Sources and Studies
in the History of Mathematics and
Physical Sciences

For other titles published in this series, go to
http://www.springer.com/series/4142

Emmy Noether (1882–1935)
(photograph courtesy of the Emmy Noether Foundation, Bar Ilan University)

Yvette Kosmann-Schwarzbach

The Noether Theorems

Invariance and Conservation Laws in the Twentieth Century

Translated by Bertram E. Schwarzbach

Springer

Yvette Kosmann-Schwarzbach
Centre de Mathématiques Laurent Schwartz
École Polytechnique
91128 Palaiseau
France
yks@math.polytechnique.fr

Bertram E. Schwarzbach
(Translator)

ISBN 978-0-387-87867-6 e-ISBN 978-0-387-87868-3
DOI 10.1007/978-0-387-87868-3
Springer New York Dordrecht Heidelberg London

Mathematics Subject Classification (2010): 01-02, 22-03, 22E70, 49-03, 49S05, 70-03, 70H03, 70H33, 83-03

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

In memory of Yseult ע״ה
who liked science as well as history

Preface

What follows thus depends upon
a combination of the methods of
the formal calculus of variations
and of Lie's theory of groups.

Emmy Noether, 1918

This book is about a fundamental text containing two theorems and their converses which established the relation between symmetries and conservation laws for variational problems. These theorems, whose importance remained obscure for decades, eventually acquired a considerable influence on the development of modern theoretical physics, and their history is related to numerous questions in physics, in mechanics and in mathematics. This text is the article “Invariante Variationsprobleme” by Emmy Noether, which was published in 1918 in the *Göttinger Nachrichten*, and of which we present an English translation in Part I of this book.

The translation of Noether’s article is followed, in Part II, by a detailed analysis of its inception, as well as an account of its reception in the scientific community. As the background to Noether’s research, we sketch some developments in the theory of invariants in the nineteeth century which culminated in the definition and study of differential invariants, we discuss several works in mechanics dating from the beginning of the twentieth century in which Sophus Lie’s infinitesimal methods in the theory of groups began to be applied, and we show that the immediate motivation for her work was related to questions arising from Einstein’s general theory of relativity of 1915. We then summarize the contents of Noether’s article in modern language. In the subsequent chapters, we review the way in which Noether’s contemporaries, the mathematicians Felix Klein, David Hilbert and Hermann Weyl, and the physicists Einstein and Wolfgang Pauli, acknowledged or failed to acknowledge her contribution; then we outline the quite different diffusions of her first and second theorems. Finally, we outline the genuine generalizations of Noether’s results that began to appear after 1970, in the field of the calculus of variations and in the theory of integrable systems.

The present edition is based on the second edition of *Les Théorèmes de Noether. Invariance et lois de conservation au XX^e^ siècle* (Palaiseau: Éditions de l’École Polytechnique, 2006). For this English edition, the French text has been considerably revised and augmented, with much new information and additional references.

Paris, July 2010

Acknowledgments

It is a pleasure to thank the many colleagues and friends who have made the writing of this book possible. When the French version first appeared in 2004, it was at the insistance of Pierre Cartier whose subsequent support was unfailing, with Alain Guichardet's generous help, and with the much appreciated contribution of Laurent Meersseman who had translated the "Invariante Variationsprobleme" into French and written a draft of the first section of a commentary.

During the preparation of the French and English texts, I received information and assistance on various questions from Alain Albouy, Sophie Bade, Henri Besson, Katherine Brading, Harvey Brown, Leo Corry, Thibault Damour, Jean Eisenstaedt, Benjamin Enriquez, Robert Gergondey, Patrick Iglésias-Zemmour, Michel Janssen, Christoph Kopper, Franco Magri, Hartmann Römer, Peter Roquette, Volodya Rubtsov, David Rowe, Erhard Scholz, Jean-Marie Souriau, Jim Stasheff, Alexander Vinogradov, Salomon Wald, Stefan Waldmann and Scott Walter. To all of them—and to any colleagues whose names I may have omitted inadvertently—I am very grateful. I especially thank Peter Olver who made many useful remarks on the manuscript of the first French edition, based on his thorough knowledge of the subject, and again answered my questions when I was preparing the second edition in 2006. The three referees for the first edition of the French text made valuable recommendations and suggestions, from which I benefitted. The help I received from David Schimmel in translating several difficult German texts remains a treasured memory of this learned friend who died in a road accident in 2007.

