi}\int_\gamma p\,dq = \frac{1}{2\pi}\int_\gamma \frac{\partial W}{\partial q}\,dq && \text{by (a) on page 633} \\ &= \frac{\omega}{2\pi}\int_\gamma \sqrt{\frac{2\alpha}{\omega^2} - q^2}\,dq && \text{(only looking like computational double-talk)} \\ &= \frac{\alpha}{\pi\omega}\int_\gamma \cos^2\theta\,d\theta && \text{using the substitution } q = \sqrt{2\alpha/\omega^2}\,\sin\theta. \end{aligned}$$

Inserting the limits of integration that correspond to a closed curve going once around the invariant torus, we can compute that

$$J = \frac{\alpha}{\pi\omega}\int_0^{2\pi} \cos^2\theta\,d\theta = \frac{\alpha}{\omega},$$

and we again obtain $H = \omega J$, and thereby determine the frequency of all orbits.

When we actually calculate W, say by writing out the formula for $\int \cos^2\theta\,d\theta$, equations (a) and (b) on page 633 can be used to obtain q and p. The results are

$$q = \frac{\sqrt{2E}}{\omega}\sin\omega\varphi, \qquad p = \sqrt{2E}\cos\omega\varphi,$$

where the first formula agrees with formula (b′) on page 551, while the two formulas together illustrate the relationship indicated at the top of page 625.

♦ For the inverse square force in polar coordinates (r,θ) we have to find the functions J_r, J_θ on the 2-dimensional tori. For any potential function V we have the general formula (page 552)

$$W(r,\theta,\alpha,\alpha_\theta) = \int \sqrt{2m(\alpha - V(r)) - \frac{\alpha_\theta{}^2}{r^2}}\,dr + \alpha_\theta\theta,$$

and we get

$$J_r = \frac{1}{2\pi}\int\limits_{\gamma} p_r\,dr = \frac{1}{2\pi}\int\limits_{\gamma} \sqrt{2m(E - V(r)) - \frac{{\alpha_\theta}^2}{r^2}}\,dr,$$

$$J_\theta = \frac{1}{2\pi}\int\limits_{\gamma} p_\theta\,d\theta = \frac{1}{2\pi}\int\limits_{\gamma} \alpha_\theta\,d\theta = \alpha_\theta,$$

illustrating the general fact that when we have separated the variables in the Hamilton–Jacobi equation—usually the only way that we can solve it—the integrals for the J_i simplify, containing just one $p_i\,dq^i$ apiece (compare page 640). Problem 2 gives some further consequences.

For an inverse square force $V(r) = -k/r$, we have to compute the integral

$$I = \frac{1}{2\pi}\int\limits_{\gamma} \sqrt{2mE + \frac{2mk}{r} - \frac{{\alpha_\theta}^2}{r^2}}\,dr.$$

A computation using rather interesting elementary trickery specifically related to the problem is given in Problem 3, but, following the lead of Sommerfeld [1], the computation is usually effected by contour integration (Problem 4), yielding

$$\begin{aligned} J_r &= -\alpha_\theta + k\sqrt{-m/2E} \\ &= -J_\theta + k\sqrt{-m/2E}. \end{aligned}$$

We can write this as

$$H = E = \frac{-k^2 m}{2(J_r + J_\theta)^2},$$

which brings up a second important point. Obviously

$$\frac{\partial H}{\partial J_r} = \frac{\partial H}{\partial J_\theta} \qquad \text{and thus} \qquad \nu_r = \nu_\theta.$$

This means that the periods for the variation of r and θ are the same, and it follows that all orbits *must be closed.* This situation is called "degeneracy", and would also be the case for any rational relation between the periods [in dimensions $n > 2$ we can have different degrees $d \le n$ of degeneracy, with $d = n$ guaranteeing that all orbits are closed].

For a general central force in 2 dimensions, we usually don't have degeneracy, and most orbits are dense in the tori, instead of being closed curves. In such non-degenerate cases, the tori are well-determined, no matter what initial functions $f_1, \ldots, f_n$ we might start with, since they are simply the closures of orbits. But in degenerate cases they usually will not be well-determined—they are tori of dimension less than n, which can be contained in different ways in n-dimensional tori. For example, for an inverse square force, we can separate variables not only in polar coordinates, but also in the elliptic coordinates of Addendum 18A, and these lead to different sets of tori. The situation is similar for the planar harmonic oscillator on page 626, where we can separate variables in polar coordinates as well as in cartesian coordinates. (Nice exercises for the eager reader.)

Action-angle variables on symplectic manifolds. Finally, we end with a proof of the action-angle variables theorem on symplectic manifolds in general, without any additional assumptions. While the proof is more abstract, it is very direct, and only the disembodied ghost of the generating function remains.

ACTION-ANGLE VARIABLES THEOREM. Let (M, ω) be a symplectic manifold of dimension $2n$ with functions $f_1, \ldots, f_n, H \colon M \to \mathbb{R}$ satisfying the hypothesis of the invariant tori theorem, and let D be an open disc $D \subset \mathbb{R}^n$ and U a neighborhood of an invariant torus $C_{\mathbf{a}}$ given by the product neighborhood lemma. Then there is a symplectic coordinate system

$$(\varphi^1, \ldots, \varphi^n, J_1, \ldots, J_n)$$

on U for which the φ's are multiple-valued coordinates mod 2π on the various tori $C_{\mathbf{b}} \subset U$ and the J's are constant on these tori (automatically implying that $\partial H/\partial \varphi^i = 0$, as on page 634).

PROOF. It will be convenient simply to identify U with $D \times C_{\mathbf{a}}$ via the product neighborhood lemma, which tells us that for the projection

$$\pi_D \colon U = D \times C_{\mathbf{a}} \to D$$

the inverse image $\pi_D{}^{-1}(x)$ is an invariant torus for each $x \in D$.

We let X_j be the vector fields used in the proof of the invariant tori theorem, while $\bar{\varphi}^i$ will now be used to denote the φ^i we obtained in that theorem, so that $(\bar{\varphi}^1, \ldots, \bar{\varphi}^n, f_1, \ldots, f_n)$ is the coordinate system for U determined in the proof of the product neighborhood lemma. Then (cf. the remark after the proof of the invariant tori theorem) we can write

$$\frac{\partial}{\partial \bar{\varphi}^i} = \sum_{k=1}^{n} a_{ik} X_k$$

for certain functions a_{ik} that are constant on each invariant torus.

In the expression for the 2-form ω in the $(\bar{\varphi}, f)$ coordinate system, there are no terms involving the $d\bar{\varphi}^i \wedge d\bar{\varphi}^j$, since $\omega = 0$ on the invariant tori. For the coefficient of the $d\bar{\varphi}^i \wedge df_j$ term we have, since $X_k \lrcorner\, \omega = df_k$,

$$\omega\left(\frac{\partial}{\partial\bar{\varphi}^i}, \frac{\partial}{\partial f_j}\right) = \sum_{k=1}^{n} a_{ik}\omega\left(X_k, \frac{\partial}{\partial f_j}\right) = \sum_{k=1}^{n} a_{ik}\, df_k\left(\frac{\partial}{\partial f_j}\right) = a_{ij},$$

and thus

$$\omega = \sum_{i,j=1}^{n} a_{ij}\, d\bar{\varphi}^i \wedge df_j \; + \sum_{i,j=1}^{n} b_{ij}\, df_i \wedge df_j$$

for certain functions b_{ij}. Since $d\omega = 0$, we have

$$\frac{\partial b_{ij}}{\partial\bar{\varphi}^k} = \frac{\partial a_{kj}}{\partial f_i} - \frac{\partial a_{ki}}{\partial f_j},$$

and the right side doesn't depend on the $\bar{\varphi}^i$, so the derivatives of b_{ij} are constant along the $\bar{\varphi}^i$ parameter curves, and thus must be 0, since the parameter curves are closed. Thus the b_{ij} as well as the a_{ij} are constant on each invariant torus.

Now write ω in the form

$$\begin{aligned}
\omega &= \sum_{i=1}^{n} d\bar{\varphi}^i \wedge \left(\sum_{j=1}^{n} a_{ij}\, df_j\right) + \sum_{i,j=1}^{n} b_{ij}\, df_i \wedge df_j \\
&= \sum_{i=1}^{n} d\bar{\varphi}^i \wedge A_i + B \qquad \text{for } A_i = \sum_{j=1}^{n} a_{ij}\, df_j,\; B = \sum_{i,j=1}^{n} b_{ij}\, df_i \wedge df_j.
\end{aligned}$$

Since the a_{ij} and b_{ij} are constant on each invariant torus, we can regard A_i and B as forms on D, i.e., there are 1-forms α_i and a 2-form β on D such that

$$A_i = \pi_D{}^*\alpha_i, \qquad B = \pi_D{}^*\beta.$$

From

$$0 = d\omega = \sum_{i=1}^{n} d\bar{\varphi}^i \wedge \pi_D{}^* d\alpha_i \; + \; \pi_D{}^* d\beta$$

we can conclude that $d\alpha_i = 0$ and $d\beta = 0$, which implies that there are functions I_i and a 1-form γ on the disc D for which[1]

$$\alpha_i = dI_i, \qquad \beta = d\gamma.$$

Finally, we set

$$J_i = (I_i \circ \pi_D) = \pi_D{}^* I_i, \qquad \text{noting that} \qquad dJ_i = A_i.$$

[1] We now have $\omega = d\vartheta$ for $\vartheta = -\sum_{i=1}^{n}(I_i \circ \pi_D)\, d\bar{\varphi}^i + \pi_D{}^*\gamma$, suggesting a proof analogous to the case of T^*M, but we follow a somewhat different route.

Since we have

$$\omega = \sum_{i=1}^{n} d\bar{\varphi}^i \wedge dJ_i + B,$$

the matrix of ω in the $(\bar{\varphi}, f)$ coordinate system is

$$\left(\begin{array}{c|c} 0 & \dfrac{\partial J_i}{\partial f_j} \\ \hline -\dfrac{\partial J_i}{\partial f_j} & b_{ij} \end{array}\right)$$

Since ω is nonsingular, the determinant of this matrix is nonzero, so the same must be true of the determinant of $\left(\partial J_i/\partial f_j\right)$, so $(\bar{\varphi}, J)$ is also a coordinate system.

We now adjust for the arbitrary choice of the origins of our angular coordinates in the proof of the product neighborhood lemma by writing γ as

$$\gamma = \sum_{i=1}^{n} g_i \, dI_i, \qquad \text{and then setting} \qquad \varphi^i = \bar{\varphi}^i + (g_i \circ \pi_D).$$

This gives us

$$\begin{aligned} \sum_{i=1}^{n} d\varphi^i \wedge dJ_i &= \sum_{i=1}^{n} d\bar{\varphi}^i \wedge dJ_i + \sum_{i=1}^{n} d(g_i \circ \pi_D) \wedge dJ_i \\ &= \sum_{i=1}^{n} d\bar{\varphi}^i \wedge dJ_i + \sum_{i=1}^{n} d(g_i \circ \pi_D) \wedge d(I_i \circ \pi_D) \\ &= \sum_{i=1}^{n} d\bar{\varphi}^i \wedge dJ_i + {\pi_D}^* d\gamma = \sum_{i=1}^{n} d\bar{\varphi}^i \wedge A_i + B = \omega. \end{aligned}$$

Thus we have found symplectic coordinates (φ, J) with the J_i constant on the invariant tori, and with the φ^i, like the $\bar{\varphi}^i$, multiple-valued functions mod 2π on these tori.

To make our result resemble the T^*M case more closely, we might make use of the footnote on the previous page, or we can note that

$$\omega = d\theta \qquad \text{for} \qquad \theta = -\textstyle\sum_{i=1}^{n} J_i \, d\varphi^i,$$

and thus

$$J_i = \frac{1}{2\pi} \int_{\gamma_i} \theta,$$

where the γ_i are the parameter curves for the φ^i, but now traversed in the reverse direction. ❖

Background. At the end of this long development, we give a brief historical account, which will indicate how the various strands of our theorem came about and were woven together, and also lead us into the final chapter.

• The source usually quoted for Liouville's theorem, Liouville [1] of 1855, was basically just an announcement of the result. The first published proof I know of occurs in Whittaker [1; §148], originally published in 1904, with a note stating that the theorem is "essentially the application to Hamilton's partial differential equation of the well-known method for finding a Complete Integral of a non-linear partial differential equation of the first order."

Perhaps in line with this view, in Whittaker's proof the step establishing equation (3) on page 616 relies on a forbidding looking thicket of previous results, and the rest of the argument doesn't look especially inviting either. Our proof is based on Pars [1; §22.14 and §25.7] of 1965, except that Pars' proof for equation (3), though straightforward, is nevertheless strangely complicated; the simplification used here comes from the paper Jost [1], which will be mentioned again a little later on, in a more important context.

• Action-angle type variables were first used in 1860 by Delaunay [1] to study the motion of the moon, and for a long time the "Delaunay variables" (for which one may consult Abraham and Marsden [1], Boccaletti and Pucacco [1], and Fasano and Marmi [1]) were used almost exclusively by astronomers.

Early in the 20^{th} century, however, they aroused the interests of physicists when Sommerfeld noted that Neils Bohr's quantum conditions on the orbit of an electron around the nucleus could be formulated as requiring the action variables for the orbit to be an integer multiple of $\hbar = h/2\pi$ (the "royal road to quantization" in Sommerfeld [1]). Einstein [1] in 1917 first drew attention to the integrals $\int_\gamma \sum_k p_k \, dq^k$ in the non-separable case, and by 1924 a thorough presentation of action-angle variables was available in Max Born [1].

The "action-angle" terminology itself was introduced by Schwarzschild [1] in 1916; an English translation may be found in Duck and Sudarshan [1], which also describes the harrowing circumstances under which Schwarzschild worked (while also obtaining the famous Schwarzschild solution in general relativity).

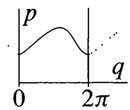
Finally, we note that the action-angle variables had always been used in connection with the Hamilton–Jacobi equation, rather than integrals in involution. Audin [1] cites Mineur [1] of 1936 for the first treatment of the latter case.

• The invariant tori theorem first appeared (in Russian) in Arnold [1] of 1963, and then in Arnold and Avez [1; Appendix 26] of 1968, which begins by noting "It was pointed out long ago that ... the manifolds ... turn out to be tori, and motion along them is quasi-periodic ... " (unfortunately, no mention is made of who did the pointing). The theorem then made its way into Arnold [2] of

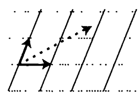
1978, the usual reference, as part of a long exposition in Chapter 10, modestly entitled simply *Liouville's theorem on integrable systems*. The physicist R. Jost gave the first treatment for arbitrary symplectic manifolds in Jost [1] of 1968, where the author likewise modestly suggested that it might be regarded merely as a "sort of commentary" on Arnold's proof.

In addition to the path provided by Jost's paper, other approaches to the modern theorem have been given, and considerable tinkering has been applied to the details, so the basic ideas appear in numerous guises. The proof given here follows Bolsinov and Fomenko [1], which also points to other approaches.

Although the tori provided mathematicians with a wonderful conceptual simplification, physicists, as pointed out previously, had simply discussed things in terms of multiple-valued functions. For example, in the 1-dimensional case, a rotation γ was typically pictured not as a circle on $S^1 \times \mathbb{R}$ but as a repeating



(a)


(b)

function on $\mathbb{R} \times \mathbb{R}$, as in (a) from Born [1]. In other words, physicists basically looked at a covering map $\tilde{\gamma}$ in the universal covering space. Similarly, Born's book contains a representation (b) of the generators of the discrete subgroup G on page 622, and its partition of $\mathbb{R}^n$, essentially describing the universal covering map from $\mathbb{R}^n$ onto these unmentioned tori.

• In any case, the initial excitement for physicists died down quite rapidly, as it soon became clear that despite the striking initial success of the Bohr theory of the atom, it was inadequate, eventually to be replaced by wave mechanics.

Nevertheless, aside from the fact that action-angle variables are immensely important for modern advances in "classical" mechanics, other ideas involved in analyzing the Bohr theory also became important. In particular, one of the pot-holes in Sommerfeld's royal road to quantization was the fact that it only worked when one used the "right" action-angle coordinates, raising the question of how one could characterize these coordinates. This led to the idea of adiabatic invariants in mechanics, which we will consider in the next chapter.

As a result of the mongrel pedigrees for these results, the theorem on invariant tori often bears the name of Liouville or the combination Liouville-Arnold, and the action-angle variables theorem can similarly be found labeled with some combination of Liouville, Arnold, and Jost (with Mineur lurking in the wings).

With some justification, or perhaps just out of exasperation, one might simply end up calling this whole circle of ideas the theory of Liouville integrability.

PROBLEMS

1. Let G be a discrete subgroup of $\mathbb{R}^n$.

(a) If $G \neq \{0\}$, consider an element $\mathbf{b}_1$ of G that is closest to 0, and let V_1 be the subspace generated by $\mathbf{b}_1$. Show that $G \cap V_1$ is generated over $\mathbb{Z}$ by $\mathbf{b}_1$.
(b) If G contains any elements not in V_1, show that it contains an element $\mathbf{b}_2$ closest to V_1. (Hint: For any element not in V_1, there is another the same distance away that is also close to 0.)
(c) If V_2 is the subspace generated by V_1 and $\mathbf{b}_2$, show that $G \cap V_2$ is generated over $\mathbb{Z}$ by $\mathbf{b}_1$ and $\mathbf{b}_2$.
(d) Continue this process inductively to show that G is generated over $\mathbb{Z}$ by $k \leq n$ elements (note: the induction is on k, not on n).

2. Let $C \subset \mathbb{R}^2$ be the closed curve $C = f^{-1}(0)$ where $\operatorname{grad} f = (f_x, f_y) \neq 0$ on C, and let ν be the outward pointing unit normal on C.

(a) Show that for some $\delta > 0$ we have $\langle \operatorname{grad} f, \nu \rangle > \delta$ on C, and use this fact, together with the fact that a tubular neighborhood of C can be defined in terms of ν, to conclude that for $J(h) =$ area enclosed by $f^{-1}(h)$, we have

$$\lim_{h \to 0} \frac{J(h)}{h} \neq 0.$$

Thus, if we use f as a coordinate for a neighborhood of C, then

$$\det\bigl(\partial J/\partial f\bigr) \neq 0 \qquad \text{on } C.$$

(b) In the examples on pages 550–553, after fixing $\alpha = E$, separation of variables for the Hamilton–Jacobi equation for W put it into the form

$$W_1(q^1, \alpha_1, \ldots, \alpha_n) + W_2(q^2, \alpha_2, \ldots, \alpha_n) + \cdots + W_n(q^n, \alpha_n),$$

where α_i is the constant value of p_i; this means that each J_j will depend only on $f_j, \ldots, f_n$. Show that whenever we can write W in this way we have

$$\det \frac{\partial J_j}{\partial f_k} = \prod_{i=1}^{n} \frac{\partial J_i}{\partial f_i} \neq 0,$$

so that (q, J) will be a coordinate system.

3.[1] Writing the term $\sqrt{}$ in the definition of I on page 636 as $\bigl(\sqrt{}\bigr)^2/\sqrt{}$, we get

$$\frac{2\pi I}{\sqrt{2m}} = \int_\gamma \frac{E + k/2r}{\sqrt{E + k/r - \alpha_\theta{}^2/2mr^2}} + \int_\gamma \frac{k/2r}{\sqrt{\cdots}} - \int_\gamma \frac{\alpha_\theta{}^2/2mr^2}{\sqrt{\cdots}}.$$

[1] From Calkin [1].

(a) The first integral is zero over the closed path γ.
(b) Problem 4-11 introduces the true anomaly θ and eccentric anomaly $\tilde{\theta}$ with

$$\text{(i)} \qquad r/a = 1 - \varepsilon\cos\tilde{\theta},$$
$$\text{(ii)} \qquad a(1-\varepsilon^2)/r = 1 + \varepsilon\cos\theta,$$

where, by the equations on pages 124–125,

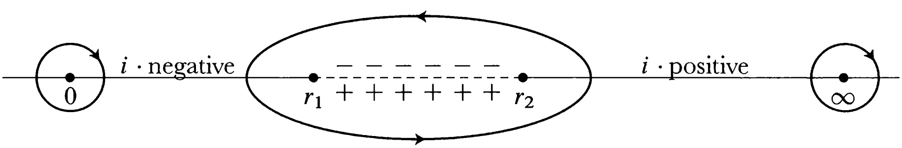
$$a = -k/2E,$$
$$\varepsilon = \sqrt{1 + 2{\alpha_\theta}^2 E/mk^2}.$$

Using (i) as a substitution for the second integral and (ii) as a substitution for the third, show that

$$2\pi I = \sqrt{-mk^2/2E}\int_\gamma d\tilde{\theta} \;-\; \alpha_\theta\int_\gamma d\theta$$
$$= 2\pi\big(\sqrt{-mk^2/2E} \;-\; \alpha_\theta\big).$$

4. The integrand $\sqrt{\cdots}$ in the definition of I on page 636 has two real roots $0 < r_1 < r_2$, so to obtain a complex analytic function we have to make a slit in $\mathbb{C}$ along the segment $[r_1, r_2]$ of the real axis.

(a) Since $\sqrt{\cdots} = p_r = m\dot{r}$, the trajectory of the particle is a path going from $r = r_1$ to $r = r_2$ and back again, so we have to integrate counterclockwise along a path surrounding $[r_1, r_2]$, as in the figure, where the rows of $+$ signs and $-$ signs show the proper sign for the square root on the two halves of the



orbit. Along $\{x > r_2 + 0\cdot i\}$ the values of $\sqrt{\cdots}$ will thus be numbers iy for real $y > 0$, while along $\{x < r_1 + 0\cdot i\}$ they will be iy for real $y < 0$, and the integral will be the negative of the sum of those around the clockwise circles.
(b) The residue of $-\sqrt{\cdots}$ at 0 is $-i\alpha_\theta$, and the residue at ∞ is $ik\sqrt{-2m/E}$, leading to the same formula for J_r as before.

