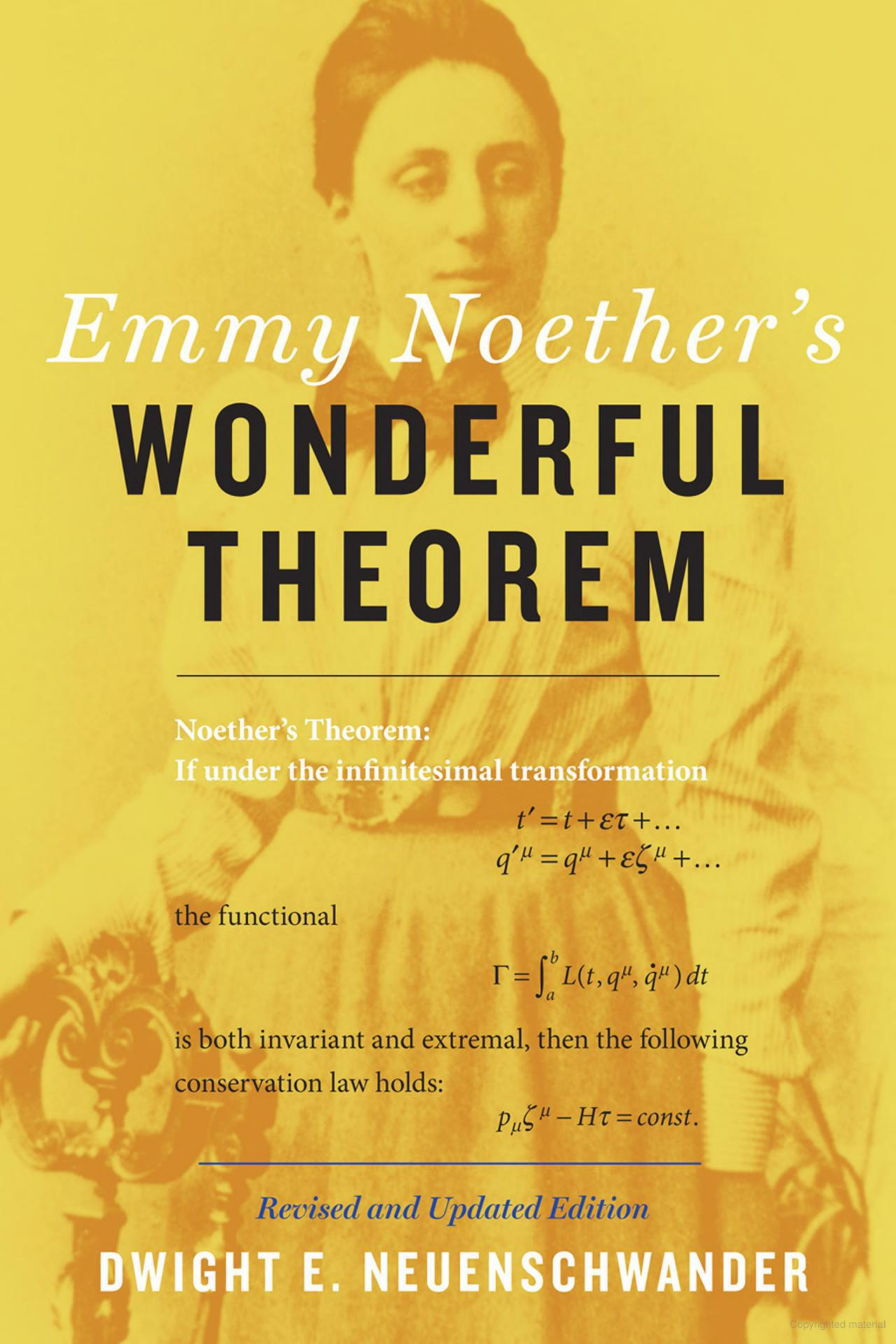

Emmy Noether's WONDERFUL THEOREM

Noether's Theorem:
If under the infinitesimal transformation

$$t' = t + \varepsilon\tau + \dots$$
$$q'^{\mu} = q^{\mu} + \varepsilon\zeta^{\mu} + \dots$$

the functional

$$\Gamma = \int_a^b L(t, q^{\mu}, \dot{q}^{\mu})\, dt$$

is both invariant and extremal, then the following conservation law holds:

$$p_{\mu}\zeta^{\mu} - H\tau = const.$$

Revised and Updated Edition

DWIGHT E. NEUENSCHWANDER

Emmy Noether's Wonderful Theorem

Revised and Updated Edition

Dwight E. Neuenschwander
Southern Nazarene University

Johns Hopkins University Press
Baltimore

Printed in the United States of America on acid-free paper
9 8 7 6 5 4 3 2 1

Johns Hopkins University Press
2715 North Charles Street
Baltimore, Maryland 21218-4363
www.press.jhu.edu

Library of Congress Control Number: 2016953397

A catalog record for this book is available from the British Library.

ISBN-13: 978-1-4214-2267-1
ISBN-10: 1-4214-2267-0

Frontispiece: Emmy Noether (1882–1935). Photo courtesy Bryn Mawr College Library.

Johns Hopkins University Press uses environmentally friendly book materials, including recycled text paper that is composed of at least 30 percent post-consumer waste, whenever possible.

Contents

II WHEN FUNCTIONALS ARE INVARIANT

III THE INVARIANCE OF FIELDS

This page intentionally left blank

Preface

This is a pedagogical book. It aims to introduce Noether's theorem to students who do not understand it already.

When as an undergraduate student I first met the elegant connections between symmetries and conservation laws through Lagrangian mechanics, I changed my major to physics and never looked back. In due time I learned how those connections in mechanics are special cases of a deeply profound theorem published by Emmy Noether in 1918. Noether's theorem as a work of mathematical physics is also a work of poetic beauty.

The creative act, wrote Jacob Bronowski, occurs twice. The first occurs in the mind of the original creator or discoverer. The second occurs in the mind of the appreciator, who re-creates the discovery afresh and sees its significance:

> The poem or the discovery exists in two moments of vision: the moment of appreciation as much as that of creation; for the appreciator must see the movement, wake to the echo which was started in the creation of the work ... In the moment of appreciation we live again the moment when the creator saw and held the hidden likeness. ... When a theory is at once fresh and convincing, we do not merely nod over someone else's work. We re-enact the creative act, and we ourselves make the discovery again. [Bronowski (1956) 19]

As noted above, this book is written especially for physics students to whom Noether's theorem and its related topics are new. The reader I have in mind is a junior or senior undergraduate physics major, or a beginning physics graduate student. Well do I remember being one of those students myself. Those memories include the frustration of trying to read a manuscript loaded with jargon that assumed a fluency I was still struggling to master. At that point in one's career, details that are incidental trifles to experts can become major sticking points for novices (see the following list of questions for examples). If some passages herein seem pedantic or repetitious to experts, I offer to them my apologies. But the offending passages are here because they address questions that were sticking points in my mind when I was a novice, and/or were questions raised by my students. Once fluency is approached, it's easy to forget the struggles. I have tried to remember them.

The Noether theorem—or, as some would say, the two Noether theorems—studied here [Noether (1918); Noether & Tavel (1971)] was for Emmy Noether a special case of her expertise in the abstract algebra of Lie groups and the study of invariants. In 1915 Albert Einstein unveiled his new general theory of relativity, and the mathematicians at Göttingen, notably David Hilbert and Felix Klein, studied it with gusto. They encountered an apparent problem in reconciling energy conservation with the new theory and asked Emmy Noether for her help. Her theorem, which we celebrate here, was the result. It resolved the problem of energy conservation in general relativity through the concept that later became known as local gauge invariance (the "second theorem"), and along the way the "first theorem" gave unified insight into the conservation laws of mechanics and electrodynamics.