Thanks are also due to Frédéric Zantonio at the Centre Polymédia de l'École Polytechnique who helped me with the reproductions. The editors of Springer New York, Ann Kostant, Hans Koelsch, and Elizabeth Loew, extended their support during the preparation of this book, for which I thank them cordially.

I am grateful to the institutions which have allowed me to reproduce documents in their collections: the Niedersächsische Staats- und Universitätsbibliothek, Göttingen, the Einstein archive in Jerusalem, and the archive of Bryn Mawr College, and to Princeton University Press for permission to quote from several volumes of *The Collected Papers of Albert Einstein*.

Last but not least, I thank the translator, Bertram E. Schwarzbach, for his work and for his advice on how to clarify and improve the French text.

Yvette Kosmann-Schwarzbach

July 2010

Contents

Part I
“Invariant Variational Problems” by Emmy Noether
Translation of “Invariante Variationsprobleme” (1918)

Invariante Variationsprobleme.

(F. Klein zum fünfzigjährigen Doktorjubiläum.)

Von

Emmy Noether in Göttingen.

Vorgelegt von F. Klein in der Sitzung vom 26. Juli 1918[1]).

Es handelt sich um Variationsprobleme, die eine kontinuierliche Gruppe (im Lieschen Sinne) gestatten; die daraus sich ergebenden Folgerungen für die zugehörigen Differentialgleichungen finden ihren allgemeinsten Ausdruck in den in § 1 formulierten, in den folgenden Paragraphen bewiesenen Sätzen. Über diese aus Variationsproblemen entspringenden Differentialgleichungen lassen sich viel präzisere Aussagen machen als über beliebige, eine Gruppe gestattende Differentialgleichungen, die den Gegenstand der Lieschen Untersuchungen bilden. Das folgende beruht also auf einer Verbindung der Methoden der formalen Variationsrechnung mit denen der Lieschen Gruppentheorie. Für spezielle Gruppen und Variationsprobleme ist diese Verbindung der Methoden nicht neu; ich erwähne Hamel und Herglotz für spezielle endliche, Lorentz und seine Schüler (z. B. Fokker), Weyl und Klein für spezielle unendliche Gruppen[2]). Insbesondere sind die zweite Kleinsche Note und die vorliegenden Ausführungen gegenseitig durch einander beein-

1) Die endgiltige Fassung des Manuskriptes wurde erst Ende September eingereicht.

2) Hamel: Math. Ann. Bd. 59 und Zeitschrift f. Math. u. Phys. Bd. 50. Herglotz: Ann. d. Phys. (4) Bd. 36, bes. § 9, S. 511. Fokker, Verslag d. Amsterdamer Akad., 27./1. 1917. Für die weitere Litteratur vergl. die zweite Note von Klein: Göttinger Nachrichten 19. Juli 1918.

In einer eben erschienenen Arbeit von Kneser (Math. Zeitschrift Bd. 2) handelt es sich um Aufstellung von Invarianten nach ähnlicher Methode.

First page of "Invariante Variationsprobleme"
Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse, 1918, pp. 235–257.