CHAPTER 22

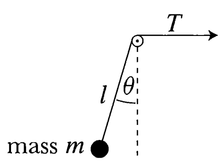
EPILOGUE

The theory of Liouville integrability may be regarded as the culmination of classical mechanics, neatly packaged and gift-wrapped, which encompasses virtually all problems that could be solved classically. It is presented, in various forms, near the end of many classical treatises on mechanics, often followed by some energy-intensive investigations into the stubbornly intransigent three body problem, which mainly seemed to suggest that classical mechanics had reached the point of diminishing returns.

Eventually, completely new mathematical methods were developed to investigate nonintegrable systems, leading to a revitalization of classical mechanics. However, these developments in "modern" classical mechanics would easily fill another entire book, which may perhaps appear some day as Mechanics II. For now, we will merely follow a few leads from the previous chapter, to give a sense of some later developments, with a rather relaxed attitude toward rigor.

Adiabatic invariants. The word *adiabatic*, from Greek *α-* (not) and *διαβαίνω* = pass through, is used in thermodynamics to indicate a process in which no heat enters or leaves a system. By a strange route, the term migrated from thermodynamics to mechanics, where an **adiabatic invariant** is a quantity that remains almost invariant when some parameter of a system changes very slowly.

Probably the earliest example of an adiabatic invariant in mechanics is due to Lord Rayleigh [1], who showed that if the length l of a pendulum is changed very slowly, then the ratio E/ν of the energy of the pendulum to its frequency remains nearly constant; more picturesquely, if the length changes infinitely slowly, then the ratio remains constant. In the situation shown in the figure,



the tension T of the pendulum string is also the force that must be exerted in order to keep the pendulum from sliding down, i.e., the force needed to keep the length l unchanged.

We are considering the case of small oscillations, so that we have the approximation, which we sloppily write as an equality,

$$\cos\theta = 1 - \tfrac{1}{2}\theta^2,$$

and thus for the tension T we have (as on page 210)

$$T = mg\cos\theta + ml\theta'^2 = mg - \tfrac{1}{2}mg\theta^2 + ml\theta'^2.$$

In order to pull the pendulum bob up extremely slowly we will need to exert a force F just a tiny bit larger than T, and we will simply take $F = T$. In one complete period the length of the string changes by a very small amount, $l \to l + \delta l$ (where $\delta l < 0$). If $\langle \ \rangle$ denotes an average over this period, the total work done on the pendulum system over this period will be very close to

$$-\delta l \cdot \langle F\rangle = -\delta l \cdot \langle T\rangle = -mg \cdot \delta l + \tfrac{1}{2}mg \cdot \delta l\,\langle\theta^2\rangle - m \cdot \delta l\,\langle\theta'^2\rangle.$$

The $-mg \cdot \delta l$ term is the increase in potential energy, and the remaining work appears in the increase of the kinetic energy of the oscillating pendulum,

$$\delta E = \tfrac{1}{2}mg \cdot \delta l\,\langle\theta^2\rangle - m \cdot \delta l\,\langle\theta'^2\rangle.$$

We can save ourselves from more involved calculations (compare Problem 1) by noting that the average values of the potential energy and the kinetic energy of the pendulum over a period are the same, and thus equal to half the total energy E, so that

$$\frac{m}{2}gl\langle\theta^2\rangle = \frac{m}{2}l^2\langle\theta'^2\rangle = \frac{E}{2}.$$

Dividing by l and substituting these values into the formula for δE then gives

$$\delta E = -\frac{E}{2l}\delta l,$$

which leads in the limit as $\delta l \to 0$ to the equation

$$\frac{dE}{E} = -\frac{1}{2}\frac{dl}{l} \implies E = \text{constant} \cdot l^{-1/2},$$

and since the length of the period of the pendulum is $2\pi\sqrt{l/g}$, we thus have

$$E = \nu \cdot \text{constant}.$$

This argument is taken[1] from Born [1]. Of course, one might have a sense of unease about this derivation, which doesn't take into account the possibility that the tiny errors introduced by taking the average over a period, during which the length of the pendulum is changing by a very small amount, might accumulate into something significant over the very large number of periods. But Rayleigh and other physicists certainly didn't worry about it.

---

[1] I can't make any sense of Rayleigh's argument, which, among other things, ignores the $ml\theta'^2$ term in T, but still comes up with the same result, though later on, when formulating problems in terms of Lagrangian mechanics, Rayleigh includes this term.

Rayleigh's pendulum, in a slightly different form, was but the first of a series he discussed before using a generalization to derive the Stefan–Boltzmann law for black body radiation, which Boltzmann had derived from principles of thermodynamics and electromagnetism. Paul Ehrenfest, a student of Boltzmann, extended Rayleigh's idea to other questions of thermodynamics and later on he was led to apply ideas about adiabaticity to [the old] quantum mechanics. This resulted, by a tangled path[1] (in which a pivatol role was played by Einstein, virtually the only physicist who had paid any attention to Ehrenfest's work), in what became known as Ehrenfest's "adiabatic hypothesis": since quantum transitions are caused only by influences that vary very rapidly, like light and molecular impacts, the only quantities that should be quantized are those that remain constant under the influence of ordinary phenomenon that vary much more slowly; these quantities were then dubbed "adiabatic invariants".

In view of Sommerfeld's derivations of Bohr's quantum conditions by the quantization of the action variables J_i, this suggested that the action variables should have this property of being adiabatic invariants, which was in fact demonstrated, with about the same degree of rigor as Rayleigh's paper, by Burgers [1], a student of Ehrenfest.

In the end, most methods of the old quantum mechanics became irrelevant, but adiabatic invariants play an important role in modern quantum mechanics as well. In the following sections, by contrast, we will only be considering some aspects of adiabatic invariants in classical mechanics, with a little more attention to mathematical proofs.

The averaging principle. Integrable systems are quite special, and classical mechanics might well owe its existence to the lucky fact that many important simple systems are of this type. But eventually one has to deal with more general systems, and the basic approach is by considering perturbations, systems that are very close to integrable ones. For example, starting from a Liouville integrable system for some Hamiltonian H, written in condensed notation as

$$\dot{J}_i = 0, \qquad \dot{\varphi}^i = \nu_i(J) = \frac{\partial H}{\partial J_i} \qquad\qquad J = (J_1, \dots, J_n),$$

we can consider a nearby "perturbed" system, again in condensed notation,

$$\text{(a)} \qquad \dot{J}_i = \varepsilon f_i(\varphi, J), \qquad \dot{\varphi}^i = \nu_i(J) + \varepsilon g_i(\varphi, J)$$

for $\varepsilon \ll 1$, where f_i and g_i are periodic with period 2π in the φ variables, which are functions on S^1.

[1] See Laidler [1], Mehra and Rechenberg [1; 230–237], and the biography of Ehrenfest by Klein [1]; detailed examination of Ehrenfest's work is given in Navarro and Perez [1], [2] and Perez [1].

More generally, we can consider a system of $n+m$ equations of the form (a) where f and g are functions of (φ, J) for n functions φ^i on S^1 and m functions J_i on $\mathbb{R}$. Note that the J_i are increasing slowly, at the same rate as the very small ε, while the φ^i are increasing at a large rate compared to ε (provided $\nu_i \neq 0$), thus going through many cycles as the J_i slowly changes. So we might hope that a good approximation to the J_i can be found simply by solving the "averaged equations"

$$\dot{X} = \varepsilon \bar{f}(X) \qquad \text{that is,} \quad \dot{X}_i = \varepsilon \bar{f}_i(X) \qquad i = 1, \ldots, m,$$

where $\bar{f} = (\bar{f}_1, \ldots, \bar{f}_m)$ is the average of $f(\varphi, X)$ over a complete cycle of the φ^i,

$$\bar{f}(X) = \frac{1}{(2\pi)^n} \int_0^{2\pi} \cdots \int_0^{2\pi} f(\varphi, X)\, d\varphi^1 \ldots d\varphi^n.$$

In other words, we might hope that the averaged equation has a solution $\bar{J}$ that is close to the solution J of our original equation for small ε. We don't intend to get into the study of perturbation theory here, but it is hardly surprising that this approach might also be useful for studying adiabatic changes.

To get some idea of how $\bar{J}$ compares to J, we consider the simplest example, with $n = m = 1$ and $\nu \neq 0$ constant, as well as $g = 0$, so that we simply have

$$\dot{J} = \varepsilon f(\varphi), \qquad \dot{\varphi} = \nu,$$

and $\varphi(t) = \varphi(0) + \nu t$ while $\bar{f}$ is just the constant $\bar{f} = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} f(\varphi)\, d\varphi$. The averaged equation is then

$$\dot{X} = \varepsilon \bar{f},$$

and the solution with the same initial condition as J is

$$\bar{J}_\varepsilon(t) = J(0) + \varepsilon \bar{f} t.$$

Setting $h = f - \bar{f}$, we can write

$$J(t) - J(0) = \varepsilon \int_0^t f(\varphi(0) + \nu t)\, dt = \varepsilon \int_0^t \bar{f}\, dt + \varepsilon \int_0^t h(\varphi(0) + \nu t)\, dt$$

$$= \varepsilon \bar{f} t + \frac{\varepsilon}{\nu} \int_{\varphi(0)}^{\varphi(0)+\nu t} h(u)\, du,$$

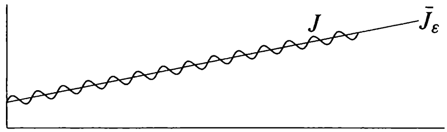
so that

$$J(t) - \bar{J}_\varepsilon(t) = \frac{\varepsilon}{\nu} \int_{\varphi(0)}^{\varphi(0)+\nu t} h(u)\, dt.$$

Since h is periodic, and thus bounded, we have

$$|J(t) - \bar{J}_\varepsilon(t)| < \varepsilon c$$

for some constant c. For small ε the function J is thus the sum of an "evolution" part $\bar{J}_\varepsilon$ with derivative $\varepsilon\bar{f}$, and a small oscillatory part.



A more specific example, showing the difficulty of a direct analysis, is given in Wells and Siklos [1], which considers a slow change of parameter for the harmonic oscillator,

$$H(q,p,\omega(t)) = \tfrac{1}{2}(p^2+\omega(t)^2q^2), \qquad \text{for } \omega(t)^2 = 1+\varepsilon t$$

(we might imagine, for example, that ω is varied by slowly heating an oscillating spring). For $x = -\varepsilon^{-2/3}(1+\varepsilon t)$, the harmonic oscillator equation $\ddot{q}+(1+\varepsilon t)q = 0$ reduces to

$$q''-xq = 0, \quad (q'' = d^2q/dx^2) \qquad \text{[i.e., } g = q\circ x^{-1} \text{ satisfies } g''(x) = xg(x) = 0.]$$

This is the *Airy equation*, whose solutions can be written in terms of the two special Airy functions $\mathrm{Ai}(x)$ and $\mathrm{Bi}(x)$. Using known asymptotic expansions for the Airy functions, Wells and Siklos show that

$$J = H/\omega = \tfrac{1}{2}+\tfrac{1}{4}\varepsilon\omega^{-3}\sin\theta\cos\theta + O(\varepsilon^2) \qquad \text{for} \quad \theta = \frac{2\omega^3-1}{3\varepsilon},$$

so that in this case the evolution term is 0, and J is clearly an adiabatic invariant (although we haven't yet defined exactly what that means).

The "averaging principle" asserts, or assumes, that the motion of equation (a) is always such a sum of an "evolution" part, obtained from the averaged equation, plus a small oscillatory part. It is only an "aspirational" principle (it ain't necessarily so), and we will examine just one, quite special, case where it is true.

An averaging theorem for one-dimensional systems. We consider the system

$$\dot{J}_i = 0, \qquad \dot{\varphi} = \nu(J) \qquad\qquad i = 1,\dots,m$$

for a given initial value $J(0)$, where φ is defined on S^1, while J is defined on some open set $G\subset\mathbb{R}^m$. The perturbed system is

(1) $$\dot{J} = \varepsilon f(\varphi,J), \qquad \dot{\varphi} = \nu(J)+\varepsilon g(\varphi,J)$$

with f and g periodic of period 2π in φ, and the averaged system is

(2) $$\dot{X} = \varepsilon\bar{f}(X), \qquad \bar{f}(X) = \frac{1}{2\pi}\int_0^{2\pi} f(\varphi,X)\,d\varphi.$$

We denote the solution of (2) with $X(0) = J(0)$ by $\bar{J}_\varepsilon$.

Our aim is to show that for sufficiently small ε, the function $\bar{J}_\varepsilon$ is close to J. Moreover, we want it to be close for a long time, $0\le t\le 1/\varepsilon$.

Our theorem will have three basic hypotheses. The first will be some regularity conditions on f, g, and ν. The second will require that ν is bounded away from 0 on G; for example, in the case of the pendulum, we must stay away from the separatrix, where we no longer have periodic orbits. The third hypothesis is the kicker: We want to assume that for sufficiently small ε

$$\bar{J}_\varepsilon(t) \in G \quad \text{for all} \quad 0 \le t \le 1/\varepsilon,$$

which at first sight seems to be unrealistically demanding.

One case where it presents no particular problem is the simplest example examined on page 647, where $\bar{J}_\varepsilon(t) = J(0) + \varepsilon \bar{f} t$ for the constant $\bar{f}$; we only need that the segment from $J(0)$ to $J(0) + \bar{f}$ lies in G. The situation is even simpler if $\bar{f} = 0$, so that $\bar{J}_\varepsilon = J(0)$ for all ε! As a matter of fact, the very situation where we will be applying our theorem is one where $\bar{f} = 0$; in other words, we have a very restrictive theorem, which just happens to be enough for our purposes.

For convenience, from now on we adopt the standard abuse of notation, and simply write $\bar{J}$ instead of $\bar{J}_\varepsilon$, so you have to remember that the ε is really there.

For technical reasons, we also need to introduce one slight complication. Given a (small) number $d > 0$, we will let G_d denote the points of G for which the open disc of radius d around that point is contained in G (alternately, G_d is the set of points at distance $\ge d$ from the boundary of G). We will need to consider G_d rather than G itself, because $\bar{J}$ might still be in G at a time when J has already passed over the boundary of G, so that the equations aren't even defined.

Finally, we will use the following lemma.

LEMMA. Let x be an $\mathbb{R}^m$-valued C^1 function, and let $a, b \ge 0$. Suppose that

$$|x'(t)| \le a|x(t)| + b,$$

using $|\ |$ for the norm in $\mathbb{R}^m$. Then

$$|x(t)| \le \big(|x(0)| + bt\big)e^{at}.$$

PROOF. If y satisfies

$$y'(t) = ay + b, \qquad y(0) = |x(0)|,$$

then $|x(t)| \le y(t)$. But

$$y(t) = C(t)e^{at}$$

for some function C, and we find that $C'(t)e^{at} = b$, or $C'(t) = e^{-at}b$; since $C(0) = |x(0)|$, we thus have $C \le |x(0)| + bt$. ❖

SIMPLE AVERAGING THEOREM. Let $G \subset \mathbb{R}^m$ be an open set, and let $d > 0$. Consider the solution J of the equation

$$\dot{J} = \varepsilon f(\varphi, J), \qquad \dot{\varphi} = \nu(J) + \varepsilon g(\varphi, J) \tag{1}$$

on $S^1 \times G$, with f and g periodic of period 2π in φ, for the initial value $\mathring{J}$, and the averaged system

$$\dot{X} = \varepsilon \bar{f}(X), \qquad \bar{f}(X) = \frac{1}{2\pi}\int_0^{2\pi} f(\varphi, X)\, d\varphi \tag{2}$$

on G for the same initial value $\mathring{J}$.

Suppose that

(1) the functions f, g, and their partial derivatives are bounded on $S^1 \times G$, and similarly for ν on G;
(2) the function ν is bounded away from 0 on G;
(3) for sufficiently small ε, the solution $\bar{J}$ of (2) is defined on $[0, 1/\varepsilon]$ and

$$\bar{J}(t) \in G_d \quad \text{for all} \quad 0 \le t \le 1/\varepsilon.$$

Then for all sufficiently small ε we have

$$\max_{t \in [0,1/\varepsilon]} |J(t) - \bar{J}(t)| = O(\varepsilon)$$

(implicit in this equation is the fact that $J(t)$ is defined on $[0, 1/\varepsilon]$).

PROOF. If we set

$$P(\varphi, J) = J + \varepsilon S(\varphi, J)$$

for a function S, then equation (1) gives

$$\begin{aligned}\dot{P} &= \dot{J} + \varepsilon \frac{\partial S}{\partial \varphi}\dot{\varphi} + \varepsilon \frac{\partial S}{\partial X}\dot{J} \\ &= \varepsilon\left[f(\varphi, J) + \nu(J)\frac{\partial S}{\partial \varphi}\right] + \varepsilon^2 g \frac{\partial S}{\partial \varphi} + \varepsilon^2 f \frac{\partial S}{\partial J}.\end{aligned}$$

In particular, we will choose

$$S(\varphi, J) = -\frac{1}{\nu(J)}\int_0^{\varphi} f(\phi, J) - \bar{f}(J)\, d\phi,$$

so that

$$f(\varphi, J) + \nu(J)\frac{\partial S}{\partial \varphi} = \bar{f}(J),$$

and we then simply get

$$\begin{aligned}\dot{P} &= \varepsilon \bar{f}(J) + O(\varepsilon^2)\\ &= \varepsilon \bar{f}(P) + O(\varepsilon^2).\end{aligned}$$

Comparing with

$$\dot{\bar{J}} = \varepsilon \bar{f}(\bar{J}),$$

we see that $x = P - \bar{J}$ satisfies

$$|x'| \leq \varepsilon A|x| + O(\varepsilon^2),$$

where A is a bound for the derivatives of $\bar{f}$.

The Lemma then shows that there is a constant C such that

$$(*) \qquad |P(t) - \bar{J}(t)| \leq C\varepsilon e^{C\varepsilon t} \quad \text{for } t = O(1/\varepsilon).$$

Since we also have $|J(t) - P(t)| = O(\varepsilon)$ for all t, it looks as if we have the conclusion in the statement of the theorem.

However, there is one delicate point. This argument is only true for $t < \tau$, where τ is the first time that $P(t)$ hits the boundary of G (if ever). So we have to show that τ is at least of order $1/\varepsilon$. To do this, we note that $(*)$ implies that for small enough ε we have $|P(\tau) - \bar{J}(\tau)| < d/2$. So $\dot{P}(t) \in G_d$ for $t \in [0, 1/\varepsilon]$, and this implies that J remains in G on the same interval, for small enough ε. Thus, $\bar{J}$, P, and J all stay in G for times of order $1/\varepsilon$. ❖

This theorem first appeared in Arnold [2; §52C, D]. The proof given here is based on a simplified proof in Lochak and Meunier [1], where a somewhat more general version is given, just the beginning of a very extensive coverage.

Adiabatic invariance of J. Now we will apply this result to study a Hamiltonian $H(q, p, \lambda)$ with a parameter, for a 1-dimensional system. It is convenient to investigate slow changes of the parameter by letting λ be the "slow time", defined by $\lambda = \varepsilon t$ for small ε, and then considering the solutions $t \mapsto (q(t), p(t))$ of the system

$$\dot{q} = \frac{\partial H}{\partial p}(p(t), q(t), \varepsilon t), \qquad \dot{p} = -\frac{\partial H}{\partial q}(p(t), q(t), \varepsilon t)$$

with given initial conditions $(p(0), q(0))$.

DEFINITION. A quantity $A(q, p, \lambda)$ is called an **adiabatic invariant** of this system if

$$\lim_{\varepsilon \to 0} \max_{t \in [0, 1/\varepsilon]} \left| A\big(q(t), p(t), \varepsilon t\big) - A\big(q(0), p(0), 0\big)\right| = 0.$$

In other words, keeping the equations close to the original equations will keep the solutions close for a long time.

Our aim is to show that when the Hamiltonians $H(q, p, \lambda)$ have action-angle variables, the action variable $J(q, p, \lambda)$ is an adiabatic invariant. We are now working on a cotangent bundle T^*M (presumably T^*S^1), so, as on page 633, we have the function $J(q, p, \lambda)$ and a type 2 generating function $S(q, p, \lambda)$ (multiple-valued mod 2π) with

$$(1) \qquad p = \frac{\partial S}{\partial q}(q, J(q, p, \lambda), \lambda), \qquad \varphi = \frac{\partial S}{\partial J}(q, J(q, p, \lambda), \lambda).$$

However, since $\lambda = \varepsilon t$, this is now a time-dependent canonical transformation, and the new Hamiltonian has the form

$$(2) \qquad H_0 + \frac{\partial S}{\partial t} = H_0 + \varepsilon \frac{\partial S}{\partial \lambda},$$

where H_0 is $H(q, p, 0)$ expressed in terms of the (φ, J) variables [i.e., composed with g^{-1} for the map $g(q, p, \lambda) = (\varphi(q, p, \lambda), J(q, p, \lambda), \lambda)$ for φ given by (1).] Note that although the function S is multiple-valued, the function $\partial S/\partial \lambda$ is single-valued, and periodic.