After publishing the 1918 paper Emmy Noether went on to become a founder of modern abstract algebra. Graduate students in mathematics are familiar with, for example, Noetherian rings. While abstract algebra in the language of ascending chain conditions and unique factorization domains deserves deep and genuine respect, physics students are more at home with Newton's laws and Maxwell's equations, Lorentz transformations, and de Broglie waves, all expressed in the mathematics of analysis. Fortunately for those of us coming to the conversation from a physics background, Emmy Noether's wonderful theorem can be approached and appreciated in the language of calculus and vector spaces. As a beautiful organizing principle of post-introductory physics, Noether's theorem deserves to be widely known among all physics students, novices and senior physicists alike.

As mentioned, of special interest to me are connections between physics concepts. Noether's theorem stands bright and clear like a magnificent summit in an impressive mountain range of ideas. But the peaks and valleys around a prominent peak are part of the landscape too. Likewise, in these pages we explore topics that run alongside Noether's theorem, interesting for their own sake and which, I think, can be more deeply appreciated with Noether's theorem in the background.

This second edition presents the opportunity to delve into the distinction between Noether's "two" theorems (the second theorem extends the first); to add a few more exercises, references, and technical details; to correct errors in the first edition; and to offer a more vivid picture of Emmy Noether's life and influence. This edition, like the first one, is offered as an expression of appreciation. Thank you for joining me in this adventure of making Emmy Noether one of our intellectual companions.

Acknowledgments

To all my teachers I owe a huge debt of gratitude, and this book offers an installment towards repaying it. Instrumental in my Noether's theorem journey were James Gibbons, my undergraduate theoretical mechanics professor at the University of Southern Colorado (now Colorado State University–Pueblo) who introduced Lagrangian dynamics to me and turned me towards a physics major; John David Logan (University of Nebraska), who at Kansas State University taught a graduate course on Noether's theorem with his own textbook [Logan (1977)] and still graciously answers my questions; Larry Weaver of the KSU physics department, who saw my interests and steered me towards Professor Logan's course; and the late Ari Kyrala, for his expressive insights in mathematics-for-physics courses at Arizona State University.

I thank my students for their good cheer and questions in my courses where these notes were developed. Student collaborators in undergraduate research projects featuring Noether's theorem have so far included Nathan Adams, Kevin Cornelius, Jorge Carmona-Reyes, Lucas Dallen, Will Holmes, Curtis McCully, Mohammed Niazi, Reza Niazi, Keith Slinker, Shawna Starkey-York, Geoffrey Taylor, Brian Turner, and Johnnie Renee West.

I appreciate SNU faculty colleagues Brent Eskridge for cheerfully sharing his mastery of LaTeX, Mark Winslow for his enthusiastic, practical encouragement, and Lee Turner for lending his knowledge of math history. With deep respect I thank The Catalysts, an organization of SNU science alumni, for their steadfast support. Thanks extends off campus to Don Lemons of Bethel College (North Newton, Kansas), Don Salsibury of Austin College (Sherman, Texas), and Raghunath Acharya (Arizona State University).

I thank the readers who corresponded with me about the first edition. Deep appreciation also goes to Camilla MacKay and Rachel Appel of the Bryn Mawr College Library; Gareth Peers of Science Photo Library; the Center for History of Physics of the American Institute of Physics; to former Johns Hopkins University Press editor Trevor Lipscombe and his staff for giving the first edition its chance, and Greg Nicholl for turning my rough sketches into crisp line drawings. To Vincent Burke and the current editor Tiffany Gasbarrini and her staff I am grateful for making this second edition possible.

Above all, I thank my family for their support. Book writing projects enter the lives of an author's family too.

Questions

PRIMARY QUESTIONS

What is a functional?
What is "symmetry"? What is "invariance"? What are "conservation laws"?
How are symmetry, invariance, and conservation laws related?
What are Lagrangians and Hamiltonians? Which is more fundamental?
What are generalized coordinates and their velocities?
What are "canonically conjugate" variables?
What are continuous symmetries? What are discrete symmetries?
Who was Noether of Noether's theorem?
What is the difference between Noether's "first" and "second" theorems?

AUXILIARY QUESTIONS

Where does Hamilton's principle come from?
Are Hamilton's principle and Fermat's principle related?
What does traditional notation such as "$\delta\varphi$" mean?
What is the distinction between "stationary" and "extremal" functionals?
Why is a classical mechanics Lagrangian kinetic *minus* potential energy?
Why do some vector components carry superscripts and others have subscripts?
What are tensors?
What is "gauge invariance"? Why distinguish global from local gauge invariance?
What are internal symmetries?
What is "minimal coupling"? What are "covariant derivatives"?
What is the "Jacobian" of a transformation?
What are Legendre transformations and what are they good for?
What is "phase space"? What is it good for?
Why are complex variables used to describe wave functions?
Why do complex scalar field Lagrangian densities lack the $\frac{1}{2}$ of real scalar fields?

Why does Noether's theorem consider *infinitesimal* transformations?
What are "unitarity" and "Hermitian" operators? Why do we need them?
What are "equations of continuity"? How do they describe conservation locally?
What are "proper" and "improper" conservation laws?
What is a group? What does SU(N) mean and what is it for?
What is the role of Hamilton-Jacobi theory?

Part I

WHEN FUNCTIONALS ARE EXTREMAL

This page intentionally left blank

Chapter 1

Symmetry

To define the idea of symmetry is certainly not simple. I shall not try to make an all-encompassing or precise definition, against which a contradiction can quickly be brought. . . . The idea is still alive and growing. We don't know all that the concept implies. I like best the idea of the seventeenth century philosopher Leibniz. . . . For Leibniz, symmetry is related to the indiscernibility of differences. Once you walk into the hall of a Palladian building, you can't quite remember whether you turned left or right.
—Philip Morrison, "On Broken Symmetries," in Judith Wechsler, ed., *On Aesthetics in Science, 1981*

1.1 Symmetry, Invariances, and Conservation Laws

The conservation of energy, linear momentum, angular momentum, and electric charge are among the most fundamental principles of physics. Have you ever wondered why nature cherishes these quantities so much that she conserves them? Asking "why" in this context may be a bottomless question. But we can connect these conservation laws to deeper principles through the elegant theorem published by Emmy Noether in 1918.[1] Noether's theorem relates a huge class of conservation laws to symmetries of space and time and "internal" variables.

Let's see how one of these familiar conservation laws could be connected to a symmetry, using for now only pre-Noether concepts encountered in an introductory physics course. Consider, for example, the conservation of

[1] The original: Noether (1918); Mort Tavel's translation: Noether & Tavel (1971); a translation also appears in Kosmann-Schwarzbach (2010).

linear momentum. Let a particle of mass m move in one dimension along an x-axis. Newton's second law says

$$F = \frac{dp}{dt} \tag{1.1}$$

where F denotes the net force, t the time, and $p = mdx/dt$ the momentum. For the momentum to be conserved—to not change with time—the net force on the particle must vanish. What does this have to do with symmetry?