Hamilton's equations for this new Hamiltonian are

$$(3) \qquad \begin{cases} \dot{J} = \varepsilon f(\varphi, J, \lambda) & f = -\dfrac{\partial}{\partial \varphi}\left(\dfrac{\partial S}{\partial \lambda}\right) \\ \dot{\lambda} = \varepsilon & \\ \dot{\varphi} = \nu(J, \lambda) + \varepsilon g(\varphi, J, \lambda) & g = \dfrac{\partial}{\partial J}\left(\dfrac{\partial S}{\partial \lambda}\right). \end{cases}$$

We want to apply the averaging theorem, for the case $m = 2$ (with J and λ being J_1 and J_2), to the averaged system

$$\dot{X}_1 = \varepsilon \bar{f}, \qquad \dot{X}_2 = \varepsilon$$

to get a solution $(\bar{J}, \bar{\lambda})$ with initial conditions $(J(0), 0)$. Of course, $\bar{\lambda} = \varepsilon t$, which is no news, but

$$\bar{f} = -\frac{1}{2\pi} \int_0^{2\pi} \frac{\partial}{\partial \varphi}\left(\frac{\partial S}{\partial \lambda}\right) d\varphi = 0 \qquad \text{since } \frac{\partial S}{\partial \lambda} \text{ is periodic,}$$

so $\bar{J}$ is constant, $\bar{J}(t) = J(0)$, and we can indeed apply the averaging theorem, which then says that

$$\max_{t \in [0, 1/\varepsilon]} |J(t) - J(0)| = O(\varepsilon). \quad \textbf{Q.E.D.}$$

Actually, this proof isn't completely rigorous, as it depends on the additional assumption that we had to make for the construction of action-angle variables on T^*M. However, we won't take the time to worry about this detail. The remainder of this chapter is presented somewhat in physics mode, and we relax our concern with rigor a bit (it makes the physics so much more fun!).

The Hannay angle. The generating function S in equation (1) was chosen to give the transformation $(q, p, \lambda) \mapsto (\varphi, J, \lambda)$ in our first proof of the existence of action-angle variables. It will now be somewhat more useful for us to consider the inverse transformation $(J, \varphi, \lambda) \mapsto (p, q, \lambda)$, where we have also flipped the order of the variables, and a type 2 generating function S for this transformation. Consulting the recipe given at the bottom of page 580, and remembering that Q and P are now p and q, while q and p are J and φ, we obtain the function $q(J, \varphi, \lambda)$ with

$$(1') \qquad p = \frac{\partial S}{\partial q}(q(J,\varphi,\lambda), J, \lambda), \qquad \varphi = \frac{\partial S}{\partial J}(q(J,\varphi,\lambda), J, \lambda).$$

Instead of setting $\lambda = \varepsilon t$, as before, now we simply let $\lambda = t$, and we write the new Hamiltonian analogous to (2) a bit more carefully

$$(2') \qquad H_0 + \frac{\partial S}{\partial t}(q(J,\varphi,t), J, t),$$

with Hamilton's equations for this new Hamiltonian being

$$(3') \qquad \begin{cases} \dot{J} = f(J,\varphi,t) & f = -\dfrac{\partial}{\partial \varphi}\left(\dfrac{\partial S}{\partial t}(q(J,\varphi,t), J, t)\right) \\[2ex] \dot{\varphi} = \nu(J,t) + g(J,\varphi,t) & g = \dfrac{\partial}{\partial J}\left(\dfrac{\partial S}{\partial t}(q(J,\varphi,t), J, t)\right). \end{cases}$$

In all of these more carefully written equations, $\dfrac{\partial S}{\partial t}$ simply denotes $D_3 S$, the partial derivative of S with respect to its third variable, and $\dfrac{\partial S}{\partial t}(q(J,\varphi,t), J, t)$ must be distinguished from

$$\frac{\partial\big(S(q(J,\varphi,t), J, t)\big)}{\partial t} \quad \text{i.e.,} \quad \frac{\partial \mathbf{S}}{\partial t}(J,\varphi,t) \;\text{ for }\; \mathbf{S}(J,\varphi,t) = S(q(J,\varphi,t), J, t).$$

For the latter we have

$$\begin{aligned} \frac{\partial \mathbf{S}}{\partial t}(J,\varphi,t) &= \frac{\partial S}{\partial q}(q(J,\varphi,t), J, t)\frac{\partial q}{\partial t}(J,\varphi,t) + \frac{\partial S}{\partial t}(q(J,\varphi,t), J, t) \\ &= p\frac{\partial q}{\partial t}(J,\varphi,t) + \frac{\partial S}{\partial t}(q(J,\varphi,t), J, t), \end{aligned}$$

so that we can write, with some arguments omitted,

$$\frac{\partial S}{\partial t}(q(J,\varphi,t),J,t) = \frac{\partial \boldsymbol{S}}{\partial t} - p\frac{\partial q}{\partial t}. \tag{4}$$

In particular, suppose that our system returns its original state after some time T. We will then have

$$\int_0^T \frac{\partial \boldsymbol{S}}{\partial t}\,dt = \boldsymbol{S}(T) - \boldsymbol{S}(0) = 0,$$

so that

$$\int_0^T \frac{\partial S}{\partial t}(q(J,\varphi,t),J,t)\,dt = -\int_0^T p(J,\varphi,t)\frac{\partial q}{\partial t}\,dt, \tag{5}$$

a formula that will be very useful for analysing cyclic processes.

While the adiabatic invariance of the action variable J had long been a subject of interest, not much attention was paid to the general behavior of the angle variable φ. Then in a rather strange reversal, quantum mechanics led to new considerations in classical mechanics. In 1984 Berry [1] pointed out that the standard description of adiabatic processes in quantum mechanics was incomplete, missing an additional term for the change of the quantum phase that explained (or at any rate, accounted for) mysterious phenomena like the "Aharonov-Bohm effect". This led Hannay [1] (as well as Berry [2]) to consider the analogous question for the classical case, involving the change of the angle variable φ.

From the equation for $\dot\varphi$ in (3′), we see that

$$\varphi(T) - \varphi(0) = \int_0^T \nu(J,t)\,dt + \frac{\partial}{\partial J}\int_0^T \frac{\partial S}{\partial t}(q(J,\varphi,t),J,t)\,dt.$$

The first term, the "dynamical angle", is what we'd expect from the fact that $\dot\varphi = \nu(J)$, while the second term is a correction that is necessary because J changes with time.

Suppose now that the solution γ of Hamilton's equations for some value $\mathring{J}$ of J traverses a closed curve C for $0 \le t \le T$. Then by (5) we have

$$\varphi(T) - \varphi(0) = \int_0^T \nu(J,t)\,dt - \left.\frac{\partial}{\partial J}\right|_{J=\mathring{J}} \int_0^T p(J,\varphi,t)\frac{\partial q}{\partial t}\,dt,$$

where $\varphi(T)$ really means $\varphi(\gamma(T))$, and $\nu(J,t)$ really means $\nu\big(J(\gamma(t)),t\big)$, and so forth through the whole gamut.

We naturally might expect for the adiabatic case, where φ is making many circuits as J varies slowly in time, that we can approximate the integrand in the second integral by its averaged value,

$$\left\langle p\frac{\partial q}{\partial t}\right\rangle(J,t) = \frac{1}{2\pi}\int_0^{2\pi} p(J,\varphi,t)\frac{\partial q}{\partial t}(J,\varphi,t)\,d\varphi$$

so that, in the adiabatic limit,

$$\varphi(T)-\varphi(0) = \int_0^T \nu(J,\lambda),d\lambda \;-\; \frac{\partial}{\partial J}\bigg|_{J=\mathring{J}} \int_0^T \left\langle p\frac{\partial q}{\partial t}\right\rangle(J,t)\,dt.$$

We won't try to give any justification for this assumption here, instead adopting the physicist's approach, laconically expressed in this passage from Hannay [1] concerning similar considerations for the adiabatic invariance of J:

> Although the adiabatic principle ... is well defined and widely realised physically (Landau and Lifschitz 1976), it appears (Arnold 1978) to be surprisingly difficult to eliminate the mathematical loopholes ... I shall take it for granted.

The quantity

$$\Delta\varphi(C) = -\frac{\partial}{\partial J}\bigg|_{J=\mathring{J}} \int_0^T \left\langle p\frac{\partial q}{\partial t}\right\rangle(J,t)\,dt,$$

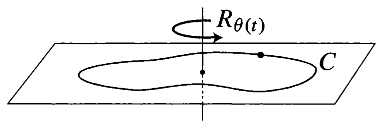
where the integral is sometimes written as $\int_C \langle p\, d_C q\rangle$ or something similar, is called the **Hannay angle**, or the "geometric angle", since it depends on the shape of the closed curve C. Thus we have, in the adiabatic limit,

$$\varphi(T)-\varphi(0) = \underbrace{\int_0^T \nu(J,\lambda)\,d\lambda}_{\text{dynamical angle}} \;+\; \underbrace{\Delta\varphi(C)}_{\text{geometric angle}}.$$

Note that although we have used the term "adiabatic limit", our situation is really different from that considered for the adiabatic invariance of J, as reflected in the fact that we have been working with $\lambda = t$. There actually is no additional parameter λ involved here. The interesting point is that even though φ is changing much faster than J, repeating its position over and over again, the variation in J can result in a cumulative mod 2π change of φ.

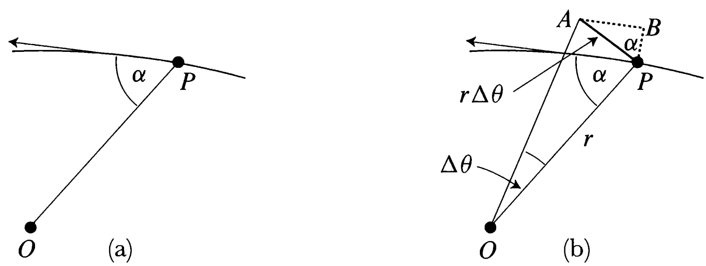
To illustrate the Hannay angle, and partly to consider the evidence for our averaging assumption, we will consider two classical phenomena that are often discussed in connection with the Hannay angle, even though their analysis doesn't necessarily use the formula for it.

The Hannay hoop. Consider a planar hoop C, along which a bead, of mass $m = 1$ for convenience, is sliding without friction, so that its velocity is constant,



and suppose that the entire hoop is being rotated very slowly in its own plane, i.e., about an axis perpendicular to the plane; at time t the hoop will be rotated by the rotation $R_{\theta(t)}$ through an angle of $\theta(t)$, where $\theta'(t)$ is very small.

♦ We will begin by analysing this situation in the same spirit as the analysis of Rayleigh's pendulum, using a hands-on elementary mechanics approach,[1] so that we can see what is actually happening (or what would be happening if we could actually realize this completely idealized situation). Looking at the plane of the hoop from above (a), suppose that O is the point of the plane



about which we are rotating the hoop, P is the position of the bead at some time, and α is the angle between OP and the tangent vector to the hoop at P. If the hoop is rotated (b) by a small angle $\Delta\theta$, the position P on the hoop goes to a position A with AP very nearly perpendicular to OP and having length very close to $r\Delta\theta$. Decomposing the displacement PA into PB perpendicular to the tangent of the curve and BA in the direction of the tangent, we have for their lengths

$$PB = r\Delta\theta\cos\alpha, \qquad BA = r\Delta\theta\sin\alpha.$$

The displacement PB causes a force that pushes the hoop perpendicularly against the bead (for $\alpha \neq \pi/2$), which keeps the bead moving along with the hoop [one of those implicit assumptions inherent in this idealized problem]. On the other hand, since the bead is sliding frictionlessly, the displacement PA in the direction along the hoop has no effect on the bead. Consequently, if s is the arclength along the hoop, measured from some fixed point, the bead falls behind the hoop by the amount $\Delta s = -r\Delta\theta\sin\alpha$.

---

[1] From Calkin [1; pp. 183–184].

The bead is making many circuits around the hoop during the time that the hoop rotates by the small amount $\Delta\theta$, and for a hoop of length ℓ the averaged value of Δs is

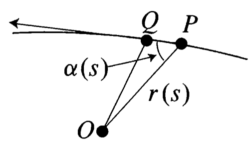
$$\langle \Delta s \rangle = -\left(\frac{1}{\ell}\int_0^\ell r(s)\sin\alpha(s)\,ds\right)\cdot\Delta\theta,$$

where we take the liberty of writing $=$ instead of $\approx$ "in the adiabatic limit".

Now we want to use the formula

$$\text{(A)} \qquad 2A = \int_0^\ell r(s)\sin\alpha(s)\,ds.$$

This can be seen geometrically by noting that for a small change δs of s, the quantity $r(s)\sin\alpha(s)\,\delta s$ is very close to twice the area of the triangle OPQ when



the arclength from P to Q is δs. A more formal derivation is given in Problem 2. Using (A), we now have

$$\langle \Delta s \rangle = -(2A/\ell)\Delta\theta.$$

If the slowly rotating hoop goes through a complete revolution, so that at the end of the process $\Delta\theta = 2\pi$, we then get

$$\langle \Delta s \rangle = -4\pi A/\ell.$$

Thus the amount the bead ends up behind, averaged over all initial positions of the bead, is $-4\pi A/\ell$.

This is often written as

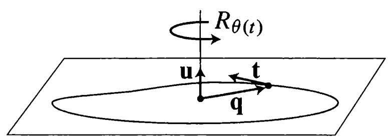
$$\langle \Delta\varphi \rangle = -8\pi^2 A/\ell^2 \qquad \text{for} \quad \varphi = \frac{2\pi}{\ell}\cdot s$$

(note that φ is not the polar coordinate angle around O, but the angle coordinate for the action-angle variables, which we will use later [page 660]). For a circular hoop we have $\langle \Delta\varphi \rangle = -2\pi$. This is obviously correct in the case where the axis of rotation is the center of the circle, since we then have $\alpha = \pi/2$ at all times, and the bead isn't being affected by the hoop at all. Conversely, if $\langle \Delta\varphi \rangle = -2\pi$, then the hoop must be circular, since we can write

$$\Delta\varphi = -2\pi + 2\pi\left(1 - \frac{4\pi A}{\ell^2}\right),$$

where the expression in parentheses is, by the isoperimetric problem, always positive for a non-circular hoop.

♦ For a more detailed analytical treatment using Lagrange's equations,[1] we consider the hoop as our configuration space, with the discussion of the displacements on the bottom of page 656 now magically subsumed under the hypothesis that the bead is confined to move on the hoop. For a number s_0, let $\mathbf{q}(s_0)$ be [the vector from the origin O to] the point on the hoop whose distance from some fixed point is s_0, where the distance is being measured by the arclength along the hoop. Then $\mathbf{q}'(s_0)$ will be the unit tangent vector $\mathbf{t}(s_0)$ to the hoop at that point. At time t let the bead be at the point $\mathbf{q}(s(t))$ on the hoop for some function s. Then in the non-rotating coordinate system, it is at the point $\mathbf{Q}(t) = R_{\theta(t)}\mathbf{q}(s(t))$, where we now regard $R_{\theta(t)}$ as a 3×3 matrix.



If $\mathbf{u}$ is the unit vector at O perpendicular to the plane, then for any vector $\mathbf{v}$ in the plane we can write

$$\left(\frac{d}{dt}R_{\theta(t)}\right)(\mathbf{v}) = \dot{\theta}(t)\,\mathbf{u} \times R_{\theta(t)}(\mathbf{v}) = \boldsymbol{\omega}(t) \times R_{\theta(t)}, \quad \text{say,}$$

so that the derivative of $\mathbf{Q}$ is given by

$$\begin{aligned}\dot{\mathbf{Q}}(s(t)) &= \frac{d}{dt}\Big[R_{\theta(t)}\mathbf{q}(s(t))\Big] \\ &= R_{\theta(t)}\,\mathbf{t}(s(t))\dot{s} + \boldsymbol{\omega}(t) \times R_{\theta(t)}(\mathbf{q}(s(t))) \\ &= R_{\theta(t)}\Big[\mathbf{t}(s(t))\dot{s} + \boldsymbol{\omega}(t) \times \mathbf{q}(s(t))\Big].\end{aligned}$$

The Lagrangian is just the kinetic energy, so

$$\begin{aligned}L(s, \dot{s}, t) &= \frac{1}{2}\big|\mathbf{t}(s)\dot{s} + \boldsymbol{\omega}(t) \times \mathbf{q}(s)\big|^2 \\ &= [\mathbf{t}(s)\dot{s} + \boldsymbol{\omega}(t) \times \mathbf{q}(s)] \bullet [\mathbf{t}(s)\dot{s} + \boldsymbol{\omega}(t) \times \mathbf{q}(s)],\end{aligned}$$

where we use $\bullet$ to indicate inner products, since $\langle\ \rangle$ has been preëmpted for averaged values. Omitting arguments of functions, as usual for working with Lagrangians, note that, with $m = 1$, the conjugate momentum $p = v$ to s is

(a) $$v = p = \frac{\partial L}{\partial \dot{s}} = \mathbf{t} \bullet [\mathbf{t}\dot{s} + \boldsymbol{\omega} \times \mathbf{q}]$$

[1] Cf. Marsden and Ratiu [1; §8.7].

and Lagrange's equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{s}}\right) = \frac{\partial L}{\partial s}$$

becomes

$$\frac{d}{dt}\mathbf{t} \cdot [\mathbf{t}\dot{s} + \boldsymbol{\omega} \times \mathbf{q}] = [\mathbf{t}\dot{s} + \boldsymbol{\omega} \times \mathbf{q}] \cdot [\mathbf{t}'\dot{s} + \boldsymbol{\omega} \times \mathbf{t}].$$

Using $\mathbf{t} \cdot \mathbf{t}' = 0$ and the vanishing of several triple-product terms $\mathbf{a} \cdot [\mathbf{b} \times \mathbf{c}]$, we obtain

$$\ddot{s} - [\boldsymbol{\omega} \times \mathbf{q}] \cdot [\boldsymbol{\omega} \times \mathbf{t}] + \mathbf{t} \cdot [\dot{\boldsymbol{\omega}} \times \mathbf{q}] = 0,$$

with the second and third terms being the centrifugal force and the Euler force. Since $\boldsymbol{\omega} = \dot{\theta}\mathbf{u}$, we get finally

$$\ddot{s} = \dot{\theta}^2 \mathbf{q} \cdot \mathbf{t} - m\ddot{\theta}|\mathbf{q}| \sin\alpha,$$

where α is the angle between $\mathbf{q}$ and $\mathbf{t}$, as before.

The integral form of the remainder in Taylor's formula gives

$$\begin{aligned} s(t) &= s(0) + \dot{s}(0)t \\ &\quad + \int_0^t (t-\tau)\Big[\dot{\theta}(\tau)^2 \mathbf{q}(s(\tau)) \cdot \mathbf{t}(s(\tau) - \ddot{\theta}(\tau)|\mathbf{q}|(s(\tau)) \sin\alpha(s(\tau))\Big]\, d\tau. \end{aligned}$$

Now we'd like to average, replacing the quantities involving s in the integral by their averages around the hoop. Assuming that $\dot{\theta}$ and $\ddot{\theta}$ are both small in comparison to the velocity of the bead, it appears that there is a theorem[1] showing that for large T, like the time for the hoop to make one complete revolution, the quantity $s(T) - s(0) - \dot{s}(0)T$ is approximately

$$\int_0^T (T-\tau)\Big[\dot{\theta}(\tau)^2 \frac{1}{\ell}\int_0^\ell \mathbf{q}(s) \cdot \mathbf{t}(s)\, ds - \ddot{\theta}(\tau)\frac{1}{\ell}\int_0^\ell |\mathbf{q}|(s) \sin\alpha(s)\, ds\Big]\, d\tau.$$

Assuming this, we note that since $\mathbf{q}(s)\cdot\mathbf{t}(s) = \frac{1}{2}(d/ds)[\mathbf{q}(s)\cdot\mathbf{q}(s)]$, the integral of this term over the whole hoop vanishes, and the other integral, involving the Euler force, is $2A$, by equation (A) on page 657 again, so that

$$\begin{aligned} s(T) &\approx s(0) + \dot{s}(0)T - \frac{2A}{\ell}\int_0^T (T-\tau)\ddot{\theta}(\tau)\, d\tau \\ &= s(0) + \dot{s}(0)T + \frac{2A}{\ell}\dot{\theta}(0)T - \frac{4\pi A}{\ell} \qquad \text{using integration by parts.} \end{aligned}$$

[1] Marsden and Ratiu [1] refer to Hale [1], though it seems one might have to read half the book to extract the result.

For the special case where $\dot{\theta}(0) = 0$, we simply have

$$s(T) \approx s(0) + \dot{s}(0)T - \frac{4\pi A}{\ell},$$

where $s(0) + \dot{s}(0)T$ would be the total arclength of the hoop if the bead traveled with constant velocity $\dot{s}(0)$, with the correction term

$$\Delta s = -\frac{4\pi A}{\ell}.$$

In the general case, we must find the average displacement over all initial positions, for a bead with constant velocity v_0, say. Equation (a) gives

$$v_0 = \dot{s}(0) + \dot{\theta}(s(0))q(s(0)) \sin\alpha(s(0)),$$

so we can write

$$s(T) - s(0) - v_0T \approx -\dot{\theta}(0)q(s(0)) \sin\alpha(s(0))T + \frac{2A}{\ell}T - \frac{4\pi A}{\ell},$$

and the first term on the right averages out over the hoop to cancel the second term, so the average shift, over all initial positions of the bead, is again $-4\pi A/\ell$.

♦ Finally, we want to consider how this question fits in with the Hannay angle. Our manifold is now just the hoop C, and the Hamiltonian is

$$H(\mathbf{q}, \mathbf{p}) = \frac{|\mathbf{p}|^2}{2},$$

with the bead having constant velocity $p = |\mathbf{p}|$ with respect to C. The action, defined by $J = \dfrac{1}{2\pi}\displaystyle\int_C p\,ds$, is just

(1) $$J = \frac{1}{2\pi}p\ell,$$

and on C we have

(2) $$\varphi = \frac{2\pi}{\ell} \cdot s,$$

where the fraction $2\pi/\ell$ is used because, by definition (page 621), φ goes once around the invariant torus on the interval $[0, 2\pi]$.