Let the particle's interactions with the world be expressed through some potential energy function $U(x)$, whose negative gradient equals the force. Write Newton's second law as the differential equation

$$-\frac{\partial U}{\partial x} = \frac{dp}{dt} \tag{1.2}$$

and approximate the left-hand side as $-\Delta U/\Delta x$. When the particle goes from x to $x + \varepsilon$ then eq. (1.2) may be written

$$-\left(\frac{U(x+\varepsilon) - U(x)}{\varepsilon}\right) = \frac{dp}{dt}. \tag{1.3}$$

To have $dp/dt = 0$ as $\varepsilon \to 0$, we must have $U(x+\varepsilon) - U(x) \sim \varepsilon^s$, where $s > 1$. For then, as the particle moves from one location to another, the change in U goes to zero faster than ε as $\varepsilon \to 0$, and the particle's momentum shows no measurable change. A tangible example you may recall from your introductory physics course may be a glider moving smoothly over a horizontal air track: "here" is identical to "over there." Evidently, the conservation of linear momentum follows from the system being unchanged by a spatial translation.

To say it another way, it means that if you *do* want the particle's momentum to change, a mere translation through space won't do it; space, itself, is homogeneous. So translational "sameness" signifies a *symmetry* of space. If you want the momentum to change then you have to make the space *over there* different from space *here*, for example by having the glider move against a spring.

Notice that our momentum example forms an "if-then" statement. It does not claim that the particle's environment is translationally symmetric; indeed, whenever a particle's momentum has been changed, the symmetry of space is spoiled, or "broken," at least for one spatial dimension. For instance, the Earth's gravitational field breaks the symmetry between the vertical and horizontal dimensions.

How can "symmetry" be defined in a way that makes it a *quantitative* concept? What *is* symmetry? For the kind of answer we need here, let

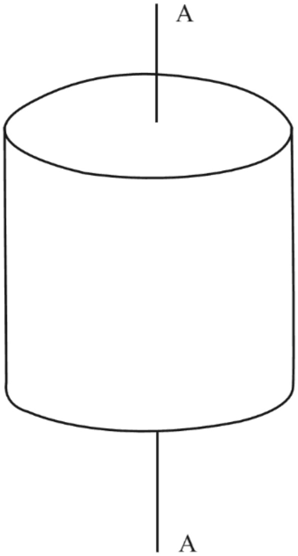

Figure 1.1: *A cylinder with axis of symmetry AA. When rotated about this axis, it looks the same before and after the rotation. That's symmetry.*

us look at a cylinder. I invite your attention to the cylinder's axis AA in Figure 1.1. After you rotate the cylinder about this axis, it looks the same as it did before. Unless you paint a dot on the cylinder, or scratch a mark on its surface to break the symmetry, you cannot tell that the cylinder has been rotated. Rotating it *and not being able to tell it's been rotated* captures the essence of what we mean by "symmetry." To make symmetry quantitative, we need to carry out some operation, or transformation, and see if we can detect a difference. If the difference is too small to detect within some infinitesimal tolerance, then we say the system is "invariant under the transformation."

A clockwise rotation of the cylinder by 90° is equivalent, from an observer's perspective, to moving one's location 90° counterclockwise. Thus a transformation can also be a change of reference frame, from an old coordinate system to a new one. If a quantity survives such transformations unchanged, that quantity is said to be invariant. The appearance of the cylinder is invariant under rotation about the AA axis. The existence of invariance reveals an underlying symmetry: the cylinder appears unchanged under the rotation because it is symmetric about the axis of rotation.

"Conservation" as in "conservation of energy" is not the same as "invariant." They are related—and the exploration of that relation forms the substance of our subject—but they are not synonymous. The momentum or energy of a system of particles may be conserved but not necessarily invariant. For example, imagine one billiard ball approaching another. In the reference frame of the billiard table, prior to the collision the cue ball moves and the eight ball sits at rest, and the momentum of the system is nonzero. But in the center-of-mass reference frame, the system's total momentum sums to zero because the cue ball and eight ball approach one another with opposite momentum. In both frames, the collision is analyzed using conservation of momentum within each frame. The table frame sees nonzero momentum, but in the center-of-mass frame sees zero momentum. In this example, momentum is conserved *within* in each frame, but is not invariant *between* them.

"Invariant" means that a quantity's numerical value is not altered by a coordinate transformation. "Conserved," in contrast, means that within a given coordinate system the quantity does not change throughout a process. "Invariance" compares a quantity between reference frames. "Conservation" compares the quantity before and after a process within a reference frame. Noether's theorem relates conservation to invariance, and thus to symmetry.

Coordinate systems or reference frames are not part of nature. They are maps that we introduce into the solution of a problem for our convenience. Therefore, the content of an equation that is supposed to express a truth about nature must transcend the choice of this or that reference frame. One says the equation must be written "covariantly," or the expression is "covariant." In the preceding example of billiard balls, in a given frame the momentum of the two balls adds up to the same total before and after the collision. In the billiard table frame using unprimed vectors, one writes the conservation of momentum as the vector sum before collision equals the vector sum after collision:

$$[\mathbf{p}_{cue} + \mathbf{p}_8]_{before} = [\mathbf{p}_{cue} + \mathbf{p}_8]_{after}. \tag{1.4}$$

In the center-of-mass frame using primes on the vectors, the same physics content is expressed as

$$[\mathbf{p}'_{cue} + \mathbf{p}'_8]_{before} = [\mathbf{p}'_{cue} + \mathbf{p}'_8]_{after}. \tag{1.5}$$

Never mind that before the collision the table frame measures $\mathbf{p}_8 = \mathbf{0}$ and the center-of-mass frame measures $\mathbf{p}'_8 \neq \mathbf{0}$; the equation expressing conservation of momentum

$$\left[\sum_n \mathbf{p}_n\right]_{before} = \left[\sum_n \mathbf{p}_n\right]_{after} \tag{1.6}$$

contains the same content whatever the reference frame; the relationship it describes holds in either frame. The equation is written covariantly. The quantities in the equation transform under a change of coordinates by the same rules as the coordinates themselves.

We will see that conservation of energy, conservation of linear momentum, and conservation of angular momentum are related to invariance under time translations, space translations, and rotations, respectively. These invariances signify underlying spacetime symmetries: the homogeniety of time, the homogeniety of space, and the isotropy of space. The conservation of electric charge emerges from a more abstract symmetry called "gauge invariance." We can go farther: the conservation of more esoteric "charges," such as quark color charge or weak isospin, involve invariances that hold under transformations among so-called internal degrees of freedom. In the first part of this book we consider spacetime symmetries only. Internal degrees of freedom are considered later.

You will have noticed that I have not said *what* are the quantities whose invariance leads to conservation laws. These quantities are called "functionals." In the functional we have a powerful concept that puts almost all of physics into a common language. Everything wonderful that I am going to relate comes through these functionals.

I have organized this book into four parts. The remainder of Part I (chapters 2 and 3) introduces functionals. As a mathematical machine, by definition a functional Γ takes a *function* as input and produces a *real number* as output. That sounds like the task of a definite integral! While a functional as an abstract concept need not be a definite integral, all the functionals considered in this book are expressed as definite integrals. If you pick an input function $x(t)$, stuff it into the input slot of a functional $\Gamma(\ )$ and turn the crank, you get a real number $\Gamma(x)$. But with a different input function $w(t)$ you get a different output number $\Gamma(w)$. Suppose you want to find the function that produces a maximum or minimum value for Γ. One says the functional is to be made an *extremal*, or, as some say, made *stationary.*[2] As chapter 3 describes, the function that makes Γ an extremum is the solution to a differential equation called the Euler-Lagrange equation.