In computing the Hannay angle, instead of using t on $[0, T]$, where T is the time for a complete rotation of the hoop, we will use the corresponding angle θ through which the hoop has been rotated for $0 \le \theta \le 2\pi$. We then have, with $\mathbf{t}$ denoting the unit tangent vector to C,

$$\begin{aligned}
\Delta\varphi &= -\frac{\partial}{\partial J}\int_0^{2\pi}\left\langle \mathbf{p}\cdot\frac{\partial \mathbf{q}}{\partial\theta}\right\rangle(J,\theta)\,d\theta \\
&= -\frac{\partial}{\partial J}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi}\mathbf{p}(J,\varphi,\theta)\cdot\frac{\partial \mathbf{q}}{\partial\theta}(J,\varphi,\theta)\,d\varphi\right)d\theta \\
&= -\frac{\partial}{\partial J}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi}\frac{2\pi J}{\ell}\mathbf{t}(\theta)\cdot\frac{\partial \mathbf{q}}{\partial\theta}(J,\varphi,\theta)\,d\varphi\right)d\theta \qquad \text{by (1)} \\
&= -\frac{1}{\ell}\int_0^{2\pi}\left(\int_0^{2\pi}\mathbf{t}(\theta)\cdot\frac{\partial \mathbf{q}}{\partial\theta}(J,\varphi,\theta)\,d\varphi\right)d\theta \\
&= -\frac{2\pi}{\ell^2}\int_0^{2\pi}\left(\int_0^{\ell}\mathbf{t}(\theta)\cdot\frac{\partial \mathbf{q}}{\partial\theta}(J,s,\theta)\,ds\right)d\theta \qquad \text{using the substitution (2)} \\
&= -\frac{2\pi}{\ell^2}\int_0^{2\pi}\left(\int_0^{\ell}|\mathbf{q}|(s)\sin\alpha(s)\,ds\right)d\theta \\
&= -\frac{8\pi^2 A}{\ell^2} \qquad \text{by equation (A) on page 657 once again.}
\end{aligned}$$

Foucault's pendulum revisited. At first sight, the classical Foucault pendulum would seem to be a perfect candidate for analysis in terms of the Hannay angle. After all, here we have the angle φ of the pendulum varying rapidly while the whole system undergoes an extremely slow change due to the rotation of the earth. On the other hand, it is not the total change in φ that interests us, but the change of the angle ϕ of the plane in which the pendulum swings. We can analyse Foucault's pendulum in terms of action-angle variables,[1] but the Foucault pendulum phenomenon will really turn out to be just a cousin of the Hannay angle phenomenon.

The angle φ ducks out of the picture as soon as we use the equations $(*)$ on page 389, where we now indicate the latitude in that equation by ℓ instead of λ,

[1] From Khein and Nelson [1].

and abbreviate the term $\omega \sin \ell$ by ϖ:

$$x'' = -\alpha^2 x + (2\varpi)y' \tag{1}$$
$$y'' = -\alpha^2 y - (2\varpi)x'.$$

We note that these are precisely Lagrange's equations for the Lagrangian

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - \frac{1}{2}\alpha^2(x^2 + y^2) + \varpi(x\dot{y} - y\dot{x}).$$

Introducing polar coordinates (ρ, ϕ) in the (x, y)-plane,

$$x = \rho \cos \phi, \qquad y = \rho \sin \phi,$$

we express L in polar coordinates by

$$L = \frac{1}{2}(\dot{\rho}^2 + \rho^2\dot{\phi}^2) - \frac{1}{2}\alpha^2\rho^2 + \varpi\rho^2\dot{\phi}.$$

The conjugate momenta are then

$$p_\rho = \frac{\partial L}{\partial \dot{\rho}} = \dot{\rho}, \qquad p_\phi = \frac{\partial L}{\partial \dot{\phi}} = \rho^2(\dot{\phi} + \varpi),$$

and the corresponding Hamiltonian H is

$$H = \frac{{p_\rho}^2}{2} + \frac{{p_\phi}^2}{2\rho^2} - \varpi p_\phi + \frac{1}{2}(\alpha^2 + \varpi^2)\rho^2. \tag{2}$$

Since ϕ is cyclic in L, the quantity p_ϕ is a constant. Note that H involves no parameter, and doesn't even depend on time, and our variables are now ρ and ϕ.

We now look for the action variables, just as in the case of a central force on page 636,

$$J_\rho = \frac{1}{2\pi}\int_\gamma p_\rho \, d\rho$$
$$J_\phi = \frac{1}{2\pi}\int_\gamma p_\phi \, d\phi = p_\phi.$$

Setting $H = E$ in (2), we obtain

$$p_\rho = \pm\frac{1}{\rho}\sqrt{-{p_\phi}^2 + 2(E + \varpi p_\phi)\rho^2 - (\alpha^2 + \varpi^2)\rho^4}.$$

In terms of

$$r = \rho^2, \qquad a = \frac{p_\phi{}^2}{\alpha^2 + \varpi^2}, \qquad b = \frac{E + \varpi p_\phi}{\alpha^2 + \varpi^2}$$

we can write

$$J_\rho = \frac{1}{4\pi}\sqrt{\alpha^2 + \varpi^2}\int_\gamma \sqrt{-\frac{a}{r^2} + 2\frac{b}{r} - 1}\, dr.$$

This can be evaluated by contour integration, as in Problem 21-4, although now there are four real roots to worry about. The result is

$$J_\rho = \frac{1}{2}\sqrt{\alpha^2 + \varpi^2}\left(b - \sqrt{a}\right) = \frac{1}{2}\left(\frac{E + \varpi p_\phi}{\sqrt{\alpha^2 + \varpi^2}} - |p_\phi|\right),$$

where the $\sqrt{a}$, involving $\sqrt{p_\phi{}^2}$, gives rise to the $|p_\phi|$ term.[1] Solving this for E gives

(3) $$H = E = (2J_\rho + |J_\phi|)\sqrt{\alpha^2 + \varpi^2} - J_\phi \varpi,$$

expressing H in terms of the action-angle variables.

For solutions with J_ϕ negative we have $\partial|J_\phi|/\partial J_\phi = -1$, so Hamilton's equation for ϕ is

$$\dot{\phi} = \frac{\partial H}{\partial J_\phi} = -\sqrt{\alpha^2 + \varpi^2} - \varpi = -\omega_1, \quad \text{say,}$$

so that for some constant C we have

(4) $$\phi(t) = -\omega_1 t + C.$$

For solutions with J_ϕ positive, we have $\partial|J_\phi|/\partial J_\phi = 1$ and we get

$$\dot{\phi} = \frac{\partial H}{\partial J_\phi} = \sqrt{\alpha^2 + \varpi^2} - \varpi = \omega_2, \quad \text{say,}$$

so that

(5) $$\phi(t) = \omega_2 t + C.$$

[1] As Khein and Nelson [1], [2] point out, this absolute value sign will be essential in this case, but is almost always neglected (as in the calculations on page 636), although in most cases that doesn't lead to any difficulties.

Remembering that $\varpi = \omega \sin \ell$ for the angular velocity ω of the earth, we find that in both cases the total change in ϕ in one day's time is

$$\begin{aligned}\phi\left(\frac{2\pi}{\omega}\right) - \phi(0) &= \pm\frac{2\pi}{\omega}\sqrt{\alpha^2 + \varpi^2} - \frac{2\pi}{\omega} \cdot \omega \sin \ell \\ &= \pm\frac{2\pi}{\omega}\sqrt{\alpha^2 + \varpi^2} - 2\pi \sin \ell.\end{aligned}$$

The first term, depending on the sign of J_ϕ, is the dynamical angle. The second term is the "Hannay angle",

$$\Delta\phi = -2\pi \sin \ell,$$

showing that the angular rate of change of ϕ is $-\omega \sin \ell$ in both cases.

To connect this perhaps somewhat confusing discussion with the Foucault pendulum, note that Foucault's pendulum is just one of the solutions for this Hamiltonian, and for any solution the value of J_ϕ will depend on the initial conditions, though we actually have to be a bit careful about what that means.

In the case of the Foucault pendulum, we release the bob from rest, taking care not to impart any sidewise motion. But J_ϕ for the Foucault pendulum is still positive (in the northern hemisphere), because the "rest" point from which we start it already has its own positive J_ϕ. To get $J_\phi = 0$, we have to give the bob an initial angular velocity $\dot{\phi} = -\varpi$ (as reckoned in our [actually rotating] coordinate system).

On the other hand, knowing the value of J_ϕ for the Foucault pendulum, or even the sign of J_ϕ, doesn't really matter for our problem, since every value of J_ϕ ends up giving the same Hannay angle, and even $J_\phi = 0$ gives the same Hannay angle, though its dynamical angle is 0.

These considerations can be illuminated by the investigations of Onnes that were mentioned previously on page 391. All solutions to (1) starting at rest with initial condition $\rho(0) = \rho_0$ can be written in terms of two normal modes. In both normal modes the pendulum bob traces out a circle of radius ρ_0, with mode 1 going clockwise, having $J_\phi < 0$, and mode 2 going counter-clockwise, having $J_\phi > 0$. Since every solution is a combination of the two normal modes, with the same Hannay angle, every solution has that same Hannay angle. Many more details, including the behavior of the ρ variable, and specific data for Foucault's pendulum, can be found in Khein and Nelson [1].

PROBLEMS

1. For small oscillations we have $\cos\theta = 1 - \frac{1}{2}\theta^2$ and the pendulum equation, $\theta'' + (g/l)\theta = 0$, has the solution

$$\theta = A\cos\omega t, \qquad \omega^2 = g/l.$$

(a) The energy E, the sum of the kinetic and potential energy, is

$$E = \tfrac{1}{2}mglA^2$$

and the tension is

$$T = mg - \tfrac{1}{2}mgA^2\cos^2\omega t + mgA^2\sin^2\omega t.$$

(b) The average $\langle F\rangle = \langle T\rangle$ over a complete period is $(\omega/2\pi)\int_{-\pi/\omega}^{\pi/\omega} F$. Using the fact that $\int_{-\pi}^{\pi}\cos^2 t\,dt = \int_{-\pi}^{\pi}\sin^2 t\,dt = \pi$, show that

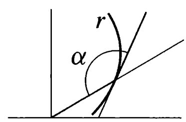
$$\langle F\rangle = mg + \tfrac{1}{4}mgA^2.$$

(c) Conclude that

$$-\delta l\cdot\langle F\rangle = -mg\cdot\delta l - \tfrac{1}{4}mgA^2\cdot\delta\ell \implies \delta E = -\tfrac{1}{4}mgA^2\cdot\delta l \implies \frac{\delta E}{E} = -\frac{1}{2}\frac{dl}{l}.$$

2. Given a function r, consider the curve $\theta \mapsto (r(\theta)\cos\theta, r(\theta)\sin\theta)$ (the graph of r in polar coordinates). Problem 9-5 shows that

$$\tan\alpha = \frac{r(\theta)}{r'(\theta)}.$$



(The α in Problem 9-5 is the supplement of the α in this figure, and tan is the same for both.)

For the region R bounded by the graph C of r, the area A is (as on page 630)

$$A = \int_R dr\wedge d\theta = \int_R d(\tfrac{1}{2}r^2\,d\theta) = \int_C \tfrac{1}{2}r^2\,d\theta = \frac{1}{2}\int_0^{2\pi} r(\theta)^2\,d\theta.$$

Using the substitution

$$s = \int_{\theta_0}^{\theta}\sqrt{r(\theta)^2 + r'(\theta)^2}\,d\theta, \qquad ds = \sqrt{r(\theta)^2 + r'(\theta)^2}\,d\theta$$

show that for C of length ℓ we have

$$A = \frac{1}{2}\int_0^{\ell}\frac{r(s)\,ds}{\sqrt{1 + \left[\dfrac{r'(s)}{r(s)}\right]^2}}, \qquad \text{leading to} \qquad 2A = \int_0^{\ell} r(s)\sin\alpha(s)\,ds.$$

SUPPLEMENT
A PDE PRIMER

The material of this Supplement is basically the beginning of the first chapter of DG, Vol. 5.

When we consider an ordinary differential equation $u'(x) = f(x, u(x))$, we find that there are solutions u with any desired value for $u(x_0)$, this dependence on the "initial condition" $u(x_0)$ usually manifesting itself, if we explicitly solve the equation, by the presence of an arbitrary constant of integration. Equations of order n, on the other hand, will involve n constants of integration.

When we solve a PDE, we usually obtain arbitrary *functions* in the answer. For example, to be as simple-minded about the thing as we can, we note that the equation

$$\frac{\partial u}{\partial y}(x, y) = 0$$

has the solutions $u(x, y) = A(x)$; the only restrictions on A are ones which follow from restrictions we might choose to place on u (e.g., that u be differentiable with respect to x). The equally stupid looking, but actually quite important, second order equation

$$\frac{\partial^2 u}{\partial x \partial y}(x, y) = 0$$

leads to

$$\frac{\partial u}{\partial x}(x, y) = \alpha(x),$$

and hence to

$$u(x, y) = A(x) + B(y), \qquad A'(x) = \alpha(x).$$

Without belaboring the point any further, we simply note that when we look for precise theorems, we should expect the hypotheses to reflect the presence of these "arbitrary functions" in the same way that the precise theorem for ordinary differential equations reflects the presence of arbitrary constants.

We will first consider equations that involve a function u on $\mathbb{R}^n$ and only its first partial derivatives u_{x_i}. For simplicity of writing, and convenience of visualization, we will first deal exclusively with the case of $\mathbb{R}^2$, denoting a typical point of $\mathbb{R}^2$ by (x, y) and adopting the standard notation

$$u_x = p, \quad u_y = q.$$

By a **first order PDE** we then mean an equation of the form

$$F\big(x, y, u(x, y), u_x(x, y), u_y(x, y)\big) = 0,$$

or, to use the standard abbreviated form,

$$F(x, y, u, p, q) = 0.$$

It will be convenient to denote the various partial derivatives of F by F_x, F_y, F_u, F_p, and F_q. Naturally, the function $F\colon \mathbb{R}^5 \to \mathbb{R}$ shouldn't be too badly behaved; for example, it wouldn't be very interesting if F were never 0. Just what hypotheses we really need will come out soon enough. To begin with, we might imagine that F is differentiable and satisfies $F_p \neq 0$ or $F_q \neq 0$, so that by the implicit function theorem we can solve for p in terms of q, or *vice versa*. Our main result is, that we can always completely reduce any first order PDE to a system of ordinary differential equations. This holds both in a "practical" and in a theoretical sense: We can actually write down a system of ordinary differential equations whose solutions, if we can find them, will give us the solution of our original problem; and the method by which this is done enables us to state and prove exact theorems. We will not deal at the very outset with the most general first order PDE, but will approach it in stages.

We consider first the most general **linear first order PDE**

$$(1) \qquad A(x, y)u_x(x, y) + B(x, y)u_y(x, y) = C(x, y)u(x, y) + D(x, y).$$

Usually this is simply written

$$A(x, y)u_x + B(x, y)u_y = C(x, y)u + D(x, y),$$

with the arguments (x, y) appearing in A, B, C, and D just to emphasize that we are not considering an equation like $A(x, y, u(x, y))u_x + \cdots$.

Consider the vector field X on $\mathbb{R}^2$ defined by

$$(2) \qquad X = A\frac{\partial}{\partial x} + B\frac{\partial}{\partial y}.$$

The value of X at (x_0, y_0) is

$$A(x_0, y_0)\left.\frac{\partial}{\partial x}\right|_{(x_0,y_0)} + B(x_0, y_0)\left.\frac{\partial}{\partial y}\right|_{(x_0,y_0)};$$

using the standard identification of the tangent space $\mathbb{R}^2{}_{(x_0,y_0)}$ with $\mathbb{R}^2$, we can also write

$$X(x_0, y_0) = \big(A(x_0, y_0), B(x_0, y_0)\big).$$

We will call X the **characteristic vector field** of equation (1); the integral curves of this vector field are called the **characteristic curves** of equation (1). Thus $c = (c_1, c_2)$ is a characteristic curve if and only if

$$\text{(3)} \qquad \frac{dc_1(t)}{dt} = A(c(t)), \quad \frac{dc_2(t)}{dt} = B(c(t)).$$

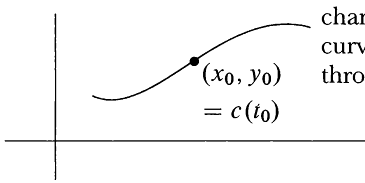
We then have, for any C^1 function $u\colon \mathbb{R}^2 \to \mathbb{R}$,

$$\begin{aligned}\frac{du(c(t))}{dt} &= u_x(c(t))\frac{dc_1(t)}{dt} + u_y(c(t))\frac{dc_2(t)}{dt} \\ &= A(c(t)) \cdot u_x(c(t)) + B(c(t)) \cdot u_y(c(t)).\end{aligned}$$

So any solution u of equation (1) satisfies

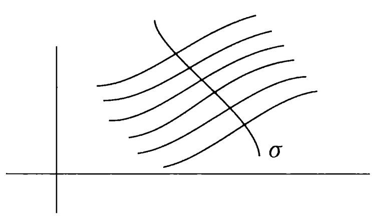
$$\text{(4)} \qquad \frac{du(c(t))}{dt} = C(c(t)) \cdot u(c(t)) + D(c(t)) \qquad \text{for any characteristic curve } c.$$

For any fixed characteristic curve $t \mapsto c(t)$, equation (4) is an ordinary differential equation for the function $u \circ c$. Consequently, $u \circ c$ is uniquely determined

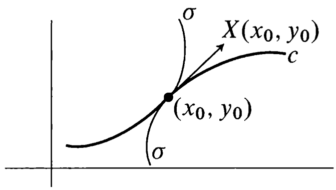


once $u(c(t_0))$ is specified. In other words, once we prescribe a value $u(x_0, y_0)$ for a solution u of equation (1), the solution u will then be completely determined along the characteristic curve c through (x_0, y_0).

Now suppose we have any curve σ which cuts a family of characteristic curves.



If we arbitrarily specify the values of u at each point of σ, then the solution u will be determined in a neighborhood of σ. Moreover, we ought to be able to produce this solution u simply by solving equation (4) for each of the characteristic curves through each point of σ. Of course, we clearly have to rule out the possibility that a portion of σ itself is a characteristic curve, for then we could not arbitrarily specify the values of u along σ. We even have to rule out the possibility that σ is tangent to some integral curve c at some point $(x_0, y_0) = c(t_0)$; for in this case, the directional derivative $X(x_0, y_0)(u)$ would



be determined both by equation (4) and (in a possibly conflicting way) by the arbitrarily assigned values of u along σ. We must thus assume that the vectors

$$\sigma'(s) = \big(\sigma_1{}'(s), \sigma_2{}'(s)\big) \qquad \text{and} \qquad \big(A(\sigma(s)), B(\sigma(s))\big)$$

are always linearly independent. Equivalently, we must require that

$$0 \neq \det\begin{pmatrix} \sigma_1{}'(s) & A(\sigma(s)) \\ \sigma_2{}'(s) & B(\sigma(s)) \end{pmatrix} = \sigma_1{}'(s)B(\sigma(s)) - \sigma_2{}'(s)A(\sigma(s))$$

for all s. In particular, $\sigma'(s) \neq (0,0)$ so σ is an imbedding. Although we will later have a much more general result, we summarize this information in a theorem, in order to get all the details cleaned up before we carry the discussion any further.

1. THEOREM. Let A, B, C, and D be C^k functions defined in an open set $U \subset \mathbb{R}^2$, and let $\sigma\colon [a,b] \to U$ be a one-one C^k curve such that

$$\sigma_1{}'(s)B(\sigma(s)) \neq \sigma_2{}'(s)A(\sigma(s)) \qquad \text{for all } s \in [a,b].$$

Let $\mathring{u}\colon [a,b] \to \mathbb{R}$ be a C^k function. Then there is a C^k function u, defined in a neighborhood V of $\sigma([a,b])$, such that u satisfies

$$(1) \qquad A \cdot u_x + B \cdot u_y = C \cdot u + D \qquad \text{on } V,$$

with the initial condition

$$u(\sigma(s)) = \mathring{u}(s) \qquad \text{for all } s \in [a,b].$$

Moreover, any two functions u with this property agree on a neighborhood of $\sigma([a,b])$.

PROOF. There is a C^k map

$$\gamma\colon [a,b]\times(-\varepsilon,\varepsilon)\to U$$

such that each curve

$$t\mapsto \gamma(s,t)$$

is a characteristic curve with

$$\gamma(s,0)=\sigma(s).$$

Clearly

$$\frac{\partial\gamma}{\partial s}(s,0)=\sigma'(s)=(\sigma_1{}'(s),\sigma_2{}'(s))$$
$$\frac{\partial\gamma}{\partial t}(s,0)=(A(\sigma(s)),\,B(\sigma(s))).$$

So, by the hypothesis on σ, the Jacobian of γ at $(s,0)$ is always nonsingular; consequently, if ε is sufficiently small, then γ is a C^k diffeomorphism onto a neighborhood V of $\sigma([a,b])$.

By choosing ε still smaller, if necessary, we can insure that for each $s\in[a,b]$ there is a C^k function $\beta_s\colon(-\varepsilon,\varepsilon)\to\mathbb{R}$ satisfying

$$\begin{cases}\dfrac{d\beta_s(t)}{dt}=C(\gamma(s,t))\cdot\beta_s(t)+D(\gamma(s,t))\\[2mm] \beta_s(0)=\mathring{u}(s)\end{cases}$$

[this is just the equation (4) which should be satisfied by $u\circ c$ along the integral curve $t\mapsto\gamma(s,t)$]. We would actually like to know that $\beta_s(t)$ is C^k as a function of s and t; in other words, if we define $\beta\colon[a,b]\times(-\varepsilon,\varepsilon)\to\mathbb{R}$ by

$$\beta(s,t)=\beta_s(t),$$

then we would like to know that β is C^k. To prove this, we must use the fact (see, e.g., DG, Vol. 1, Prob. 5-5) that we can solve the equation "depending on parameters"

$$\begin{cases}\alpha(0,s,r)=r \qquad \text{for } r\in\mathbb{R}\\[2mm] \dfrac{\partial}{\partial t}\alpha(t,s,r)=C(\gamma(s,t))\cdot\alpha(t,s,r)+D(\gamma(s,t)),\end{cases}$$

for a C^k function α, so that

$$\beta(s,t)=\alpha(t,s,\mathring{u}(s))$$

is also C^k.