Part II (chapters 4 and 5) studies the conditions for invariance of the functional under transformations of the independent and/or the dependent variables. When we change t to some new t' and x to some new x', so that $x(t) \to x'(t')$, then $\Gamma(x'(t'))$ may or may not be the same number as $\Gamma(x(t))$. We find that Γ meets our formal definition of invariance if and only if a fundamental invariance identity is satisfied.[3] The plot lines of "extremal"

[2]The distinction between "extremal" and "stationary" is discussed in section 5.3.

[3]The version of the invariance identity we present here is due to Rund [Rund (1972)] and Trautman [Trautman (1967)], who streamlined the Noether theorem proof and notation.

and “invariance” converge in Noether’s theorem (the “first” theorem), as shown in chapter 5. In chapter 5 we also turn the problem around and see how to find transformations that leave a given functional invariant.

Chapter 6 begins Part III by enlarging our program to *fields*, functions of space and time. There we see how charge conservation comes from *global* gauge invariance. Chapter 7 imposes *local* gauge invariance on fields, and we see how far we can push local gauge invariance using Noether’s first theorem. As an introduction to an important application of these ideas, we also examine in chapter 7 how local gauge invariance, when applied to “internal” degrees of freedom or generalized “charges,” becomes a dynamical principle that leads to an understanding of fundamental forces as the exchange of “gauge bosons.”

Chapter 8 extends the program of locally variable (or gauge) transformations to the functional’s dependent and independent variables, leading beyond Noether’s first theorem to the second theorem. The second theorem contains the first as a special case, provides constraints on differential operators, and expresses conservation laws for such systems in terms of so-called covariant derivatives. In clarifying conservation law issues for the coupled matter-field systems of relativistic gravitation, Emmy Noether helped David Hilbert, Felix Klein, and Albert Einstein put the finishing touches on the general theory of relativity in 1915.

Part IV (chapters 9 and 10) considers other applications of invariance that are not part of Noether’s theorem proper, but that share its vocabulary of functionals, transformations, invariances, and conservation laws. Chapter 9 reexamines invariance and conservation in the language of phase space. Whereas Noether’s theorem produces conservation laws given equations of motion and an invariance, Hamilton-Jacobi theory uses equations of motion and conservation laws and produces transformations. Through Hamilton-Jacobi theory, the possibilities offered by an indefinite integral version of the functional are developed. From deep within classical mechanics they suggestively point the way to quantum mechanics, as developed in chapter 10.

Because of the central role of conservation laws, one could argue that Noether’s Theorem offers a strategic unifying principle for most if not all of physics. Although I never had the honor of meeting her personally, I would be remiss if I did not introduce you to Emmy Noether.

1.2 Meet Emmy Noether

“She was not clay, pressed by the artistic hands of God into a harmonious form, but rather a chunk of human primary rock into which he had blown his creative breath of life.”—Hermann Weyl, memorial address for

Emmy Noether, April 26, 1935 in Auguste Dick, *Emmy Noether, 1882–1935*, translated by H. I. Biocher 1981, 112–152

In the judgment of the most competent living mathematicians, Fräulein Noether was the most significant creative mathematical genius thus far produced since the higher education of women began. In the realm of algebra, in which the most gifted mathematicians have been busy for centuries, she discovered methods which have proved of enormous importance in the development of the present-day younger generation of mathematicians.
—Albert Einstein, letter to the *New York Times*, May 5, 1935 in Alice Calaprice, *The Quotable Einstein*, 1996

Amalie Emmy Noether was born on March 23, 1882, the first child of Jewish parents Max (1844–1921) and Amalie Ida Kauffman Noether (1852–1915).[4] Emmy's mother came from a highly educated family that included university scholars in law and history. Emmy's father was descended from iron wholesalers. The first in his line to earn a PhD, Max became a distinguished mathematician, earning his doctorate from the University of Heidelberg, and qualified there in 1870 as a *Privatdozent*.[5] In 1875 Max moved to the University of Elrangen, where Felix Klein (1849–1925)[6] had in 1872 launched the *Erlanger Programm*, which treated geometry as the study of properties of a space that are invariant under a group of transformations. For example, in Euclidean geometry, lengths and angles are unchanged under rotations. This outlook offered a unified approach to classifying the non-Euclidean geometries that had proliferated by the late nineteenth century, pointed to the possibility of new geometries defined from diverse transformation groups [Burton (2011) 602–603], and presumably created the environment that nurtured Emmy's mathematical interests, in which invariance theory played a prominent role. Klein's *Erlanger Programm* brought world prominence to the mathematics department at Erlangen. Max Noether served there the rest of his life. In 1886 Klein accepted a chair at the University of Göttingen, and he brought David Hilbert (1862–1943) to Göttingen in 1895. That move would have effects on Emmy Noether 20 years later, and on readers of this book over a century later.

From 1889 to 1896 Emmy attended the Städtischen Höheren Töchterschule in Erlangen, where she studied languages and piano. She was fond of dancing. As a student she originally planned to teach French and

---

[4]During Max Noether's generation the family name's spelling was changed from Nöther to Noether. Emmy had three younger brothers, Alfred (1883–1918), Fritz (1884–1941), and Gustav Robert (1889–1928) [Brewer & Smith (1982)].

[5]An unsalaried university lecturer with the right to teach independently and advise research students. Lecturers received fees from students rather than a university salary.

[6]Klein is perhaps best known among students for the "Klein bottle."

English, and in 1890 passed the Bavarian State Examination, qualifying her to teach these languages—in schools for women. But her interests had turned increasingly to the eloquent language of mathematics.

This interest was evidently encouraged by her father and his mathematician colleagues, notably Paul Gordan (1837–1912).[7] As Hermann Weyl described in his 1935 memorial address, Emmy developed a close "mathematical kinship" with her father:

> Clebsch has introduced Riemann's ideas into the geometric theory of algebraic curves and [Max] Noether became, after Clebsch had passed away young, his executor in this matter: he succeeded in erecting the whole structure of algebraic geometry of curves on the basis of the so-called Noether residual theorem. . . .
>
> [Max] Noether's residual theorem was later [in the 1920s] fitted by Emmy into her general theory of ideals in arbitrary rings. This scientific kinship of father and daughter—who became in a certain sense his successor in algebra, but stands beside him independent in her fundamental attitude and in her problems—is something extremely beautiful and gratifying. [Dick (1981); Brewer & Smith (1982); James (2002)]

When Gordan passed away in 1912, Max and Emmy Noether wrote his obituary for *Mathematische Annalen*.

After completing grammar school, Emmy wanted to attend university to study mathematics. In the culture of that time and place, women were not allowed to matriculate into German universities,[8] although they could enroll in courses with the professor's permission. As late as 1898, the Academic Senate at the University of Erlangen declared that to admit women as students "would overthrow all academic order." However, in 1900 the spunky Emmy got permission to attend lectures at Erlangen, one of two young women among 986 students during her first semester. There Emmy attended lectures until 1902, and in July 1903 she passed the *matura* examinations, necessary (but not sufficient) for matriculation into university.

That fall she registered at the University of Göttingen,[9] where she attended lectures by distinguished mathematicians who to this day have name recognition among physics students. These included Karl

[7]This is the same Gordan of the beloved Clebsch-Gordan coefficients, which grew out of the study of Lie groups and find application to the quantum addition of angular momentum. Rudolf F. A. Clebsch (1833–1872) was a pioneer in invariant theory, and at Giessen collaborated with Paul Gordan.

[8]Women were allowed to enroll in universities in 1861 in France, 1879 in England, and 1885 in Italy. [Brewer & Smith (1982)]

[9]Evidently Emmy was only allowed to audit university classes during this time.

Figure 1.2: *Emmy Noether (1882–1935). This photo was taken sometime before she entered Göttingen University.* (Science Photo Library)

Schwarzschild (1873–1916), Hermann Minkowski (1864–1909), Felix Klein, and David Hilbert.[10] After she had been one semester at Göttingen, Erlangen University relented in their policy against women and allowed female students to matriculate there and take examinations with the same rights as male students. Emmy returned home to Erlangen and entered the university as a degree-earning student in October 1904. She listed mathematics as her program of study. Academic order was *not* overthrown.

At Erlangen, Paul Gordan became Emmy's PhD advisor. Emmy was Gordan's only doctoral student throughout his distinguished career. She completed the PhD in 1907. Dr. Noether's dissertation was titled *On Complete Systems of Invariants for Ternary Biquadratic Forms.*

In his 1935 memorial tribute to Emmy Noether, Hermann Weyl divided her career into three epochs. Epoch 1, 1908–1919, was a time of collaboration with eminent senior mathematicians. The Noether's theorem celebrated in this book was a product of this epoch. In the second epoch, 1920–1926, Emmy Noether became an eminent mathematician in her own right, one whom later generations would call the "mother of abstract algebra" [Tent (2008)], developing the general theory of ideals and Noetherian rings [Moore (1967) 189–196]. By the time of the third epoch, 1927–1935,

[10]The Schwarzschild of the Schwarzschild metric in general relativity, the Minkowski of Minkowskian spacetime; most physics majors first hear of Hilbert in the context of "Hilbert space" in quantum mechanics.

Noether was recognized as a leader of the "Noetherian school" of algebraists and continued developing noncommunative algebras such as hypercomplex numbers. During that epoch she also found herself a political refugee.

The Noether theorem discussed in these pages was published in 1918 and thus belongs to the first epoch of Emmy Noether's career, the time of collaboration with top mathematicians such as Paul Gordan, Felix Klein, and David Hilbert. The early part of that era saw her hard at work, as Gordan's doctoral student, on the theory of invariants of binary and higher-order forms. A binary form, for example, is a second-order polynomial $p(x, y)$ in two variables x and y of the form

$$p(x, y) = ax^2 + bxy + cy^2. \tag{1.7}$$

If a, b, and c are constants from a specified set of numbers, $p(x, y)$ is an "algebraic form." If these constant coefficients are replaced by functions that include derivatives of x and y, $p(x, y)$ is a "differential form." A large question in the study of quadratic and higher-order forms asks what quantities remain invariant under a change of variables.[11] In her dissertation, Noether calculated and tabulated 331 invariants among "ternary biquadratic forms"[Brewer & Smith (1982)].

Between 1908 and 1915 Dr. Noether worked without salary at the Erlangen Mathematical Institute. She carried out her own research, attended mathematics conferences, was elected to an Italian mathematical society, and as her father's health declined delivered lectures for him as needed. When Ernst Fischer (1875–1959) joined the Erlangen mathematics faculty after Gordan's retirement, Fischer and Noether collaborated on research. Gordan's methods had been algorithmic and computational—finding invariants by cranking through the possibilities. In contrast, Fischer brought Hilbert's abstract algebra approach to mathematics, along the lines of Hilbert's 1888 paper on basis theory [Byers (1999) 5], methods quickly taken up by Emmy Noether.[12] In 1919 Noether's curriculum vitae credits Fischer—whose specialty was also invariance theory—as influencing her to more abstract ways of mathematical thinking, as she mastered approaches closer to Hilbert's.

Noether became so adept at Hilbert's methods of invariance theory that in April 1915, shortly after her mother died, she moved to Göttingen at the invitation of Hilbert and Klien. Her biographers remark that "already, Emmy Noether was recognized for the extreme generality and abstractness

[11] One can see a natural application to differential equations and conservation laws, by setting the $p(x, y)$ of eq. (1.7) equal to zero.

[12] Upon reading one of Hilbert's proofs of a finite basis for specific invariants, Gordan allegedly remarked "Das ist nicht Mathematik; das ist Theologie." It's not certain that Gordan actually said this, or if he did, whether it was in jest, a form of praise, or a criticism [Brewer & Smith (1982)].

of approach which would eventually be seen as her most distinguishing characteristic"[Brewer & Smith (1982) 13].

At Göttingen, women at that time were still not allowed to hold lecturing positions. When Hilbert argued for Noether as a lecturer, prominent nonmathematics members of the faculty senate objected: "What will our soldiers think when they return to the university and find that they are expected to learn at the feet of a woman?" Hilbert famously responded "*Meine Herren*, I do not see that the sex of the candidate is an argument against her admission as a *Privatdozent.* After all, the senate is not a bathhouse!" David Hilbert did not win the argument that day. But determined to keep Emmy Noether at Göttingen, he improvised. Lectures would be announced under his name, but delivered by Fräulein Noether [Reid (1972) 143].

At the end of June and the beginning of July 1915, Albert Einstein spent about a week in Göttingen delivering six lectures on his not-quite-finished general theory of relativity. Noether wrote to Fischer back at Erlangen, "invariant theory is trump here; ... Hilbert is planning to lecture next week on his Einsteinian differential invariants, and to understand that, the Göttingen people must certainly know something!"[Brewer & Smith (1982) 12]. By "invariant theory," Noether meant differential invariants, in contrast to algebraic invariants. In 1935 Weyl recalled [Kosmann-Schwarzbach (2010) 77],

> Hilbert at that time was over head and ears in the general theory of relativity, and for Klein, too, the theory of relativity and its connections with his old ideas of the Erlangen program[13] brought the last flareup of his mathematical interests and mathematical production.

Even after Einstein unveiled the finished general theory of relativity in November 1915, a problem persisted. Hilbert and Klein encountered a puzzle with energy conservation in Einstein's theory.[14] With her invariant-theoretic knowledge of differential forms, Noether was able to help them. In an exchange of letters, Hilbert wrote to Klein, "Emmy Noether, whose help I sought in clarifying questions concerning my energy law" and Klein wrote to Hilbert, "you know that Fraülein Noether continues to advise me in my work" [Pais (1982) 276]. Noether resolved the problem about energy conservation in general relativity, and along the way proved the theorems we study in these pages. The result was the Noether's theorem we celebrate

---

[13]Since general relativity was by this time treating gravitation as the curvature of spacetime, one can see how general relativity as geometry in a four-dimensional pseudo-Riemannian metric space would be attractive to the authors of the *Erlanger programm.* "Pseudo-Riemannian" means the metric is not positive-definite.

[14]See chapter 8.

here, published in 1918 as *Invariant Variation Problems* [Noether (1918); Noether & Tavel (1971); Kosmann-Schwarzbach (2010)], writing:

> Hilbert enunciates his assertion to the effect that the failure of proper laws of conservation[15] of energy is a characteristic feature of the "general theory of relativity." In order for this assertion to hold good literally, therefore, the term "general relativity" should be taken in a broader sense than usual, and extended also to the foregoing groups depending on N arbitrary functions.

In other words, general relativity is a theory of coupled systems exhibiting local gauge invariance.

Hermann Weyl summarized,

> For two of the most significant sides of the general theory of relativity . . . she gave at that time the genuine and universal mathematical formulation. [Brewer & Smith (1982) 13; Dick (1981)]

Albert Einstein was one of the first to appreciate Noether's theorem, describing it in a letter to Hilbert as "penetrating mathematical thinking"[Pais (1982)].