Now the solution u, if it exists, clearly must be the C^k function

$$u(x,y) = \beta(\gamma^{-1}(x,y)) \quad \text{or equivalently} \quad u(\gamma(s,t)) = \beta(s,t).$$

To prove that u really is a solution, we note that through any point $(x,y) \in V$ there is a characteristic curve $t \mapsto \gamma(s,t)$, and that

$$\begin{aligned}\frac{du(\gamma(s,t))}{dt} = \frac{d\beta(s,t)}{dt} &= C(\gamma(s,t)) \cdot \beta(s,t) + D(\gamma(s,t)) \\ &= C(\gamma(s,t)) \cdot u(\gamma(s,t)) + D(\gamma(s,t)),\end{aligned}$$

while we also have

$$\begin{aligned}\frac{du(\gamma(s,t))}{dt} &= u_x(\gamma(s,t)) \cdot \frac{\partial \gamma_1}{\partial t}(s,t) + u_y(\gamma(s,t)) \cdot \frac{\partial \gamma_2}{\partial t}(s,t) \\ &= u_x(\gamma(s,t)) \cdot A(\gamma(s,t)) + u_y(\gamma(s,t)) \cdot B(\gamma(s,t)),\end{aligned}$$

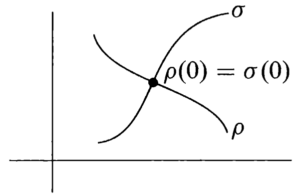
since $t \mapsto \gamma(s,t)$ is a characteristic curve. ❖

Notice that Theorem 1 involves exactly the sort of "arbitrary function" that our general considerations would lead us to expect: in a neighborhood of the "initial curve" $u(\sigma(s)) = \mathring{u}(s)$. The only requirement is that σ be nowhere tangent to a characteristic curve; we will express this by saying that σ is **free** (sometimes the term "non-characteristic" is used, but this seems a little misleading). In general, the problem of finding a solution of a PDE with an appropriate initial condition is called the "Cauchy problem" for this equation. Thus we have solved the Cauchy problem for the linear PDE (1) for any initial condition along any free curve. In particular, we can solve the Cauchy problem along the x-axis $\sigma(s) = (s,0)$ if the x-axis is free, which is equivalent to the condition that $B \neq 0$ along the x-axis. In this case we can use the given equation (1) to solve for u_y in terms of u_x along the x-axis:

$$u_y = -\frac{A}{B}u_x + \frac{C}{B}u + \frac{D}{B}.$$

If we were interested in the Cauchy problem only along the x-axis, then we could simply demand this very natural condition in our hypotheses, and not mention the characteristic curves at all; but the characteristic curves are still the most important ingredient in the proof, and their generalizations will play decisive roles in all other equations we discuss.

If our initial curve σ actually happens to be a characteristic curve (thus failing in the worst possible way to be free), then we will be unable to solve the Cauchy problem, and this inability will be manifested in the worst possible way: the possible initial condition along σ is almost uniquely determined—it is determined by the value at only one point, by the equation (4). On the other hand, if we are given an initial condition $\mathring{u}$ along σ which does satisfy (4), then there will be infinitely many solutions u with this initial condition; for we can consider any free curve ρ with $\rho(0) = \sigma(0)$, and choose any initial data ϕ along ρ



with $\phi(0) = \mathring{u}(0)$. Thus, the characteristic curves are the places where different solutions agree.

From Theorem 1 we can see immediately that an arbitrary linear first order PDE has, in common with the simple-minded equation $\partial u/\partial y = 0$, a property which sharply distinguishes it from an *ordinary* differential equation

$$u'(x) = f(x, u(x)).$$

For the ordinary differential equation, any solution u will clearly be at least one time more differentiable than f is, and if f is analytic, the solution will also be analytic (cf. DG, Vol. 1, Prob. 6-9). But there are solutions of the equation in Theorem 1 which are only C^l $(1 \le l \le \infty)$ even when A, B, C, D are C^k $(l < k \le \omega)$. For we may choose σ to be a C^k curve and $\mathring{u}$ to be a function which is C^l, but not C^{l+1}; then the solution u cannot be C^{l+1}, since its restriction to the C^k curve σ is not C^{l+1}.

We next consider the most general **quasi-linear first order PDE**

$$A(x, y, u(x, y))u_x(x, y) + B(x, y, u(x, y))u_y(x, y) = C(x, y, u(x, y)),$$

or, more briefly,

$$A(x, y, u)u_x + B(x, y, u)u_y = C(x, y, u).$$

The functions A, B, and C are now defined on $\mathbb{R}^3$, and we consider the vector field X *in* $\mathbb{R}^3$ defined by

$$X = A\frac{\partial}{\partial x} + B\frac{\partial}{\partial y} + C\frac{\partial}{\partial z}. \tag{2}$$

This vector field will be called the **characteristic vector field** of equation (1); the integral curves of X are called the **characteristic curves** equation (1). Thus $c = (c_1, c_2, c_3)$ is a characteristic curve if and only if

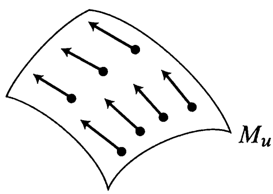
$$(3) \qquad \frac{dc_1(t)}{dt} = A(C(t)), \quad \frac{dc_2(t)}{dt} = B(c(t)), \quad \frac{dc_3(t)}{dt} = C(c(t)).$$

The slight discrepancy between this terminology and that adopted in the linear case is easily explained. Notice that if A and B depend only on x and y, then all characteristic vectors $X(x_0, y_0, z_0)$ have the same projection on the (x, y)-plane, namely $(A(x_0, y_0), B(x_0, y_0))$. So the characteristic curves of a linear equation are really the projections on the (x, y)-plane of the characteristic curves in $\mathbb{R}^3$.

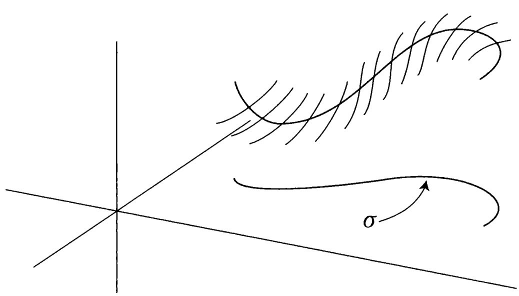
For the quasi-linear PDE (1), the characteristic curves in $\mathbb{R}^3$ have the following significance. Any C^1 function $u\colon \mathbb{R}^2 \to \mathbb{R}$ determines a surface $M_u = \{(x, y, u(x, y))\} \subset \mathbb{R}^3$, and the vector

$$(u_x(x, y),\ u_y(x, y),\ -1)$$

is normal to M_u at $(x, y, u(x, y))$. Equation (1) is therefore equivalent to saying that $X(x, y, u(x, y))$ lies in the tangent space of M_u at $(x, y, u(x, y))$. So the



characteristic vectors at the various points of M_u give a vector field on M_u. Thus M_u is the union of integral curves of this vector field; that is, M_u is the union of characteristic curves. If we are given an arbitrary initial condition $\mathring{u}$ along an initial curve σ in $\mathbb{R}^2$, then we ought to be able to construct a solution u passing through the curve $s \mapsto (\sigma_1(s), \sigma_2(s), \mathring{u}(s))$ in $\mathbb{R}^3$ simply by taking the union of the characteristic curves through all the points of this curve. We will clearly have to require that the vectors $(\sigma_1{}'(s), \sigma_2{}'(s))$ and $(A(\sigma_1(s), \sigma_2(s), \mathring{u}(s)), B(\sigma_1(s), \sigma_2(s), \mathring{u}(s)))$ are linearly independent for all s.



2. THEOREM. Let A, B, and C be C^k functions defined in an open set $U \subset \mathbb{R}^3$. Let $\sigma\colon [a,b] \to \mathbb{R}^2$ be a one-one C^k function, and $\mathring{u}\colon [a,b] \to \mathbb{R}$ a C^k function such that $(\sigma_1(s), \sigma_2(s), \mathring{u}(s)) \in U$ for all $s \in [a,b]$. Suppose moreover that

$$\sigma_1{}'(s) \cdot B(\sigma_1(s), \sigma_2(s), \mathring{u}(s)) \neq \sigma_2{}'(s) \cdot A(\sigma_1(s), \sigma_2(s), \mathring{u}(s)) \qquad \text{for all } s \in [a,b].$$

Then there is a C^k function u, defined in a neighborhood V of $\sigma([a,b])$, which satisfies the equation

$$A(x,y,u)u_x + B(x,y,u)u_y = C(x,y,u) \qquad \text{on } V, \tag{1}$$

with the initial condition

$$u(\sigma(s)) = \mathring{u}(s) \qquad \text{for all } s \in [a,b].$$

Moreover, any two functions u with this property agree on a neighborhood of $\sigma([a,b])$.

PROOF. Now there is a C^k function $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ with

$$(*) \qquad \begin{cases} \alpha(0,s,r) = r & \text{for } r \in \mathbb{R}^3 \\ \dfrac{\partial}{\partial t}\alpha_1(t,s,r) = A(\alpha(t,s,r)) \\ \dfrac{\partial}{\partial t}\alpha_2(t,s,r) = B(\alpha(t,s,r)) \\ \dfrac{\partial}{\partial t}\alpha_3(t,s,r) = C(\alpha(t,s,r)). \end{cases}$$

Let

$$\beta(s,t) = \alpha(t, s, \sigma_1(s), \sigma_2(s), \mathring{u}(s)),$$

so that β is also C^k. In particular,

$$\begin{aligned} \beta(s,0) &= (\sigma_1(s), \sigma_2(s), \mathring{u}(s)) \\ &= \bullet, \qquad \text{for short} \end{aligned}$$

[so for each s, the curve $t \mapsto \beta(s,t)$ is a characteristic curve through $\bullet$]. If we define

$$\gamma(s,t) = (\beta_1(s,t), \beta_2(s,t)) \in \mathbb{R}^2,$$

then the Jacobian of γ at $(s,0)$ is

$$\begin{pmatrix} \dfrac{\partial\beta_1}{\partial s}(s,0) & \dfrac{\partial\beta_1}{\partial t}(s,0) \\[2ex] \dfrac{\partial\beta_2}{\partial s}(s,0) & \dfrac{\partial\beta_2}{\partial t}(s,0) \end{pmatrix} = \begin{pmatrix} \sigma_1{}'(s) & \dfrac{\partial\alpha_1}{\partial t}(0,s,\bullet) \\[2ex] \sigma_2{}'(s) & \dfrac{\partial\alpha_2}{\partial t}(0,s,\bullet) \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1{}'(s) & A(\bullet) \\ \sigma_2'(s) & B(\bullet) \end{pmatrix} \qquad \text{by } (*),$$

assumed to be nonsingular. So if ε is sufficiently small, $\gamma\colon [a,b]\times(-\varepsilon,\varepsilon)\to\mathbb{R}^2$ is a C^k diffeomorphism onto a neighborhood V of $\sigma([a,b])$.

The solution u, if it exists, clearly must be the C^k function

$$u(x,y) = \beta_3(\gamma^{-1}(x,y)) \quad \text{or equivalently} \quad u(\gamma(s,t)) = \beta_3(s,t).$$

To prove that u is a solution, we note that for any point $(x,y)\in V$, there is a characteristic curve $t\mapsto\beta(s,t)$ through $(x,y,u(x,y))$, and that

$$\frac{du(\gamma(s,t))}{dt} = \frac{d\beta_3(s,t)}{dt} = C(\beta(s,t)) \qquad \text{by } (*),$$

while we also have

$$\begin{aligned}\frac{du(\gamma(s,t))}{dt} &= u_x(\gamma(s,t))\cdot\frac{\partial\gamma_1}{\partial t}(s,t) + u_y(\gamma(s,t))\cdot\frac{\partial\gamma_2}{\partial t}(s,t)\\ &= u_x(\gamma(s,t))\cdot\frac{\partial\beta_1}{\partial t}(s,t) + u_y(\gamma(s,t))\cdot\frac{\partial\beta_2}{\partial t}(s,t)\\ &\qquad\text{by definition of } \gamma\\ &= u_x(\gamma(s,t))\cdot A(\beta(s,t)) + u_y(\gamma(s,t))\cdot B(\beta(s,t)) \qquad \text{by } (*). \end{aligned}$$ ❖

We will say that the initial curve σ is **free for the initial condition $\mathring{u}$** when it satisfies

$$\sigma_1{}'(s)\cdot B(\sigma_1(s),\sigma_2(s),\mathring{u}(s)) \neq \sigma_2{}'(s)\cdot A(\sigma_1(s),\sigma_2(s),\mathring{u}(s)).$$

Thus we can solve the Cauchy problem for a quasi-linear PDE (1) for any initial condition along any curve which is free for this initial condition. (In the linear case things are simpler, since the condition that σ be free doesn't depend on the initial condition $\mathring{u}$.)

The worst way in which the initial curve $\sigma\colon [a,b]\to\mathbb{R}^2$ can fail to be free for the initial condition $\mathring{u}$ is when the vectors $\sigma'(s) = (\sigma_1{}'(s),\sigma_2{}'(s))$ and the vectors $(A(\sigma_1(s),\sigma_2(s),\mathring{u}(s)), B(\sigma_1(s),\sigma_2(s),\mathring{u}(s))) = (A(\bullet),B(\bullet))$ are everywhere linearly dependent. In this case, it is customary to say that σ is **characteristic** for $\mathring{u}$; this does *not* mean that σ is a characteristic curve (indeed, σ isn't even a curve in $\mathbb{R}^3$). If we assume that σ is an imbedding, then σ is characteristic if and only if $(A(\bullet), B(\bullet))$ is always a multiple of the tangent vector $\sigma'(s)$; by reparameterizing σ we can then arrange that

$$(A(\bullet), B(\bullet)) = \sigma'(s).$$

Then if $\mathring{u}$ is to be the initial condition for a solution u of (1) we must have

$$\begin{aligned} C(\bullet) &= {\sigma_1}'(s) \cdot u_x(\sigma(s)) + {\sigma_2}'(s) \cdot u_y(\sigma(s)) \\ &= \frac{d}{ds} u(\sigma(s)) = \frac{d}{ds} \mathring{u}(s). \end{aligned}$$

This shows that the reparameterized curve $s \mapsto (\sigma_1(s), \sigma_2(s), \mathring{u}(s))$ must be a characteristic curve; equivalently, the original curve $s \mapsto (\sigma_1(s), \sigma_2(s), \mathring{u}(s))$ must be a characteristic curve up to reparameterization in order for the Cauchy problem to be solvable when σ is characteristic for $\mathring{u}$. If our initial condition $\mathring{u}$ does have this property, then there will be infinitely many solutions u with this initial condition along σ. The characteristic curves in $\mathbb{R}^3$ are the places where the graphs of different solutions intersect; the projections of the characteristic curves onto $\mathbb{R}^2$ are the places where different solutions agree.

It should be clear once again that a quasi-linear first order PDE has solutions which are less differentiable than its coefficients.

We are now ready to consider the most general *first order* PDE

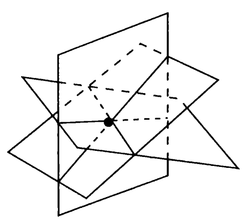
$$(1) \qquad F(x, y, u, p, q) = F(x, y, u(x, y), u_x(x, y), u_y(x, y)) = 0.$$

This equation can also be reduced to a system of ordinary differential equations, but in this case the system will involve *five* functions; the geometric analysis will be correspondingly more complicated.

At each point $(x_0, y_0, z_0) \in \mathbb{R}^3$, we can consider the set of all vectors $(a, b, -1)$ with

$$F(x_0, y_0, z_0, a, b) = 0,$$

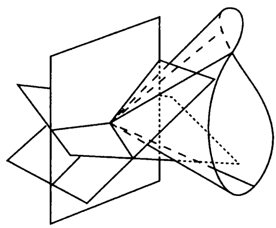
and the corresponding family $\mathcal{F}(x_0, y_0, z_0)$ of planes perpendicular to such vectors. If u is a solution of (1), and M_u is the surface $M_u = \{(x, y, u(x, y))\}$,



then the tangent space of M_u at $(x_0, y_0, u(x_0, y_0))$ is a member of the family $\mathcal{F}(x_0, y_0, u(x_0, y_0))$. In order to describe this situation more geometrically, we would like to have a more geometric way of describing the families $\mathcal{F}(x_0, y_0, z_0)$. Now the relation

$$F(x_0, y_0, z_0, a, b) = 0$$

is one equation in the two unknowns, a and b, so $\mathcal{F}(x_0, y_0, z_0)$ ought to be a one-parameter family of planes; this suggests that there is a cone $K(x_0, y_0, z_0)$, having its vertex at (x_0, y_0, z_0), such that a plane P is in $\mathcal{F}(x_0, y_0, z_0)$ if and only if P is tangent to $K(x_0, y_0, z_0)$ along a generator of this cone. If we consider a



quasi-linear equation

$$F(x, y, u, p, q) = A(x, y, u) \cdot p + B(x, y, u) \cdot q - C(x, y, u) = 0,$$

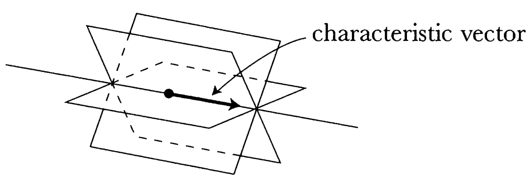
we immediately see that this is not always so. For in this case, the family $\mathcal{F}(x_0, y_0, z_0)$ consists of planes perpendicular to vectors $(a, b, -1)$ with

$$a \cdot A(x_0, y_0, z_0) + b \cdot B(x_0, y_0, z_0) = C(x_0, y_0, z_0).$$

These planes all contain the characteristic vector

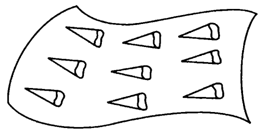
$$(A(x_0, y_0, z_0), B(x_0, y_0, z_0), C(x_0, y_0, z_0)).$$

Thus our "cone" degenerates into a straight line through (x_0, y_0, z_0), pointing in the direction of the characteristic vector at that point. Clearly things might



be even messier if the analytic properties of the function F are sufficiently nasty.

Despite these difficulties, we can obtain a great deal of geometric motivation by temporarily pretending that each family $\mathcal{F}(x_0, y_0, z_0)$ *is* determined by a cone $K(x_0, y_0, z_0)$, which happens to degenerate to a straight line in the case of a quasi-linear equation. This semi-mythical cone is called the **Monge cone** at (x_0, y_0, z_0). Having accepted this fiction, we can now imagine a field of cones in $\mathbb{R}^3$; a C^1 function $u\colon \mathbb{R}^2 \to \mathbb{R}$ is a solution of equation (1) if and only



if the corresponding surface $M_u = \{(x, y, u(x, y))\}$ is tangent to the Monge cone $K(x_0, y_0, u(x_0, y_0))$ at each point $(x_0, y_0, u(x_0, y_0))$. This gives us a field of directions at each point of M_u, namely the direction which lies along a generator of the Monge cone at that point. The integral manifolds of this field of directions could be called the "characteristic curves of the solution u". This definition is easily seen to be compatible with the one already given in the quasi-linear case, where the Monge cones degenerate to straight lines: for these straight lines must be the field of directions for any solution u, and the "characteristic curves of the solution u" are simply those characteristic curves of the quasi-linear equation which happen to lie on M_u. But in the general case, we cannot write $\mathbb{R}^3$ as a disjoint union of curves in such a way that each M_u is the union of a certain subset of these curves; we cannot describe the "characteristic curves of a solution u" at all until we already know u. This might make the concept seem rather useless, but the requisite supplementary considerations will appear quite naturally when we seek an analytic description of these geometric pictures.

How would we go about finding an analytic description of the Monge cone? Addendum 8B [and especially the material in DG, Vol. 3, Chap. 3, Addendum mentioned at the end] suggests that the Monge cone $K(x_0, y_0, z_0)$ should be the envelope of the family of planes $\mathcal{F}(x_0, y_0, z_0)$; geometrically, the generators of $K(x_0, y_0, z_0)$ should be the limits of the intersections of two planes of the family $\mathcal{F}(x_0, y_0, z_0)$, the limit being formed as the two planes approach each other. Until we explicitly say the opposite, everything we now do will be based on the assumption that these limits really exist; the ensuing discussion is consequently merely a route to discovery, and does not purport to prove anything.