Meanwhile, due to the continuing prejudice against women being professors, Hilbert's repeated attempts to gain Noether a permanent appointment were frustrated. In a letter to Hilbert, dated May 24, 1918, Einstein wrote, "It would not have done the Old Guard at Göttingen any harm, had they picked up a thing or two from her. She certainly knows what she is doing"[Calaprice (1996)]. In a letter to Felix Klein dated December 27 of that same year, Einstein objected again to Dr. Noether not being allowed to lecture officially because she happened to be female: "On receiving the new work from Fräulein Noether, I again find it a great injustice that she cannot lecture officially. I would be very much in favor of taking energetic steps in the ministry [to overturn this rule]."

On May 21, 1919, Dr. Noether submitted an application for *Habilatation*, the right to teach at the university as a *Privadozent*. As part of the application she presented a colloquium, and submitted her curriculum vitae and publications. Besides citing her list of papers on abstract algebra and invariants, in her closing paragraph she adds:

> Finally, there are two works on differential invariants and variation problems. These resulted from my assistance to Klein and Hilbert in their work on the Einsteinian general theory

[15] By "proper laws" (Hilbert's term) he meant conservation laws that could be cleanly expressed by an equation of continuity; see section 6.4.

> of relativity.... The second of these, *Invariant Variation Problems,* which is my Habilitation paper, treats arbitrary continuous Lie groups, finite or infinite, and draws conclusions from a special case of invariance relative to such a group. These general results contain as special cases some known results concerning proper intergrals from mechanics, stability theorems, and certain dependencies among field equations arising in the theory of relativity, while, on the other hand, the converses of these theorems are also given. [Brewer & Smith (1982) 15–16]

As a result of some slight liberalization of policies following World War I, the university granted Dr. Noether the right to lecture under her own name in 1919. In 1922 she was finally named an "unofficial associate professor," a purely honorary position. A subsequent teaching appointment in abstract algebra provided her with a modest salary. She remained at Göttingen for the next decade, except for visiting professorships, where she was warmly received, at Moscow in 1928–1929 and Frankfurt in the summer of 1930. By 1931 she was an associate professor at Göttingen [Brewer & Smith (1982) 75].

The second and third epochs of Emmy Noether's career lie outside the scope of this book. But her work during those years wrote new chapters in abstract algebra that made her name famous among mathematicians everywhere. Those contributions include the origination in 1920 of the *left ideals* and *right ideals*, followed by *Noetherian rings* with their "ascending chain condition" in 1921. Weyl wrote that "she changed the face of algebra by this work." Her influence extended well beyond her students. When mathematician B. L. Van der Waerden spent a year at Göttingen studying with her, he returned to Amsterdam inspired, and wrote a two-volume treatise, *Moderne Algebra*, of which Garrett Birkhoff wrote decades later with the benefit of historical perspective,

> both the axiomatic approach and much of the content of "modern" algebra dates back to before 1914. However, even in 1929, its concepts and methods were still considered to have marginal interest compared with those of analysis.... By exhibiting their mathematical and philosophical unity, and by showing their power as developed by Emmy Noether and her younger colleagues,... Van der Warden made "modern algebra" suddenly seem central in mathematics. [Brewer & Smith (1982) 19]

During the 1927–1935 epoch Dr. Noether's major publications included ideal theory of hypercomplex number systems and their applications to group representations [Brewer & Smith (1982) 146] and noncommuncative algebras.

Noether was a coeditor of the collected mathematical papers of Richard Dedekind (1831–1916), edited the correspondence of Dedekind and Georg Cantor (1845–1918), and did some editing for *Mathematische Annalen* [Dick (1981) 42]. At Göttingen she had an eager group of loyal students, and many of their mathematics discussions took place in the welcoming quarters of Emmy's modest apartment. She was "so amazingly lively!" recalled one of her students. Another remembered, "she lived in close communion with her pupils; she loved them, and took an interest in their personal affairs." The mathematician and philosopher Norbert Wiener (1894–1964)[16] recalled encountering Emmy Noether and a group of her students on a train, "probably the best woman mathematician there has ever been . . . looking like an energetic and very near-sighted washerswoman . . . and her many students flocked around her like a clutch of ducklings about a kind, motherly hen"[Brewer & Smith (1982) 40]. She was blessed with seemingly boundless energy. In her lectures she talked loudly, with large gestures, and frequently thought her way through new ideas while in front of the class—which could make her difficult to follow for listeners not used to it—but her students adored her. In middle age, with her round spectacles and ample lap, she looked like she could have been your favorite auntie—which she was to the children of her brothers. She was not sentimental, but expressed her affection in an almost rough, jolly way [Dick (1981) 46]. She was always generous in sharing ideas with her students, and in talks and publications gave them ample credit for their contributions.

The dark clouds of arrogant intolerance that gathered over Fascist Europe in the 1930s did not spare Emmy Noether. With the appointment of Adolf Hitler as chancellor of Germany in January 1933, swiftly followed by brutal consolidation of Nazi control of German institutions, Dr. Noether, like other Jewish professors in German universities, was abruptly dismissed. Herman Weyl recalled,

> A stormy time of struggle like this one we spent in Göttingen in the summer of 1933 draws people closely together; thus I have a particularly vivid recollection of these months. Emmy Noether—her courage, her frankness, her unconcern about her own fate, her conciliatory spirit—was in the midst of all the hatred and meanness, despair and sorrow surrounding us, a moral solace. . . . Her heart knew no malice.

With the assistance of the Rockefeller Foundation's Emergency Committee to Aid Displaced German Scholars, in 1933 Bryn Mawr College in Pennsylvania offered Dr. Noether a faculty position. There she was treated

[16] Among his many accomplishments, Wiener was the founder of cybernetics, the study of feedback mechanisms.

Figure 1.3: *Emmy Noether, "probably the best woman mathematician there has ever been."* (Bryn Mawr College)

with the respect she had always deserved, but which had too often been denied her before. The president of Bryn Mawr College, Marion Park, wrote to a friend in November 1933, "I am venturing to ask you whether by any lucky chance you can come down to Bryn Mawr in December and see Dr. Noether in action!" In 1934 President Park reported to the Rockefeller Foundation that the newly created Emmy Noether Fellowship would be awarded to distinguished students. In a letter of January 1935, Norbert Wiener wrote,

> Miss Noether is a great personality; the greatest woman mathematician who has ever lived; and the greatest woman scientist of any sort now living, and a scholar at least on the plane of Madame Curie. Leaving all questions of sex aside, she is one of the ten or twelve leading mathematicians of the present generation in the entire world and has founded what is certain to be the most important close-knit group of mathematicians in Germany—the Modern School of Algebraists. . . . In all the cases of German refugees, whether in this country or elsewhere, that of Miss Noether is without doubt the first to be considered. [Brewer & Smith (1982) 33–34]

In March 1935, the newly formed Institute for Advanced Study in nearby Princeton, New Jersey, offered Dr. Noether a stipend to visit the institute and conduct a weekly seminar. One wonders what might have developed in unified field theory had she and Albert Einstein, who was newly installed there, been able to collaborate . . .

Alas! following complications from surgery to remove a tumor on April 10, 1935, in which her initial recovery seemed to be going fine, Emmy Noether went into a coma. She passed from this life on April 14, at the age of 57. Over the next few weeks, tributes to her appeared in several languages. At the Bryn Mawr memorial service where Hermann Weyl spoke with such eloquence on April 26, 1935, he summarized her personality as "warm as a loaf of bread"[Dick (1981)]. Emmy Noether's ashes were interred on the campus of Bryn Mawr College, in an honored place, under the cloister walkway of the Martha Carey Thomas Library.[17]

With Emmy Noether's premature death, besides the personal loss that befell her friends and extended family, and the loss to mathematics and physics, one also mourns her loss to the culture of science. She was and still is an icon for women in the sciences, a model of integrity for not allowing injustice to turn a victim towards bitterness, and an inspiring intellectual companion for anyone who appreciates a life of the mind.[18]

Emmy Noether's achievements live beyond her physical presence. She left a great legacy to mathematics, as in her studies of invariants, abstract algebra (notably the theory of rings and ideals), hypercomplex numbers, and applications of group theory to combinatorial topology [Brewer & Smith (1982) 22], to name a few of her important contributions. To mathematical physics, she also left a splendid legacy through her powerful theorem celebrated in this book. She advised a dozen doctoral students, authored over 40 publications,[19] and inspired the "Noether school" of algebraists. Today the Emmy Noether Gymnasium in Berlin and the Emmy Noether Gymnasium in Erlangen honor her memory. In 1980 the Association for Women in Mathematics established the annual Emmy Noether Lecture, to "honor women who have made fundamental and sustained contributions to the mathematical sciences." From a place of honor among the mathematical sciences, Emmy Noether's theorem of 1918 forms a central organizing principle for the great range of physics.

---

[17]Martha Carey Thomas (1857–1935) was a first dean and second president of Bryn Mawr College. She was an active suffragist, and set a tone of excellence through her vision, tenacity, and independent spirit. When Thomas was denied admission to American universities because she happened to be female, she earned her PhD in linguistics at the University of Zürich. She and Emmy Noether must have been kindred spirits, and it is fitting that their ashes share the same hollowed ground.

[18]For a list of numerous sources on Noether's theorem (first and second), including biographical and historical resources, see Neuenschwander (2014*b*).

[19]See Brewer & Smith (1982) for a list of Noether's publications and descriptions of her doctoral students and their accomplishments; see Dick (1981) for a publications list and the dissertation titles of Noether's 13 doctoral students.

Questions for Reflection and Discussion

Q 1.a. In the argument that leads from translational invariance to the conservation of linear momentum, why was it necessary that the force be derivable from a potential energy function?

Q 1.b. List all the invariants you can think of for the following settings:
a. Newtonian mechanics, with its Galilean transformation between inertial frames;
b. Special relativity, for inertial frames related by a Lorentz transformation;
c. Classical mechanics, for the transformation from an inertial frame to a rotating frame.

Q1.c. A quantity may be conserved but not invariant. Can a quantity be invariant but not conserved?

Q 1.d. Show that if space is isotropic about two points then it is also homogeneous.

Q 1.e. Obtain a translation of Emmy Noether's 1918 paper (or the 1918 original, in German!) [Noether (1918); Noether & Tavel (1971); Kosmann-Schwarzbach (2010)]. As you progress through this text, consult Noether's paper and notice the similarities and differences between her approach and the approach followed here.

Q1.f. If conservation laws follow from symmetries, what symmetry accounts for the conservation of electric charge?

Q1.g. In this book we celebrate symmetry. However, we realize that lack of symmetry, or "symmetry breaking," is also essential to life as we know it.
a. Describe instances where symmetry breaking makes life possible and/or interesting. For specific instances to consider, contemplate asymmetry in architecture; machine design; face recognition; chemical reactions that depend on "handedness" of some molecules; traffic dynamics; and art.
b. Does the second law of thermodynamics have anything to say about the invariance (or lack thereof) of a system's evolution in time?

Exercises

1.1. Show how to express Newton's second law, $\mathbf{F} = d\mathbf{p}/dt$, in terms of the change in
a. mechanical energy, and
b. angular momentum.

c. In each case identify the criteria for the corresponding conservation law, and suggest for it an underlying symmetry of space or time.

1.2. An example from special relativity (see appendix B) illustrates the distinct roles of invariance and conservation in solving a problem. A proton of mass m moves with momentum $\mathbf{p}$ and kinetic energy K through the lab frame and collides with another proton initially at rest. Find in this frame the minimum value of K necessary for the production of a new proton and antiproton pair,

$$p + p \rightarrow 3p + \bar{p}, \tag{1.8}$$

where $\bar{p}$ denotes an antiproton. Hints: use the invariance of $E^2 - (pc)^2 = (mc^2)^2$ between the laboratory and the center-of-mass frames (c = speed of light), and the covariance of expressions for energy and momentum conservation.

Chapter 2

Functionals

Variable quantities called functionals play an important role in many problems arising in analysis, mechanics, geometry, etc. By a functional, we mean a correspondence which assigns a definite (real) number to each function (or curve) belonging to some class. —I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, translated by Richard A. Silverman, 1963

2.1 Single-Integral Functionals

At the foundation of Noether's theorem stand mathematical objects called *functionals*. To invoke Noether's (first) theorem, we make two distinct demands on a functional: (1) that it be an *extremal* (or *stationary*), and (2) that it be *invariant* under a continuous transformation. Before going there, let's get acquainted with these functionials.

Generically, a functional is a mapping from a well-defined set of functions to the real numbers (Figure 2.1). A functional is like a vending machine. Into the input slot you insert a function selected from a set of allowed possibilities. The machine clanks and grinds, and out pops a real number as the output. Definite integrals map a function to a real number, and the functionals of interest in this book are expressed as definite integrals. A few examples provide illustrations.

Example (Distance Functional): I would offend tradition and good sense to not include a standard illustrative example, the "distance functional." In the xy plane, a function $y = y(x)$ describes a path. The distance on the path from point $(x, y) = (a, y(a))$ to point $(x, y) = (b, y(b))$ is found

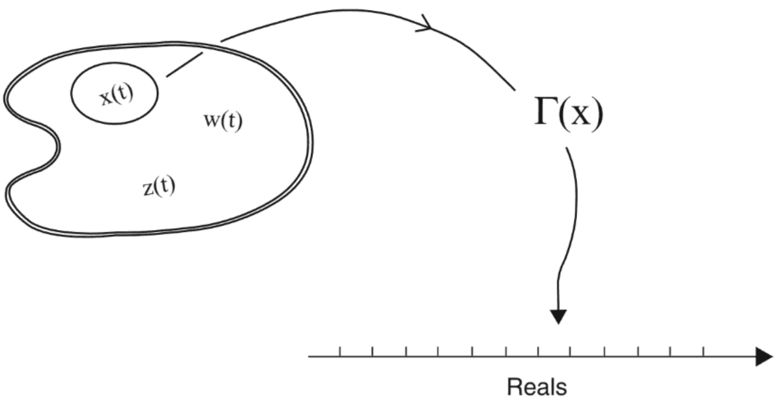

Figure 2.1: *Definition of a functional: a mapping from a set of functions* $\{x(t)\}$ *to the set of real numbers.*

by summing up little increments of length ds. With dx^2 denoting $(dx)^2$, we have

$$\begin{aligned} \text{distance} &= \int_a^b ds \\ &= \int_a^b \sqrt{dx^2 + dy^2} \qquad (2.1) \\ &= \int_a^b \sqrt{1 + y'^2}\, dx, \end{aligned}$$

where y' denotes dy/dx, which we assume exists on the interval $[a, b]$. This requirement defines the set of all permissible functions $\{y(x)\}$. The shortest path constrained to a surface, that connects two points on that surface, is called the *geodesic*.