Let us assume for the moment that the equation

$$F(x_0, y_0, z_0, a, b) = 0$$

can be solved for b in terms of a. In other words, assume there is a function ϕ with

$$\text{(i)} \qquad F(x_0, y_0, z_0, a, \phi(a)) = 0.$$

One plane of the family $\mathscr{F}(x_0, y_0, z_0)$ may be described by the equation

$$z - z_0 = a(x - x_0) + \phi(a)(y - y_0).$$

A nearby plane may be described by the equation

$$z - z_0 = (a + h)(x - x_0) + \phi(a + h)(y - y_0).$$

The points (x, y, z) in the intersection then satisfy

$$0 = h(x - x_0) + [\phi(a + h) - \phi(a)](y - y_0),$$

and hence

$$0 = (x - x_0) + \left[\frac{\phi(a + h) - \phi(a)}{h}\right](y - y_0).$$

Therefore points in the limiting intersection ought to satisfy

$$\text{(ii)} \qquad \begin{cases} z - z_0 = a(x - x_0) + \phi(a)(y - y_0) \\ \qquad 0 = (x - x_0) + \phi'(a)(y - y_0). \end{cases}$$

On the other hand, equation (i) shows that

$$\begin{aligned} 0 &= \frac{d}{da} F(x_0, y_0, z_0, a, \phi(a)) \\ &= F_p(x_0, y_0, z_0, a, \phi(a)) + \phi'(a) \cdot F_q(x_0, y_0, z_0, a, \phi(a)), \end{aligned}$$

and hence

$$\text{(iii)} \qquad \phi'(a) = -\frac{F_p(x_0, y_0, z_0, a, \phi(a))}{F_q(x_0, y_0, z_0, a, \phi(a))}.$$

From (ii) and (iii) we find that the points (x, y, z) on the Monge cone $K(x_0, y_0, z_0)$ should satisfy

$$\text{(iv)} \qquad \begin{cases} z - z_0 = a(x - x_0) + b(y - y_0), & \text{where } a \text{ and } b \text{ are numbers such that:} \\ F(x_0, y_0, z_0, a, b) = 0 & \\ \dfrac{x - x_0}{F_p} = \dfrac{y - y_0}{F_q} & [F_p \text{ and } F_q \text{ evaluated at } (x_0, y_0, z_0, a, b)]. \end{cases}$$

Now consider a solution u of (1), and let

$$z_0 = u(x_0, y_0), \qquad p_0 = u_x(x_0, y_0), \qquad q_0 = u_y(x_0, y_0).$$

The tangent plane of M_u at (x_0, y_0, z_0) consists of points (x, y, z) satisfying

$$z - z_0 = p_0(x - x_0) + q_0(y - y_0).$$

Equations (iv) show that points (x, y, z) which are on both this tangent plane and the Monge cone $K(x_0, y_0, z_0)$ ought to satisfy

(v) $$\frac{x - x_0}{F_p} = \frac{y - y_0}{F_q} = \frac{z - z_0}{p_0 F_p + q_0 F_q}$$
$$[F_p \text{ and } F_q \text{ evaluated at } (x_0, y_0, z_0, p_0, q_0)].$$

Therefore, these points ought to lie along the line through (x_0, y_0, z_0) with direction

$$(F_p, F_q, p_0 F_p + q_0 F_q) \qquad [F_p \text{ and } F_q \text{ evaluated at } (x_0, y_0, z_0, p_0, q_0)].$$

We have finally reached the stage where we can make a perfectly sensible definition, involving no assumptions at all. Let u be a solution of (1), and for a point (x_0, y_0), define z_0, p_0, and q_0 as before. We then define the **characteristic vector of** u **at** (x_0, y_0) to be the vector

(2) $$X(u; x_0, y_0) = (F_p, F_q, p_0 F_p + q_0 F_q),$$

where F_p and F_q are to be evaluated at $(x_0, y_0, z_0, p_0, q_0)$; this vector is to be considered as an element of $\mathbb{R}^3{}_{(x_0,y_0,z_0)}$. If $M_u = \{(x, y, u(x, y))\}$, then the tangent plane of M_u at (x_0, y_0, z_0) is perpendicular to the vector $(p_0, q_0, -1)$. The vector $X(u; x_0, y_0)$ clearly has this property, so every characteristic vector of u is tangent to M_u, and the set of all characteristic vectors of u forms a vector field on M_u. The integral curves of this vector field are called the **characteristic curves of the solution** u, and they are clearly curves on M_u.

A characteristic curve c of u is thus a curve in $\mathbb{R}^3$ satisfying the equations

(3) $$\begin{cases} \dfrac{dc_1(t)}{dt} = F_p(\bullet) \\ \dfrac{dc_2(t)}{dt} = F_q(\bullet) \\ \dfrac{dc_3(t)}{dt} = u_x(c_1(t), c_2(t)) \cdot F_p(\bullet) + u_y(c_1(t), c_2(t)) \cdot F_q(\bullet) \\ \qquad \text{where } \bullet = (c_1(t), c_2(t), c_3(t), u_x(c_1(t), c_2(t)), u_y(c_1(t), c_2(t))). \end{cases}$$

Now if we assume that u is C^2, then we can also obtain equations for the partials $u_x(c_1(t), c_2(t))$ and $u_y(c_1(t), c_2(t))$. For equations (3) allow us to write

$$(4)\quad \begin{cases} \dfrac{du_x(c_1(t), c_2(t))}{dt} = u_{xx}(c_1(t), c_2(t))\dfrac{dc_1(t)}{dt} + u_{xy}(c_1(t), c_2(t))\dfrac{dc_2(t)}{dt} \\ \qquad\qquad\qquad\quad = u_{xx}(c_1(t), c_2(t))F_p(\bullet) + u_{xy}(c_1(t), c_2(t))F_q(\bullet) \\ \dfrac{du_y(c_1(t), c_2(t))}{dt} = u_{yx}(c_1(t), c_2(t))F_p(\bullet) + u_{yy}(c_1(t), c_2(t))F_q(\bullet). \end{cases}$$

On the other hand, since u satisfies

$$F(x, y, u(x, y), u_x(x, y), u_y(x, y)) = 0,$$

we also have

$$(5)\quad \begin{cases} F_x + u_x F_u + u_{xx} F_p + u_{yx} F_q = 0 \\ F_y + u_y F_u + u_{xy} F_p + u_{yy} F_q = 0, \end{cases}$$

where all partials of F are evaluated at $(x, y, u(x, y), u_x(x, y), u_y(x, y))$. Thus equations (4) become

$$(6)\quad \begin{cases} \dfrac{du_x(c_1(t), c_2(t))}{dt} = -F_x(\bullet) - u_x(c_1(t), c_2(t)) \cdot F_u(\bullet) \\ \dfrac{du_y(c_1(t), c_2(t))}{dt} = -F_y(\bullet) - u_y(c_1(t), c_2(t)) \cdot F_u(\bullet). \end{cases}$$

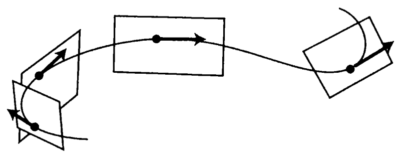
Let us now define a curve Γ in $\mathbb{R}^5$ by

$$(7)\qquad \Gamma(t) = (c_1(t), c_2(t), c_3(t), u_x(c_1(t), c_2(t)), u_y(c_1(t), c_2(t))).$$

Then equations (3) and (6) may be written

$$(8)\quad \begin{cases} \dfrac{d\Gamma_1(t)}{dt} = F_p(\Gamma(t)) \\ \dfrac{d\Gamma_2(t)}{dt} = F_q(\Gamma(t)) \\ \dfrac{d\Gamma_3(t)}{dt} = \Gamma_4(t) \cdot F_p(\Gamma(t)) + \Gamma_5(t) \cdot F_q(\Gamma(t)) \\ \dfrac{d\Gamma_4(t)}{dt} = -F_x(\Gamma(t)) - \Gamma_4(t) \cdot F_u(\Gamma(t)) \\ \dfrac{d\Gamma_5(t)}{dt} = -F_y(\Gamma(t)) - \Gamma_5(t) \cdot F_u(\Gamma(t)). \end{cases}$$

Now although the curve Γ was defined in terms of a solution u, the final equations (8) involve *only* the original equation (1). This will allow us to define geometrically meaningful objects which do not depend on knowing a solution u. We may regard a point $(x_0, y_0, z_0, a, b) \in \mathbb{R}^5$ as a plane in the tangent space $\mathbb{R}^3{}_{(x_0,y_0,z_0)}$, namely, as the plane perpendicular to the vector $(a, b, -1)$. A curve Γ in $\mathbb{R}^5$ may then be regarded as a family of planes, the plane at time t being in the tangent space of $\mathbb{R}^3$ at $c(t) = (\Gamma_1(t), \Gamma_2(t), \Gamma_3(t))$; it will be convenient to refer to this curve c as the **base curve** of Γ. An arbitrary curve Γ is called



a **strip** if the tangent vector $c'(t)$ of the base curve c always lies in the plane determined by Γ at time t. This means that

$$c'(t) = (\Gamma_1{}'(t), \Gamma_2{}'(t), \Gamma_3{}'(t)) \quad \text{is perpendicular to} \quad (\Gamma_4(t), \Gamma_5(t), -1).$$

So Γ is a strip if and only if it satisfies the **strip condition**:

$$\frac{d\Gamma_3(t)}{dt} = \Gamma_4(t)\frac{d\Gamma_1(t)}{dt} + \Gamma_5(t)\frac{d\Gamma_2(t)}{dt}. \tag{9}$$

Notice that any solution of (8) is automatically a strip. A curve Γ will be called a **characteristic strip** of the PDE (1) if Γ satisfies (8) and also

$$F(\Gamma(t)) = F(\Gamma_1(t), \Gamma_2(t), \Gamma_3(t), \Gamma_4(t), \Gamma_5(t)) = 0. \tag{10}$$

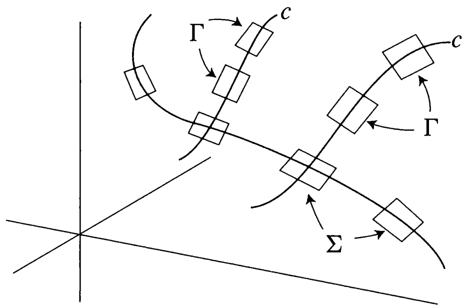
This last restriction is not as stringent as it might first seem, for if Γ satisfies (8), then

$$\begin{aligned}
\frac{d}{dt}F(\Gamma(t)) &= F_x\frac{d\Gamma_1(t)}{dt} + \cdots + F_q\frac{d\Gamma_5(t)}{dt} \\
&\qquad [\text{all partials of } F \text{ evaluated at } \Gamma(t)] \\
&= F_xF_p + F_yF_q + F_z\cdot(\Gamma_4(t)F_p + \Gamma_5(t)F_q) \\
&\qquad + F_p\cdot(-F_x - \Gamma_4(t)F_z) + F_q\cdot(-F_y - \Gamma_5(t)F_z) \\
&= 0.
\end{aligned} \tag{11}$$

So if Γ satisfies (8) and also satisfies (10) for one t, *then it satisfies* (10) *for all* t, *and is consequently a characteristic strip.*

Now how are characteristic strips related to solutions? We have seen that if u is a solution of (1), then M_u is the union of certain characteristic curves [solutions of (3)]. Moreover, if c is a characteristic curve, then the set of tangent planes of M_u along c gives the curve Γ of equation (7), which is a characteristic strip. So M_u is the union of base curves of characteristic strips.

Now suppose that we have an arbitrary curve Σ in $\mathbb{R}^5$, with base curve σ, and that $F(\Sigma(s)) = 0$ for all s. There is a unique solution of (8) through each point $\Sigma(s)$, and by the remark after equation (11), this solution is a characteristic strip. We thus obtain a family of characteristic strips Γ. The union of the



corresponding base curves c is a surface M_u, containing the base curve σ. Is it reasonable to suppose now that u is a solution of (1)? The answer is no, for there is clearly no hope unless Σ *is also a strip*. When *this* condition is satisfied, then everything works out. We will prove that if $\sigma\colon [a,b] \to \mathbb{R}^2$ is a given curve, $\mathring{u}\colon [a,b] \to \mathbb{R}$ is a given function, and $\mathring{p}, \mathring{q}\colon [a,b] \to \mathbb{R}$ are two functions satisfying

(a) $$F(\Sigma(s)) = F(\sigma_1(s), \sigma_2(s), \mathring{u}(s), \mathring{p}(s), \mathring{q}(s)) = 0,$$

and the strip condition

(b) $$\frac{d\mathring{u}(s)}{ds} = \mathring{p}(s)\frac{d\sigma_1(s)}{ds} + \mathring{q}(s)\frac{d\sigma_2(s)}{ds},$$

then there is a unique solution u of (1) satisfying

$$u(\sigma(s)) = \mathring{u}(s), \quad u_x(\sigma(s)) = \mathring{p}(s), \quad u_y(\sigma(s)) = \mathring{q}(s)$$

[naturally, (b) is a necessary consequence of these equations]. We will clearly have to assume that $\sigma'(s)$ is linearly independent of the vector obtained by projecting the characteristic vector (2) on the (x, y)-plane. In other words, we will have to require that $\sigma'(s)$ and $(F_p(\Sigma(s)), F_q(\Sigma(s)))$ are linearly independent, or that

(c) $$\sigma_1{}'(s) \cdot F_q(\Sigma(s)) \neq \sigma_2{}'(s) \cdot F_p(\Sigma(s)).$$

Before we proceed to prove the theorem, we should insert a remark about the hypotheses, which will involve σ, $\mathring{u}$, $\mathring{p}$, and $\mathring{q}$ satisfying (a)–(c). At first sight, we seem to be contradicting our basic philosophy about first order equations, for we seem to be saying that we can arbitrarily specify not only the values $\mathring{u}$ of u along σ, but *also* the values $\mathring{p}$ and $\mathring{q}$ of u_x and u_y along σ. This is not really the case, for $\mathring{p}$ and $\mathring{q}$ are practically determined by the equations (a) and (b) which they must satisfy. This is most apparent when our initial curve σ is the x-axis, $\sigma(s) = (s, 0)$. Then equation (b) already determines $\mathring{p}$. Moreover, condition (c) says that $F_q \neq 0$ along $\{(s, 0, \mathring{u}(s), \mathring{p}(s), \mathring{q}(s))\}$, so the implicit function theorem shows that equation (a) can be solved for $\mathring{q}(s)$ in terms of $\mathring{p}(s)$—there is a function ϕ with

$$F(s, 0, \mathring{u}(s), \mathring{p}(s), \phi(\mathring{p}(s))) = 0.$$

Of course, there may be several possible ϕ, but once $\mathring{q}(0)$ is determined, there will be only one continuous choice of $\mathring{q}$ satisfying (a). [In the quasi-linear case, $\mathring{q}(s)$ will actually be uniquely determined.] It is not hard to see that a similar situation prevails when σ is any curve satisfying (c): we are essentially specifying only the values $\mathring{u}$ of u along σ, and then making certain that we have a continuous choice of the limited possibilities for $\mathring{p}$ and $\mathring{q}$. In order to emphasize this point we will refer to $(\mathring{u}, \mathring{p}, \mathring{q})$ as "initial data", rather than as initial conditions.

3. THEOREM. Let F be a function of class C^k, $k \geq 3$, defined in an open set $U \subset \mathbb{R}^5$. Let $\sigma\colon [a, b] \to \mathbb{R}^2$ be a one-one C^{k-1} function, and let $\mathring{u}, \mathring{p}, \mathring{q}\colon [a, b] \to \mathbb{R}$ be C^{k-1} functions such that for all $s \in [a, b]$ we have

(a) $$\Sigma(s) = (\sigma_1(s), \sigma_2(s), \mathring{u}(s), \mathring{p}(s), \mathring{q}(s)) \in U \quad \text{and} \quad F(\Sigma(s)) = 0,$$

(b) $$\frac{d\mathring{u}(s)}{ds} = \mathring{p}(s)\frac{d\sigma_1(s)}{ds} + \mathring{q}(s)\frac{d\sigma_2(s)}{ds},$$

(c) $$\sigma_1{}'(s) \cdot F_q(\Sigma(s)) \neq \sigma_2{}'(s) \cdot F_p(\Sigma(s)).$$

Then there is a C^{k-1} function u, defined in a neighborhood V of $\sigma([a, b])$, which satisfies the equation

$$F(x, y, u(x, y), u_x(x, y), u_y(x, y)) = 0 \qquad \text{on } V$$

and also

$$u(\sigma(s)) = \mathring{u}(s), \quad u_x(\sigma(s)) = \mathring{p}(s), \quad u_y(\sigma(s)) = \mathring{q}(s), \qquad \text{for } s \in [a, b].$$

Moreover, any two functions u with this property agree on a neighborhood of $\sigma([a, b])$.

PROOF. As in the proof of Theorems 1 and 2, we use DG, Vol. 1, Prob. 5-5 to conclude that there is a C^{k-1} function $\alpha = (\alpha_1, \dots, \alpha_5)$ with

$$(*)\quad \begin{cases} \alpha(0,s,r) = r \qquad \text{for } r \in \mathbb{R}^5 \\ \dfrac{\partial}{\partial t}\alpha_1(t,s,r) = F_p(\alpha(t,s,r)) \\ \dfrac{\partial}{\partial t}\alpha_2(t,s,r) = F_q(\alpha(t,s,r)) \\ \dfrac{\partial}{\partial t}\alpha_3(t,s,r) = \alpha_4(t,s,r)\cdot F_p(\alpha(t,s,r)) + \alpha_5(t,s,r)\cdot F_q(\alpha(t,s,r)) \\ \dfrac{\partial}{\partial t}\alpha_4(t,s,r) = -F_x(\alpha(t,s,r)) - \alpha_4(t,s,r)\cdot F_u(\alpha(t,s,r)) \\ \dfrac{\partial}{\partial t}\alpha_5(t,s,r) = -F_y(\alpha(t,s,r)) - \alpha_5(t,s,r)\cdot F_u(\alpha(t,s,r)). \end{cases}$$

Let

$$\beta(s,t) = \alpha(t,s,\sigma_1(s),\sigma_2(s),\mathring{u}(s),\mathring{p}(s),\mathring{q}(s)),$$

so that β is also C^{k-1}. In particular,

$$\beta(s,0) = (\sigma_1(s),\sigma_2(s),\mathring{u}(s),\mathring{p}(s),\mathring{q}(s)) = \Sigma(s).$$

If we define

$$\gamma(s,t) = (\beta_1(s,t),\beta_2(s,t)) \in \mathbb{R}^2,$$

then the Jacobian of γ at $(s,0)$ is

$$\begin{aligned} \begin{pmatrix} \dfrac{\partial\beta_1}{\partial s}(s,0) & \dfrac{\partial\beta_1}{\partial t}(s,0) \\[2ex] \dfrac{\partial\beta_2}{\partial s}(s,0) & \dfrac{\partial\beta_2}{\partial t}(s,0) \end{pmatrix} &= \begin{pmatrix} {\sigma_1}'(s) & \dfrac{\partial\alpha_1}{\partial t}(0,s,\Sigma(s)) \\[2ex] {\sigma_2}'(s) & \dfrac{\partial\alpha_2}{\partial t}(0,s,\Sigma(s)) \end{pmatrix} \\ &= \begin{pmatrix} {\sigma_1}'(s) & F_p(\Sigma(s)) \\ {\sigma_2}'(s) & F_q(\Sigma(s)) \end{pmatrix} \qquad \text{by } (*), \end{aligned}$$

assumed nonsingular. So if ε is sufficiently small, then $\gamma\colon [a,b]\times(-\varepsilon,\varepsilon) \to \mathbb{R}^2$ is a C^{k-1} diffeomorphism onto a neighborhood V of $\sigma([a,b])$.