Example (another Distance Functional): The distance traveled on a journey from Baltimore to Erlangen depends on the path taken. To specify a path one could chart the trajectory in terms of longitude as a function of latitude. For a journey over the surface of the Earth, in spherical coordinates (r, θ, φ) the distance traveled may be computed from the definite integral

$$\text{distance} = \int_{\text{Baltimore}}^{\text{Erlangen}} ds \qquad (2.2)$$

where

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2). \qquad (2.3)$$

At constant radius R from the Earth's center, the distance is given by

$$\text{distance} = R \int_{\text{Baltimore}}^{\text{Erlangen}} \sqrt{1 + \varphi'^2 \sin^2 \theta}\; d\theta \tag{2.4}$$

and $\varphi' \equiv d\varphi/d\theta$. If the trajectory is expressed in terms of longitude as a given function of latitude, $\varphi = \varphi(\theta)$, then we can evaluate the integral and find the distance along that path between the two points on the globe. A sensible airline will find the path that makes the distance traveled a minimum, the geodesic.

Example (Fermat's Principle): A light ray passes through a medium of refractive index n. The light's speed through the medium is $v = ds/dt = c/n$, where c denotes the speed of light in vacuum. Let the ray move in an xy plane. The refractive index may vary with position so that $n = n(x, y)$. When the light goes from a fixed initial point a to a fixed final point b, the elapsed time is

$$\Delta t = \int_{t(a)}^{t(b)} dt. \tag{2.5}$$

The time increment may be written $dt = ds/v = n(x, y)ds/c$:

$$\Delta t = \frac{1}{c} \int_{x(a)}^{x(b)} n(x, y)\sqrt{1 + y'^2}\; dx. \tag{2.6}$$

For any given path $y = y(x)$, one does the integral and computes the elapsed time. Fermat's principle of geometrical optics postulates that the ray's actual trajectory between fixed points will be the one for which Δt is a minimum.[1]

You may notice in functionals that treat one spatial variable as a function of another, as with curves in the xy plane, the integral may be rewritten in terms of a parameter t. This t could be elapsed time from a point of departure, arc length along a curve, and so on. In the first distance functional example, suppose a parameterization $x \to x(t)$ and $y \to y(t)$ is introduced. The particle sets off from $(x, y) = (a, y(a))$ at time $t(a)$ and follows some curve $y = y(x)$ to arrive at $(b, y(b))$ at time $t(b)$. Denoting $\dot{x} \equiv dx/dt$, it follows that $dx = \dot{x}dt$ and similarly for dy. Now the functional may be written parameterically in terms of t as

$$\text{distance} = \int_{t(a)}^{t(b)} \sqrt{\dot{x}^2 + \dot{y}^2}\; dt. \tag{2.7}$$

[1]The "minimum" was stated by Fermat and continues in most introductory optics treatments; see the comment by Jenkins and White in section 5.3.

In the distance functional, the integrand had changed from $\sqrt{1+y'^2}$ integrated over x, to $\sqrt{\dot{x}^2+\dot{y}^2}$ integrated over t. Same information, different parametrization.

The same procedure can be done with the Fermat's principle functional:[2]

$$\begin{aligned}\Delta t &= \frac{1}{c}\int_a^b n(x,y)\sqrt{dx^2+dy^2} \\ &= \frac{1}{c}\int_{t(a)}^{t(b)} n(x(t),y(t))\ \sqrt{\dot{x}^2+\dot{y}^2}\ dt.\end{aligned} \tag{2.8}$$

A system may have more than one dependent variable; for example, in the blue space that birds fly through, a bird's location may be expressed by three spatial coordinates, which together form the components of a position vector. Let the vector components be denoted x^μ, where $\mu = 1, 2, 3$ in ordinary three-dimensional space; in special and general relativity, $\mu = 0,$ $1, 2, 3$ for the four dimensions of spacetime.[3] Denote the functional as Γ, and let its integrand L be a function of the independent variable t, the dependent variables $x^\mu(t)$, and their first derivatives with respect to t. For example, in three-dimensional space, where the coordinates of a particle in motion depend upon a parameter t such as time, Γ takes the form

$$\Gamma = \int_a^b L(t, x^\mu, \dot{x}^\mu)dt \tag{2.9}$$

where $\dot{x}^\mu \equiv dx^\mu/dt$.

Example (Hamilton's Principle): In mechanics, the functional of cardinal importance is the time integral of the difference between a system's kinetic and potential energies:

$$\Gamma = \int_a^b (K-U)dt. \tag{2.10}$$

For a single particle, in classical mechanics the kinetic energy is given by $K = \frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2)$, or in terms of so-called generalized coordinates q^μ,

[2]Note that in parametric representation of the Fermat's principle functional, the coefficient of dt is $nv/c = 1$. This often occurs in parametric representations, where the integrand is a *function*, but it's *numercial value* is a constant. This is not unusual: for an example in elementary mechanics, the conservation of energy can be written $1 = (mv^2)/2E + U(x)/E$, even though v and x are variables.

[3]See appendixes A and B for a discussion of coordinates with upper and lower indices. In Euclidean spaces the distinction between x^μ and x_μ makes no difference, but the distinction becomes important in special and general relativity, and in the geometries of curved spaces.

$K = \frac{1}{2} m g_{\mu\nu} \dot{q}^\mu \dot{q}^\nu$, where the metric tensor components $g_{\mu\nu}$ (see appendix A) turn coordinate displacements into distances[4] and repeated indices are summed. For example, in spherical coordinates,

$$K = \tfrac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 (\sin^2 \theta) \dot{\varphi}^2 \right), \tag{2.11}$$

so that the nonzero $g_{\mu\nu}$ are $g_{rr} = 1, g_{\theta\theta} = r^2, g_{\varphi\varphi} = r^2 \sin^2 \theta$. Through the $g_{\mu\nu}$ the kinetic energy may be a function of the coordinates themselves, in addition to being a function of coordinate velocities. The potential energy is a function of the coordinates and possibly the time, $U = U(t, q^\mu)$. Given a trajectory that specifies a particle's coordinates as functions of time, $q^\mu(t)$, one evaluates Γ for that trajectory. According to Hamilton's principle, a particle's *actual* trajectory between fixed times is one for which Γ is a minimum.[5] More specifically, consider a projectile falling, without air resistance, under the control of a uniform gravitational field directed vertically down, the direction opposite the $+z$ axis. Therefore the potential energy is mgz. The trajectory followed, according to Hamilton's principle, is precisely the path for which this integral is a minimum:

$$\Gamma = \int_a^b \left(\tfrac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz \right) dt. \tag{2.12}$$

Example (Special Relativity and Free Fall in Gravity-Free Spacetime): No matter which reference frame measures proper time between any two events,[6] all inertial observers can use their reference frame's time and space measurements (dt, dx, dy, dz) to compute the proper time[7] increment $d\tau$ with the aid of the invariant spacetime interval, according to

$$c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dx^2. \tag{2.13}$$

One way of stating the postulates of special relativity asserts a relativistic analog to Fermat's principle: Of all world lines through spacetime that

---

[4]In Euclidean space mapped with Cartesian coordinates, $g_{\mu\nu} = \delta_{\mu\nu}$, elements of the unit matrix, or Kronecker delta. If you are not familiar with tensors yet, for now it suffices to think of a two-index tensor as a matrix. However, under coordinate transformations there are additional requirements put on tensors that do not necessarily apply to matrices; see appendix A.

[5]Some authors say that the trajectory actually followed is the one for which Γ is *extremal*, others say *stationary*. See section 5.3.

[6]The distinction between two *places* and two *events* should be noted.

[7]As a particle's world line takes it from spacetime event a to event b, the proper time is the "wristwatch time" recorded by the particle itself; in other words, the time between the two events as measured in the reference frame where they occur at the same place. See Taylor & Wheeler (2000).