The solution u, if it exists, clearly must be the C^{k-1} function

$$u(x,y) = \beta_3(\gamma^{-1}(x,y)) \qquad \text{or equivalently} \qquad u(\gamma(s,t)) = \beta_3(s,t).$$

We claim that

$$u_x(\gamma(s,t)) = \beta_4(s,t) \qquad \text{and} \qquad u_y(\gamma(s,t)) = \beta_5(s,t).$$

This will prove that

$$F(x,y,u(x,y),u_x(x,y),u_y(x,y)) = 0;$$

for we have already seen (equation 11) that $F(\alpha(t,s,r))$ is constant for fixed s and r, while $F(\alpha(0,s,\Sigma(s))) = 0$ by (a), so that we will have

$$\begin{aligned}
0 &= F(\alpha(t,s,\Sigma(s))) \\
&= F(\beta(s,t)) = F(\beta_1(s,t),\beta_2(s,t),\beta_3(s,t),\beta_4(s,t),\beta_5(s,t)) \\
&= F(\gamma(s,t),u(\gamma(s,t)),u_x(\gamma(s,t)),u_y(\gamma(s,t))).
\end{aligned}$$

To prove the claim, we consider the function

$$\Delta = \frac{\partial\beta_3}{\partial s} - \beta_4 \cdot \frac{\partial\beta_1}{\partial s} - \beta_5 \cdot \frac{\partial\beta_2}{\partial s}.$$

We have

$$\begin{aligned}
\Delta(s,0) &= \frac{d\mathring{u}(s)}{ds} - \mathring{p}(s) \cdot \frac{d\sigma_1(s)}{ds} - \mathring{q}(s) \cdot \frac{d\sigma_2(s)}{ds} \\
&= 0 \qquad \text{by (b).}
\end{aligned}$$

Moreover,

$$\begin{aligned}
\frac{\partial\Delta}{\partial t} &= \frac{\partial^2\beta_3}{\partial s\partial t} - \frac{\partial\beta_4}{\partial t}\frac{\partial\beta_1}{\partial s} - \frac{\partial\beta_5}{\partial t}\frac{\partial\beta_2}{\partial s} - \beta_4 \cdot \frac{\partial^2\beta_1}{\partial s\partial t} - \beta_5 \cdot \frac{\partial^2\beta_2}{\partial s\partial t} \\
&= \frac{\partial}{\partial s}\left(\frac{\partial\beta_3}{\partial t} - \beta_4 \cdot \frac{\partial\beta_1}{\partial t} - \beta_5 \cdot \frac{\partial\beta_2}{\partial t}\right) \\
&\qquad + \frac{\partial\beta_4}{\partial s}\frac{\partial\beta_1}{\partial t} + \frac{\partial\beta_5}{\partial s}\frac{\partial\beta_2}{\partial t} - \frac{\partial\beta_4}{\partial t}\frac{\partial\beta_1}{\partial s} - \frac{\partial\beta_5}{\partial t}\frac{\partial\beta_2}{\partial s} \\
&= 0 + F_p \cdot \frac{\partial\beta_4}{\partial s} + F_q \cdot \frac{\partial\beta_5}{\partial s} + (F_x + F_u\beta_4)\frac{\partial\beta_1}{\partial s} + (F_y + F_u\beta_5)\frac{\partial\beta_2}{\partial s} \\
&\qquad \text{by } (*) \qquad [\text{where all partials of } F \text{ are evaluated at } \beta(s,t)] \\
&= F_x \cdot \frac{\partial\beta_1}{\partial s} + F_y \cdot \frac{\partial\beta_2}{\partial s} + F_u \cdot \frac{\partial\beta_3}{\partial s} + F_p \cdot \frac{\partial\beta_4}{\partial s} + F_q \cdot \frac{\partial\beta_5}{\partial s} \\
&\qquad\qquad - F_u \cdot \left(\frac{\partial\beta_3}{\partial s} - \beta_4 \cdot \frac{\partial\beta_1}{\partial s} - \beta_5 \cdot \frac{\partial\beta_2}{\partial s}\right) \\
&= \frac{\partial}{\partial s}(F(\beta(s,t))) - F_u \cdot \Delta \\
&= -F_u \cdot \Delta,
\end{aligned}$$

since we have already seen that $F(\beta(s,t)) = 0$. Now for each fixed s, we have an ordinary differential equation

$$\frac{\partial \Delta}{\partial t} = -F_u \cdot \Delta,$$

with the initial condition

$$\Delta(s,0) = 0,$$

so the unique solution is $\Delta(s,t) = 0$. In other words, we have shown that

$$\frac{\partial \beta_3}{\partial s} = \beta_4 \cdot \frac{\partial \beta_1}{\partial s} + \beta_5 \cdot \frac{\partial \beta_2}{\partial s}.$$

Also

$$\frac{\partial \beta_3}{\partial t} = \beta_4 \cdot \frac{\partial \beta_1}{\partial t} + \beta_5 \cdot \frac{\partial \beta_2}{\partial t} \qquad \text{by } (*).$$

On the other hand, differentiating the definition $u(\gamma(s,t)) = \beta_3(s,t)$ gives

$$\begin{aligned}
\frac{\partial \beta_3}{\partial s} &= u_x(\gamma(s,t)) \cdot \frac{\partial \beta_1}{\partial s} + u_y(\gamma(s,t)) \cdot \frac{\partial \beta_2}{\partial s} \\
\frac{\partial \beta_3}{\partial t} &= u_y(\gamma(s,t)) \cdot \frac{\partial \beta_1}{\partial t} + u_y(\gamma(s,t)) \cdot \frac{\partial \beta_2}{\partial t}.
\end{aligned}$$

These last four equations give two solutions for two linear equations in two unknowns, whose determinant

$$\det \begin{pmatrix} \dfrac{\partial \beta_1}{\partial s} & \dfrac{\partial \beta_2}{\partial s} \\[2ex] \dfrac{\partial \beta_1}{\partial t} & \dfrac{\partial \beta_2}{\partial t} \end{pmatrix}$$

is $\neq 0$ for $(s,t) \in [a,b] \times (-\varepsilon, \varepsilon)$. So the two solutions must be the same, i.e.,

$$u_x(\gamma(s,t)) = \beta_4(s,t) \qquad \text{and} \qquad u_y(\gamma(s,t)) = \beta_5(s,t),$$

as desired. ❖

We will say that the initial curve σ is **free for the initial data** $\mathring{u}, \mathring{p}, \mathring{q}$ when condition (c) in Theorem 3 is satisfied. Thus we can solve the Cauchy problem for a first order PDE (1) for any initial strip $\Sigma = (\sigma_1, \sigma_2, \mathring{u}, \mathring{p}, \mathring{q})$ for which the initial curve σ is free for the initial data $\mathring{u}, \mathring{p}, \mathring{q}$.

Again we consider the case where our initial curve σ fails to be free for the initial data $\mathring{u}, \mathring{p}, \mathring{q}$ in the worst possible way, namely when $\sigma'(s)$ and $(F_p(\Sigma(s)), F_q(\Sigma(s)))$ are everywhere linearly dependent. Once again we say that σ is **characteristic** for $\mathring{u}, \mathring{p}, \mathring{q}$. Assuming that σ is an imbedding, we can reparameterize σ so that $\sigma'(s) = (F_p(\Sigma(s)), F_q(\Sigma(s)))$. This gives us the first two equations in (8) for the curve $(\sigma_1, \sigma_2, \mathring{u}, \mathring{p}, \mathring{q})$. The third equation of (8) is just the strip condition (b). The argument on page 682 shows that these three equations imply the last two if there is a solution u of (1) with

$$u(\sigma(s)) = \mathring{u}(s), \quad u_x(\sigma(s)) = \mathring{p}(s), \quad u_y(\sigma(s)) = \mathring{q}(s).$$

So when σ is characteristic, the Cauchy problem is solvable for the initial data $\mathring{u}, \mathring{p}, \mathring{q}$ along σ only if $(\sigma_1, \sigma_2, \mathring{u}, \mathring{p}, \mathring{q})$ is a characteristic strip. When this is the case, there will be infinitely many solutions with this initial data along σ. The base curves of characteristic strips are the intersection curves of the graphs of different solutions meeting tangentially.

We can now describe the situation for first order PDE's in n variables very easily, without bothering to write down all the results as formal theorems. Consider first the quasi-linear PDE

$$\sum_{i=1}^{n} A_i(x_1, \ldots, x_n, u) \cdot u_{x_i} = C(x_1, \ldots, x_n, u).$$

The **characteristic vector field** of this equation is the vector field X in $\mathbb{R}^{n+1}$ defined by

$$X = \sum_{i=1}^{n} A_i \frac{\partial}{\partial x_i} + C \frac{\partial}{\partial z};$$

the integral curves of X are the **characteristic curves** of the equation. As in the case $n = 2$, it is clear that if u is a solution of (1), then the hypersurface

$$M_u = \{(x_1, \ldots, x_n, u(x_1, \ldots, x_n))\} \subset \mathbb{R}^{n+1}$$

is a union of characteristic curves. Now suppose we are given a one-one map

$$\sigma \colon \mathcal{D} \to \mathbb{R}^n,$$

where $\mathcal{D} \subset \mathbb{R}^{n-1}$ is a compact $(n-1)$-dimensional manifold-with-boundary, and a function $\mathring{u} \colon \mathcal{D} \to \mathbb{R}$. We can produce a solution u of (1) with

$$u(\sigma(s)) = \mathring{u}(s) \qquad \text{for all } s \in \mathcal{D}$$

by taking the union of the characteristic curves through all points $(\sigma(s), \mathring{u}(s)) \in \mathbb{R}^{n+1}$. The proof is exactly analogous to the proof of Theorem 2, except that we will now require that the matrix

$$\begin{pmatrix} D_1\sigma_1(s) & & D_{n-1}\sigma_1(s) & A_1(\sigma(s), \mathring{u}(s)) \\ \vdots & & \vdots & \vdots \\ D_1\sigma_n(s) & & D_{n-1}\sigma_n(s) & A_n(\sigma(s), \mathring{u}(s)) \end{pmatrix}$$

be nonsingular for all $s \in \mathcal{D}$. This means, first of all, that the matrix $(D_j\sigma_i(s))$ must have rank $n-1$, so that σ is an imbedding and $\sigma(\mathcal{D}) \subset \mathbb{R}^n$ is a hypersurface. In addition, the vector $(A_1(\sigma(s), \mathring{u}(s)), \ldots, A_n(\sigma(s), \mathring{u}(s)))$ must not lie in the tangent space of $\sigma(\mathcal{D})$; we express this by saying that the "initial manifold" $\sigma(\mathcal{D})$ is **free for the initial condition** $\mathring{u}$ (for linear equations the initial condition $\mathring{u}$ is irrelevant). Thus we can solve the Cauchy problem for any initial condition along an initial $(n-1)$-manifold which is free for the initial condition.

Now we consider the general first order PDE

$$F(x_1, \ldots, x_n, u(x_1, \ldots, x_n), u_{x_1}(x_1, \ldots, x_n), \ldots, u_{x_n}(x_1, \ldots, x_n)) = 0.$$

We denote the partials of F by

$$F_{x_i}, \quad F_u, \quad F_{p_i}.$$

Consider curves Γ in $\mathbb{R}^{2n+1}$ satisfying

$$\begin{cases} \dfrac{d\Gamma_i(t)}{dt} = F_{p_i}(\Gamma(t)) & i = 1, \ldots, n \\ \dfrac{d\Gamma_{n+1}(t)}{dt} = \displaystyle\sum_{i=1}^{n} \Gamma_{n+1+i}(t) \cdot F_{p_i}(\Gamma(t)) & \\ \dfrac{d\Gamma_{n+1+i}(t)}{dt} = -F_{x_i}(\Gamma(t)) - \Gamma_{n+1+i}(t) F_u(\Gamma(t)) & i = 1, \ldots, n. \end{cases}$$

As before, we easily check that if Γ satisfies these equations, then $F(\Gamma(t))$ is constant in t. A solution Γ with $F(\Gamma(t)) = 0$ for all t is called a **characteristic strip**. Now suppose we have a one-one map

$$\sigma : \mathcal{D} \to \mathbb{R}^n$$

with $\mathcal{D} \subset \mathbb{R}^{n-1}$ as before, and functions

$$\mathring{u}, \mathring{p}_1, \ldots, \mathring{p}_n : \mathcal{D} \to \mathbb{R}$$

with

$$F(\Sigma(s)) = F(\sigma_1(s), \ldots, \sigma_n(s), \mathring{u}(s), \mathring{p}_1(s), \ldots, \mathring{p}_n(s)) = 0 \qquad \text{for all } s \in \mathcal{D}.$$

Then there is a unique characteristic strip Γ through each point $\Sigma(s)$, and the union of the corresponding base curves is a hypersurface M_u. In order for the function u to be a solution to our PDE we will need two conditions, which allow us to extend the proof of Theorem 3 essentially without change. First, the matrix

$$\begin{pmatrix} D_1\sigma_1(s) & D_{n-1}\sigma_1(s) & F_{p_1}(\Sigma(s)) \\ \vdots & \vdots & \vdots \\ D_1\sigma_n(s) & D_{n-1}\sigma_n(s) & F_{p_n}(\Sigma(s)) \end{pmatrix}$$

must be nonsingular. Thus $\sigma(\mathcal{D}) \subset \mathbb{R}^n$ must be an $(n-1)$-manifold, and $(F_{p_1}(\Sigma(s)), \ldots, F_{p_n}(\Sigma(s)))$ must not lie in its tangent space—once again, we express this by saying that the initial manifold $\sigma(\mathcal{D})$ is **free for the initial data** $\mathring{u}, \mathring{p}_1, \ldots, \mathring{p}_n$. Second, we must have

$$\frac{\partial \mathring{u}}{\partial s_j} = \sum_{i=1}^{n} \mathring{p}_i \cdot \frac{\partial \sigma_i}{\partial s_j}.$$

In terms of Σ, this condition reads

$$\frac{\partial \Sigma_{n+1}}{\partial s_j} = \sum_{i=1}^{n} \Sigma_{n+1+i} \cdot \frac{\partial \Sigma_i}{\partial s_j},$$

and is called the **strip manifold condition**. If we think of a point $(x_1, \ldots, x_n, z, p_1, \ldots, p_n)$ in $\mathbb{R}^{2n+1}$ as a hyperplane in $\mathbb{R}^{n+1}{}_{(x_1,\ldots,x_n,z)}$, namely as the hyperplane perpendicular to the vector $(p_1, \ldots, p_n, -1)$, then $\Sigma\colon \mathcal{D} \to \mathbb{R}^{2n+1}$ may be regarded as a family of hyperplanes along the $(n-1)$-dimensional submanifold $\sigma(\mathcal{D})$. It is easy to see that Σ satisfies the strip manifold condition if and only if the tangent space of $\sigma(\mathcal{D})$ at any point $\sigma(s)$ always lies in the hyperplane determined by Σ at s. We may summarize by saying that we can solve the Cauchy problem for any strip manifold $(\sigma_1, \ldots, \sigma_n, \mathring{u}, \mathring{p}_1, \ldots, \mathring{p}_n)$ for which the initial $(n-1)$-dimensional submanifold $\sigma(\mathcal{D})$ is free for the initial data $\mathring{u}, \mathring{p}_1, \ldots, \mathring{p}_n$.

BIBLIOGRAPHY

This list contains all the books and journal articles that have been referenced, as well as others that I have consulted, or that might be considered classics, or that might be of interest for further reading, or a source for problems.

Complete names for the journal references given here are listed on a separate page after this Bibliography, since it is often impossible to look up journals without knowing the unabbreviated name, and the "standard" abbreviations aren't always so standard.

Names of book publishers are generally treated more casually, since most large collected works can be found only in libraries, information about publishers is usually easy to find on the web, and new printings of old classics are continually appearing, with many new and old printings now available on the web.

Abel, N. H.

[1] "Auflösung einer mechanischen Aufgabe", *J. Reine Angew. Math.*, **1** (1826), 153–157. (Also to be found in the first volume of Abel's *Œuvres Complètes*.)

Abraham, Ralph and Marsden, Jerrold E.

[1] *Foundations of Mechanics*, 2nd ed., Benjamin/Cummings, 1978.

Alciatore, David G.

[1] Web site: `http://billiards.colostate.edu`.

For the material discussed in Chapter 11, click on "Technical Proof (TP) analyses", and scroll down to the TP A section, especially the material from TP A.4.

Anh, Le Xuan

[1] *Dynamics of Mechanical Systems with Coulomb Friction*, A. K. Belyaev, trans., Springer, 2003.

Appell, Paul

[1] *Traité de Mécanique Rationnelle*, 3rd ed., Gauthier-Villars, 1909–1937.

Arnold, V. I.

[1] "ОБ ОДНОЙ ТЕОРЕМЕ ЛИУВИЛЛЯ КАСАЮЩЕЙСЯ ИНТЕГРИРУЕМЫХ ПРОБЛЕМ ДИНАМИКИ", *СИБИРСКИЙ МАТЕМАТИЧЕСКИЙ ЖУРНАЛ* **4** No. 2, 1963.

[2] *Mathematical Methods of Classical Mechanics*, K. Vogtmann and A. Weinstein, trans., Springer, 1978; 2nd ed., 1989.

[3] *Huygens and Barrow, Newton and Hooke*, Birkhäuser, 1990.

Arnold, V. I. and Avez, A.

[1] *Ergodic Problems of Classical Mechanics*, Benjamin, 1968.

Atzema, Eisso J.

[1] "Charles François Sturm's Writings on Optics", 49–65, in *Collected Works of Charles Frainçois Sturm*, J-C. Pont *et al.* eds., Birkhäuser, 2009.

Audin, Michèle

[1] *Torus Actions on Symplectic Manifolds*, 2nd edition, Birkhäuser, 2004.

Auerbach, D.

[1] "Colliding rods: dynamics and relevance to colliding balls", *Am. J. Phys.* **62** (1994), 522–525.

Avez, A. *See* Arnold, V. I. and Avez, A.

Barbour, J. and Pfister, H.
[1] *Mach's Principle. From Newton's Bucket to Quantum Gravity*, J. Barbor and H. Pfister, eds., Birkhauser, 1995.

Barger, Vernon and Olsson, Martin
[1] *Classical Mechanics. A Modern Perspective*, McGraw-Hill, 1955; 2nd ed., 1973.

Belorizky, Elie and Sivardière, Jean
[1] "Comments on the horizontal deflection of a falling object", *Am. J. Phys.* **55** (1987), 1103–1104.

Berry, M. V.
[1] "Quantal phase factors accompanying adiabatic changes", *Proc. R. Soc. Lond. A* **392** (1984), 45–57.
[2] "Classical adiabatic angles and quantal adiabatic phase", *J. Phys. A. Math. Gen.* **18** (1985), 15–27.

Bertrand, J.
[1] "Théorème relatif au mouvement d'un point attiré vers un centre fixe", *C. R. Acad. Sci. Paris* **77** (1873), 849–853.
An English translation is available at `arXiv.org/abs/0704.2396`

Beutler, G.
[1] *Methods of Celestial Mechanics*, Vol. 1, Springer, 2005.

Boccaletti, D. and Pucacco, G.
[1] *Theory of Orbits*, Vol. 1, Springer, 1996.

Bohlin, M. K.
[1] "Note sur le problème des deux corps et sur une intégration nouvelle dans le problemé des trois corps", *Bull. Astr.* **28** (1911), 113–119.

Bolsinov, A. V. and Fomenko, A. T.
[1] *Integrable Hamiltonian Systems*, Chapman & Hall/CRC, 2004.

Born, Max
[1] *Vorlesungen über Atommechanik*, Springer, 1925.
Available as *The Mechanics of the Atom*, J. W. Fisher, trans., D. R. Hartree, revisions, Bell and Sons, 1927.

Born, Max and Wolf, Emil
[1] *Principles of Optics*, 7th ed., Cambridge University Press, 1999.

Bou-Rabee, Nawaf M., Marsden, Jerrold E., and Romero, Louis A.
[1] "A geometric treatment of Jellett's egg", *Z. Angew. Math. Mech.* **85** (2005), 618-642.

Brogliato, B.
[1] *Nonsmooth Mechanics*, 2nd ed. Springer, 1999.
See also Génot, F.

Bruce, Ian
[1] Web site: `17centurymaths.com`.

Bruns, Heinrich
[1] "Das Eikonal", *Abh. Math. Phys. Kl. Königl. Sächs. Ges. Wiss.*, **21** (1895), 322–436.

Buchwald, Jed Z.
[1] *The Rise of the Wave Theory of Light*, University of Chicago Press, 1989.

Burgers, J. M.
[1] "Die adiabatischen Invarianten bedingt periodische Systeme", *Ann. Physik* **52** (1917), 195-202.

Cabannes, Henri
[1] *General Mechanics*, 2nd ed., S. P. Sutera, trans., Blasidell, 1968.

Calkin, M. G.
[1] *Lagrangian and Hamiltonian Mechanics*, World Scientific, 1966.

Calkin, M. G. and March, R. H.
[1] "The dynamics of a falling chain: I", *Am. J. Phys.* **57** (1989), 154–157.

Carathéodory, Constantin
[1] *Calculus of Variations and Partial Differential Equations of the First Order. Part I*, Robert B. Dean and Julius J. Brandstatter, trans., Holden-Day, 1965.

Casey, James
[1] "The elasticity of wood", *The Physics Teacher*, **31** (1993), 286–288.

Cayley, Arthur
[1] "On a class of dynamical problems", *Proc. R. Soc. Lond.* **8** (1857), 506–511.

Ceanga, V. and Hurmuzlu, Y.
[1] "A new look at an old problem: Newton's cradle", *J. Appl. Mechanics* **68** (2001), 575–583.

Chandrasekhar, S.
[1] *Ellipsoidal Figures of Equilibrium*, Dover, 1987. A reprint, with revisions, of the 1969 book published by Yale University Press.
[2] *Newton's* Principia *for the Common Reader*, Clarendon, 1995.

Chapman, S.
[1] "Misconception concerning the dynamics of the impact ball apparatus", *Am. J. Phys.* **28** (1960), 705–711.

Chicone, C. and Jacobs, M.
[1] "Bifurcation of critical periods for plane vector fields", *Trans. Am. Math. Soc.* **312** (1989), 433–486.

Cohen, Bernard, and Whitman, Anne

See Newton [1]. We use "Cohen-Whitman" whenever we are referring to additional commentary, as opposed to the translation itself.

Cohen, Richard J.
[1] "The tippe top revisited", *Am. J. Phys.* **45** (1977), 12–17.

Cordani, Bruno
[1] *The Kepler Problem*, Birkhäuser, 2003.

Coriolis, G.-G.

[1] *Théorie Mathématique des Effets de Jeu de Billard*, together with Coriolis' papers of 1832 and 1835, Editions Jacque Gabay, 1990.

[2] *Mathematical Theory of Spin, Friction, and Collision in the Game of Billiards*, David Nadler, trans. Privately printed, available at `www.coriolisbilliards.com`.

Courant, R. and Hilbert, D.

[1] *Methods of Mathematical Physics*, Vol. 2, Wiley & Sons, 1962.

Delaunay, C.

[1] *Théorie du Mouvement de la lune, tome I*, [*Mem. Acad. Sci.* **28**], Mallet-Bachelier, 1860; *tome II*, [*Mem. Acad. Sci.*, **29**]. Gauthier-Villars, 1867.

DG. *See* Spivak, Michael [2].

Dicke, R. H. *See* Roll, P. G.

Dijksterhuis, E.J.

[1] *Archimedes*, Princeton University Press, 1987.

Duck, Ian and Sudarshan, E. C. G.

[1] *100 years of Plank's Quantum*, World Scientific, 2000.

Dugas, Rene

[1] *A History of Mechanics*, J. R. Maddox, trans., Dover, 1988.

Einstein, A.

[1] "Zum Quantensatz von Sommerfeld und Epstein", *Verh. Dtsch. Phys. Ges.* **19**, 82-92.

[2] "Prinzipielles zur allgemeinen Relativitätstheoreie", *Ann. Physik* **55** (1918), 241–244.

Euclid

[1] *The Arabic Version of Euclid's* Optics. (Kitāb Uglīdis fī Ikhtilāf al-manāẓir), Vol. 1, Elaheh Kheirandish trans., Springer, 1999.

Fasano, Antonio and Marmi, Stefano

[1] *Analytical Mechanics*, Beatrice Pelloni, trans., Oxford University Press, 2006.

Fenchel, W.

[1] "On conjugate convex functions", *Can. J. Math.*, **1** (1949), 73–77.

Feynman, R., Leighton, R., and Sands, M.

[1] *The Feynman Lectures on Physics, Volume 1*, Addison Wesley, 1963.

Finch, Janet D. *See* Hand, Louis N.

Fomenko, A. T. *See* Bolsinov, A. V.

French, A. P.

[1] "The deflection of falling objects", *Am. J. Phys.* **52** (1984), 199.

Galilei, Galileo

[1] *Dialogues Concerning the Two Chief World Systems*, Stillman Drake, trans., University of California Press, 1967.

[2] *Dialogues Concerning Two New Sciences*, Henry Crew and Alfonso de Salvio, trans., Dover, 1954.

Garwin, Richard L.

[1] "Kinematics of an Ultraelastic Rough Ball", *Am. J. Phys.* **37** (1969), 88–92.

Gary, C. G. and Nickel, B. G.

[1] "Constants of the motion for nonslipping tippe tops and other tops with round pegs", *Am. J. Phys* **68** (2000), 821–828.

Gauss, C. F.

[1] *Werke*, Vol. 5, 498–503.

Génot, F. and Brogliato, B.

[1] "New results on Painlevé paradoxes", *Eur. J. Mech. A Solids* **18** (1999), 653–677.

Giaquinta, Mariano and Hildebrandt, Stefan

[1] *Calculus of Variations II*, Springer, 1996.

Glad, T. *See* Rauch-Wojciechowski, S.

Goldstein, Herbert

[1] *Classical Mechanics*, Addison Wesley, 1950.

Gras-Mati, Albert *See* Santos-Beniot, Julio, V.

Griffith, B. *See* Synge, J.

Guicciardini, N.

[1] "Johann Bernoulli, John Keill and the inverse problem of central forces", *Ann. of Sci.* **52** (1995), 537–575.

Hale, Jack K.

[1] *Oscillations in Nonlinear Systems*, McGraw-Hill, 1963.

Hall, E. H.

[1] "Do falling bodies move south?", *Physical Review*, **17** (1903), 179–190 and 245–254.

Hamel, Georg

[1] *Theoretische Mechanik*, Springer, 1949.

Hamilton, William Rowan

[1] *The Mathematical Papers of Sir William Rowan Hamilton*, 4 Vols., Cambridge University Press, 1931.

Vol. 1 contains the papers on optics, Vol. 2 the two main papers on dynamics.

Hand, Louis N. and Finch, Janet D.

[1] *Analytical Mechanics*, Cambridge University Press, 1998.

Hannay, J. H.

[1] "Angle variable holonomy in adiabatic excursion of an integrable Hamiltonian", *J. Phys. A. Math. Gen.* **18** (1985), 221–230.

Helmholz, A. Carl *See* Kittel, Charles

Herivel, J.

[1] *The Background to Newton's Principia*, Oxford University Press, 1965.

Herrmann, F. and Schmälzle, P.

[1] "Simple explanation of a well-known collision experiment", *Am. J. Phys.* **49** (1981), 761–764.

Herrmann, F. and Seitz, M.

[1] "How does the ball-chain work?", *Am. J. Phys.* **50** (1982), 977–981.

Herzberger, Max
[1] *Modern Geometrical Optics*, Interscience, 1958.

Hilbert, D. *See* Courant, R.

Hildebrandt, Stefan. *See* Giaquinta, Mariano.

Hjorth, P. G. *See* Knudsen, J. M.

Hugenholtz, N. M.
[1] "On tops rising by friction", *Physica (Amsterdam)*, **18** (1952), 503–514.

Hughes, Thomas J. R. *See* Marsden, Jerrold E.

Hurmuzlu, Y. *See* Ceanga, V.

Huygens, Christaan
[1] *The Pendulum Clock*, R. J. Blackwell, trans., Iowa State University Press, 1986.
[2] *Treatise on Light*, BiblioBazaar, 2007. A reprint of the translation of 1912 by Silvanus P. Thompson.

Jacobi, C. G. J.
[1] *Vorlesungen über Dynamik*, included as Vol. 8 of his *Œuvres Complètes.*

Now available as *Jacobi's Lectures on Dynamics, Second Edition, Edited by A. Clebsch*, K. Balagangadharan and Biswarup Banerjee, trans., Hindustan Book Agency, 2009.

Jacobs, M. *See* Chicone, C.

Jellett, J. H.
[1] *A Treatise on the Theory of Friction*, Hodges, Foster and Co. and MacMillan, 1872.

Available for reading or download at `google.books.com`.

Jost, R.
[1] "Winkel- und Wirkungsvariable für allgemeine mechanische Systeme", *Helvetica Physica Acta*, **41** (1968), 965–968.

Kevorkian, J.
[1] *Partial Differential Equations*, Wadsworth & Brooks/Cole, 1990; 2nd ed. Springer, 2000.

Khein, Alexander and Nelson, D. F.
[1] "Hannay angle study of the Foucault pendulum in action-angle variables", *Am. J. Phys.* **61** (1993), 170–174.
[2] "A persistent error in action-angle treatments of Hamiltonian mechanics", *Am. J. Phys.* **61** (1993), 175–176.

Kheirandish, Elaheh *See* Euclid [1].

Kittel, Charles, Knight, Walter D., Ruderman, Malvin, A., Helmholz, A. Carl, and Moyer, Burton J.
[1] *Mechanics. Berkeley Physics Course. Volume 1*, 2nd ed. McGraw-Hill, 1973.

Klein, Martin J.
[1] *Paul Ehrenfest*, North-Holland, 1970.

Kleppner, D. and Kolenkow, R.
[1] *An Introduction to Mechanics*, McGraw-Hill, 1973.

Knight, Walter D. *See* Kittel, Charles

Knudsen, J. M. and Hjorth, P. G.
[1] *Elements of Newtonian Mechanics*, 3rd ed., Springer, 2000.

Kolenkow, R. *See* Kleppner, D.

Kozlov, V. V.
[1] *Dynamical Systems* X. *General theory of Vortices*, Springer, 2003.

Krim, Jacqueline
[1] "Friction at the atomic scale", *Scientific American*, October, 1996, 74–80.

Krotkov, R. *See* Roll, P. G.

Lagrange, J.-L.
[1] *Mécanique Analytique*, Gauthier-Villars et fils, 1788.

Available in English translation as *Analytical Mechanics*, Auguste Boissonnade and Vidor N. Vagliente, eds. and trans., Kluwer, 1997.

Laidler, Keith J.
[1] "The meaning of 'adiabatic' ", *Can. J. Chem.* **72** (1994),

Landau, L. D. and Lifschitz, E. M.
[1] *Mechanics*, 3rd ed., J. B. Sykes and J. S. Bell, trans., Pergamon, 1994.

Laplace,
[1] *Œuvres Complètes*, Vol. 14, 267–277.

Lee, Hwa-Chung
[1] "The universal integral invariants of Hamiltonian systems and applications to the theory of canonical transformation", *Proc. R. Soc. Edinburgh, Sect. A.* **11** Part 3 (1946-48).

Legendre, A.-M.
[1] "Mémoire sur l'integration de quelques équations aux différences partielles", *Mémoires de l'Académie Royale des Sciences* (1789), 309–351.

Legendre's transformation had been given earlier by Euler (cf. Giaquinta and Hildebrandt [1; pg. 146]), who wrote mathematics faster than other mathematicians could read it, but Euler already has enough things named after him.

Leighton, R. B. *See* Feynman, R. *and* Neher, H. V.

Leutwyler, H.
[1] "Why some tops tip", *Eur. J. Phys.* **15** (1994), 59–61.

Lifschitz, E. M. *See* Landau, L. D.

Liouville, J.
[1] "Note sur l'intégration des équations différentielles de la Dynamique, présentée au Bureau des Longitudes le 29 juin 1853", *J. Math. Pures Appl.*, **20** (1855), 137–138.

Lochak, P. and Meunier, C.
[1] *Multiphase Averaging for Classical Systems. With Applications to Adiabatic Theorems*, Springer, 1988.

Mach, E.
[1] *The Science of Mechanics*, Open Court Publishing Co., 1974.

March, R. H. *See* Calkin, M. G.

Marmi, Stefano *See* Fasano, Antonio

Marsden, Jerrold E. *See* Abraham, Ralph *and* Bou-Rabee, Nawaf M.

Marsden, Jerrold E. and Hughes, Thomas J. R.
[1] *Mathematical Foundations of Elasticity*, Prentice-Hall, 1983.

Marsden, Jerrold E. and Ratiu, Tudor S.
[1] *Introduction to Mechanics and Symmetry*, 2nd ed., Springer, 1999.

Maxwell, James Clerk
[1] *The Electrical Researches of Henry Cavendish, edited by Maxwell*, Frank Cass & Company, Ltd., 1967.

McOwen, Robert
[1] *Partial Differential Equations*, Prentice Hall, 1996.

Mehra, Jagdish and Rechenberg, Helmut
[1] *The Historical Development of Quantum Theory*, Volume 1 (published in 2 parts), Springer, 1982.

Meunier, C. *See* Lochak, P.

Meyer, E.
[1] *Nanoscience: friction and rheology on the nanometer scale*, World Scientific, 1998.

Miller, K. S. and Ross, B.
[1] *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley & Sons, 1993.

Mineur, H.
[1] "Réduction des systèmes mécaniques à *n* degrés de liberté admettant *n* intégrales premières uniformes en involution aux systèmes à variables séparées", *J. Math. Pures Appl.* **15** (1936), 385–389.

Misner, C. W., Thorne, K. S., and Wheeler, J. A.
[1] *Gravitation*, W. H. Freeman and Company, 1973.

Moffatt, H. K. and Shimomura, Y.
[1] "Spinning eggs — a paradox resolved", *Nature* **416** (2002), 385–386.

Morin, David
[1] *Introduction to Classical Mechanics With Problems and Solutions*, Cambridge University Press, 2008.

Moser, Jürgen
[1] "On the volume elements on a manifold", *Trans. Am. Math. Soc.* **120** (1965), 286–294.

See also Siegel, C. L.

Moyer, Burton J. *See* Kittel, Charles

Mukunda, N. *See* Sudarshan, E. C. G.

Nahin, Paul J.
[1] *When Least is Best*, Princeton University Press, 2004.

Navarro, L and Perez, E.

[1] "Paul Ehrenfest on the necessity of quanta (1911): discontinuity, quantization, corpuscularity and adiabatic invariance", *Archs. Hist. Exact Sci.* **58** (2004) 97–141.

[2] "Paul Ehrenfest: The genesis of the adiabatic hypothesis, 1911-1914", *Archs. Hist. Exact Sci.* **60** (2006) 209–267.

Neher, H. V. and Leighton, R. B.

[1] "Linear Air Trough", *Am. J. Phys.* **31** (1963), 225.

Nelson, D. F. *See* Khein, Alexander

Newton, I.

[1] *The Correspondence of Isaac Newton*, Cambridge University Press, 1959–1977.

[2] *The Principia*, Translated by Bernard Cohen and Anne Whitman, University of California Press, 1999.

Nickel, B. G. *See* Gary, C. G.

Noether, Emmy

[1] "Invariante Variationsprobleme", *Nachr. Kgl. Ges. Wiss. Göttg., Math-phys. Klasse* (1918), 235–257.

English translation available on-line at `arXiv.org/abs/physics/0503066v1`.

Olsson, Martin *See* Barger, Vernon

Osgood, W. F.

[1] *Mechanics*, Macmillan, 1937; Crawford Press, 2007.

Page, Leigh

[1] *Introduction to theoretical physics*, 3rd ed., Van Nostrand, 1952.

Painlevé, P.

[1] *Leçons sur le frottement*, Hermann, 1895.

Available for reading or download at `books.google.com` (search for the title "sur le frottement" to avoid problems with the cedilla).

[2] "Sur les lois du frottement de glissement", *C. R. Acad. Sci. Paris* **121**, 112–115; **141**, 401–405; **141**, 546–552.

Palais, Richard S. and Palais, Robert A.

[1] *Differential Equations, Mechanics, and Computation*, American Mathematical Society and IAS/Park City Mathematics Institute, 2009.

Pars, L. A.

[1] *A Treatise on Analytical Dynamics*, Wiley & Sons, 1965.

Perez, E.

[1] "Ehrenfest's adiabatic theory and the old quantum theory", *Archs. Hist. Exact Sci.* **63** 81–125.

See also Navaro, L.

Persson, Anders O.

[1] "The Coriolis Effect: Four centuries of conflict between common sense and mathematics, Part I: A history to 1885", *History of Meteorology* **2** (2005). See `meteohistory.org/2005historyofmeteorology2/01persson.pdf`.

Pfister, H. *See* Barbour, J.

Planck, Max

[1] *Scientific Autobiography*, Philosophical Library, New York, 1949.

Pliskin, W. A.

[1] "The Tippe Top (Topsy-Turvy Top)", *Am. J. Phys.* **22** (1954), 28–32.

Poinsot, M.

[1] *Théorie Nouvelle de La Rotation des Corps*, Bachelier, 1851.

Pourciau, B.

[1] "On Newton's proof that inverse-square orbits must be conics", *Ann. of Sci.*, **48** (1991), 159–172.

[2] "Newton's solution of the one-body problem", *Archs. Hist. Exact Sci.* **44** (1992), 125–146.

[3] "The integrability of ovals: Newton's lemma 28 and its counterexamples", *Archs. Hist. Exact Sci.* **55** (2001), 479–499

[4] "Newton's argument for proposition 1 of the Principia", *Archs. Hist. Exact Sci.* **57** (2003), 267-311.

[5] "Force, deflection, and time: Proposition VI of Newton's Principia", *Hist. Math.* **34** (2007), 140–172.

[6] "Newton's Second Law (as Newton Understood it) from Galileo to Laplace" (to appear).

Pucacco, G. *See* Boccaletti, D.

Ratiu, Tudor S. *See* Marsden, Jerrold E.

Rauch-Wojciechowski, S., Sköldstam, M., and Glad, T.

[1] "Mathematical analysis of the Tippe Top", *Regular and Chaotic Dynamics*, **10** 2005, 333–362.

Rayleigh [Lord Rayleigh] (John William Strutt, Baron Rayleigh)

[1] "On the pressure of vibrations", *Philosophical Magazine* **3** (1902), 338–346. (Also to be found in Vol. 5 of Lord Rayleigh's *Scientific Papers*.)

Rechenberg, Helmut *See* Mehra, Jagdish

Reddingius, E.

[1] "Comment on 'The eastward deflection of a falling object'", *Am. J. Phys.* **52** (1984), 562–563.

Roll, P. G., Krotkov, R., and Dicke, R. H.

[1] "The equivalence of inertial and passive gravitational mass", *Ann. Phys. (U.S.A.)* **26** (1964), 442–517.

Romero, Louis A. *See* Bou-Rabee, Nawaf M.

Ronchi, Vasco

[1] *The Nature of Light*, Harvard University Press, 1970.

Ross, B. *See* Miller, K. S.

Routh, E. J.

[1] *Advanced Dynamics of a System of Rigid Bodies*, Dover 1955 reprint. Available from Dover Books; original 1905 printing available for reading or download at `books.google.com`.

Ruderman, Malvin, A. *See* Kittel, Charles

Sabra, A. I.
[1] *Theories of Light From Descartes to Newton*, Oldbourne, London, 1967.

Sands, M. *See* Feynman, R.

Santos-Beniot, Julio, V. and Gras-Mati, Albert
[1] "Ubiquitous drawing errors in the simple pendulum", `albertgrasmarti.org/agm/recerca-divulgacio/pendulum-TPT.pdf`

Schmälzle, P. *See* Herrmann, F.

Schrödinger, E.
[1] "An undulatory theory of the mechanics of atoms and molecules", *Physical Review* **28** (1926), 1049–1070.

Schulz-DuBois, E. O.
[1] "Foucault pendulum experiment by Kamerlingh Onnes and degenerate perturbation theory", *Am. J. Phys.* **38** (1970), 173–188.

Schwarzschild, K.
[1] "Zur Quantenhypothese", *Sitzungsber. Deut. Akad. Wiss. Berlin Kl. Math. Phys. Tech.* **16** (1916), 548–568.

Seitz, M. *See* Herrmann, F.

Shapiro, A.
[1] "Bath-Tub Vortex", *Nature*, **196** (1962), 1080–1081.

Shimomura, Y. *See* Moffatt, H. K.

Siegel, C. L. and Moser, J. K.
[1] *Lectures on Celestial Mechanics*, Springer-Verlag, 1971.

Siklos, Stephen T. C. *See* Wells, Clive G.

Sivardière, Jean. *See* Belorizky, Elie.

Sköldstam, M. *See* Rauch-Wojciechowski, S.

Smith, D. E.
[1] *A Source Book in Mathematics*, Dover Publications.

Sommerfeld, Arnold
[1] *Atombau und Spektrallinien*, 4th ed., Vieweg und Sohn, 1924

In the Mathematical Addenda, §6 [as of the 4th edition] introduces the use of contour integrals, and §13 says of the use of the method of separation of variables: Er ist für die Quantenprobleme ein wirklicher „Königsweg“. This was rendered in Goldstein [1; §9-8] as "a royal road to quantization", the standard quotation nowadays.

[2] *Mechanics*, Academic Press, 1964.

Spivak, Michael
[1] *Calculus*, 4th ed., Publish or Perish, Inc., 2008.
[2] *A Comprehensive Introduction to Differential Geometry*, 5 volumes, 3rd ed., Publish or Perish, Inc., 2005.

Strutt, John William *See* Rayleigh [Lord Rayleigh]

Sudarshan, E. C. G. *See* Duck, Ian.

Sudarshan, E. C. G. and Mukunda, N.
[1] *Classical Dynamics: A Modern Perspective*, Wiley, 1974.

Synge, J. and Griffith, B.
[1] *Principles of Mechanics*, 3rd ed., McGraw-Hill, 1959.

Terrall, Mary
[1] *The Man who Flattened the Earth: Maupertuis and the Sciences in the Enlightenment*, Chicago, 2002.

Thorne, K. S. *See* Misner, C. W.

Tobin, William
[1] *The Life and Science of Léon Foucault*, Cambridge University Press, 2003.

Trefethen, L. *et al.*
[1] "The Bath-Tub Vortex in the Southern Hemisphere", *Nature*, **207** (1965), 1084–1085.

Treschev, Dmitry and Zubelevich, Oleg
[1] *Introduction to the Perturbation Theory of Hamiltonian Systems*, Springer, 2010.

Truesdell, C.
[1] *Essays in the History of Mechanics*, Springer-Verlag, 1968.

van der Waerden, B. L.
[1] "La démonstration dans les sciences exactes de l'antiquité", *Bull. Soc. Math. Belg.* **4** (1957), 8–20.

Vilasi, Gaetano
[1] *Hamiltonian Dynamics*, World Scientific, 2001.

Weinstein, Alan
[1] "Symplectic structures on Banach manifolds", *Bull. Am. Math. Soc.* **75** (1969), 1040–1041.
[2] "Symplectic geometry", *Bull. Am. Math. Soc.* **5** (July 1981), 1–13.

Wells, Clive G. and Siklos, Stephen T. C.
[1] "The adiabatic invariance of the action variable in classical dynamics", *Eur. J. Phys.* **28** (2007), 105–112.

Wheeler, J. A. *See* Misner, C. W.

Whitman, Anne *See* Cohen, Bernard

Whittaker, E. T.
[1] *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed., Cambridge University Press, 1960.

Wintner, A.
[1] *The Analytical Foundations of Celestial Mechanics*, 1943.

Wolf, Emil. *See* Born, Max.

Wong, C. W., and Yasui, K.
[1] "Falling chains", *Am. J. Phys.* **74** (2006), 490–496.

Yasui, K. *See* Wong, C. W.

Zubelevich, Oleg *See* Treschev, Dmitry

UNABBREVIATED JOURNAL TITLES

Abh. Math. Phys. Kl. Königl. Sächs. Ges. Wiss. = Abhandlungen der Mathematisch-Physischen Classe der Königlich Sächsischen Gessellshaft der Wissenshaften

Am. J. Phys. = American Journal of Physics

Ann. of Sci. = Annals of Science

Ann. Phys. = Annals of Physics

Ann. Physik = Annalen der Physik

Archs. Hist. Exact Sci. = Archives for History of Exact Sciences

Bull. Am. Math. Soc. = Bulletin. American Mathematical Society

Bull. Astr. = Bulletin Astronomique

Bull. Soc. Math. Belg. = Bulletin. Societé Mathématique de Belgique

C. R. Acad. Sci. Paris = Comptes Rendus Académie des Sciences (Paris)

Can. J. Chem. = Canadian Journal of Chemistry

Can. J. Math. = Canadian Journal of Mathematics

Eur. J. Mech. A. Solids = European Journal of Mechanics A. Solids

Eur. J. Phys. = European Journal of Physics

Hist. Math. = Historia Mathematica

J. Appl. Mechanics = Journal of Applied Mechanics

J. Math. Pures Appl. = Journal de mathématiques pures et appliqúees

J. Phys. A. Math. Gen. = Journal of Physics. A. Mathematical and General

J. Reine Angew. Math. = Journal für die Reine und Angewandte Mathematik

Mem. Acad. Sci. = Mémoires. Académie des Sciences. Institut de France.

Nachr. Kgl. Ges. Wiss. Göttg., Math-phys. Klasse = Nachricten der Königliche Gesellschaft der Wissenshaften zu Göttingen, Mathematisch-Physischen Klasse

Proc. R. Soc. Lond. [A.] = Proceedings of the Royal Society of London. [A.]

Proc. R. Soc. Edinburgh, Sect. A. = Proceedings. Royal Society of Edinburgh. Section A. Mathematical and Physical Sciences

Sitzungsber. Deut. Akad. Wiss. Berlin Kl. Math. Phys. Tech. = Sitzungsberichte. Deutsche Akademie der Wissenshaften zu Berlin. Klasse für Mathematik, Physik, und Technik.

Trans. Am. Math. Soc. = Transactions. American Mathematical Society

Verh. Dtsch. Phys. Ges. = Verhandlungen. Deutsche Physikalische Gesellshaft

Z. Angew. Math. Mech. = Zeitschrift für Angewandte Mathematik und Mechanik

INDEX