INVARIANT VARIATIONAL PROBLEMS

(For F. Klein, on the occasion of the fiftieth anniversary of his doctorate)

by **Emmy Noether** in Göttingen

Presented by F. Klein at the session of 26 July 1918*

We consider variational problems which are invariant[A] under a continuous group (in the sense of Lie); the consequences that are implied for the associated differential equations find their most general expression in the theorems formulated in §1, which are proven in the subsequent sections. For those differential equations that arise from variational problems, the statements that can be formulated are much more precise than for the arbitrary differential equations that are invariant under a group, which are the subject of Lie's researches. What follows thus depends upon a combination of the methods of the formal calculus of variations and of Lie's theory of groups. For certain groups and variational problems this combination is not new; I shall mention Hamel and Herglotz for certain finite groups, Lorentz and his students (for example, Fokker), Weyl and Klein for certain infinite groups.[1] In particular, Klein's second note and the following developments were mutually influential, and for this reason I take the liberty of referring to the final remarks in Klein's note.

1 Preliminary Remarks and the Formulation of the Theorems

All the functions that will be considered here will be assumed to be analytic or at least continuous and continuously differentiable a finite number of times, and single-valued within the domain that is being considered.

By the term "transformation group" one usually refers to a system of transformations such that for each transformation there exists an inverse which is an element of the system, and such that the composition of any two transformations of the system is again an element of the system. The group is called a *finite continuous* [group] $\mathfrak{G}_\rho$ when its transformations can be expressed in a general form which depends analytically on ρ *essential* parameters ε (i.e., the ρ parameters cannot be represented by ρ functions of a smaller number of parameters). In the same way, one speaks of an *infinite continuous* group $\mathfrak{G}_{\infty\rho}$ for a group whose most general transformations depend on ρ essential arbitrary functions $p(x)$ and their derivatives in a way that is

* The definitive version of the manuscript was prepared only at the end of September.

[A] *gestatten*, to permit, in the sense of admitting [an invariance group] has been translated as "being invariant under [the action of] a group" (Translator's note).

[1] Hamel, Math. Ann., vol. 59, and Zeitschrift f. Math. u. Phys., vol. 50. Herglotz, Ann. d. Phys. (4) vol. 36, in particular §9, p. 511. Fokker, Verslag d. Amsterdamer Akad., 27/1 1917. For a more complete bibliography, see Klein's second note, Göttinger Nachrichten, 19 July 1918.

In a paper by Kneser that has just appeared (Math. Zeitschrift, vol. 2), the determination of invariants is dealt with by a similar method.

Y. Kosmann-Schwarzbach, *The Noether Theorems*, Sources and Studies in the History of Mathematics and Physical Sciences, DOI 10.1007/978-0-387-87868-3_1,

analytical or at least continuous and continuously differentiable a finite number of times. An intermediate case is the one in which the groups depend on an infinite number of parameters but not on arbitrary functions. Finally, one calls a group that depends not only on parameters but also on arbitrary functions a *mixed group*.[2]

Let $x_1,\dots,x_n$ be independent variables, and let $u_1(x),\dots,u_\mu(x)$ be functions of these variables. If one subjects the x and the u to the transformations of a group, then one should recover, among all the transformed quantities, precisely n independent variables, $y_1,\dots,y_n$, by the assumption of invertibility of the transformations; let us call the remaining transformed variables that depend on them $v_1(y),\dots,v_\mu(y)$. In the transformations, the derivatives of u with respect to x, that is to say $\frac{\partial u}{\partial x},\frac{\partial^2 u}{\partial x^2},\cdots$, may also occur.[3] A function is said to be an *invariant* of the group if there is a relation

$$P\left(x,u,\frac{\partial u}{\partial x},\frac{\partial^2 u}{\partial x^2},\cdots\right)=P\left(y,v,\frac{\partial v}{\partial y},\frac{\partial^2 v}{\partial y^2},\cdots\right).$$

In particular, an integral I is an invariant of the group if it satisfies the relation

$$(1)\qquad \begin{aligned} I &= \int\cdots\int f\left(x,u,\frac{\partial u}{\partial x},\frac{\partial^2 u}{\partial x^2},\cdots\right)dx \\ &= \int\cdots\int f\left(y,v,\frac{\partial v}{\partial y},\frac{\partial^2 v}{\partial y^2},\cdots\right)dy\ ^4 \end{aligned}$$

integrated over an *arbitrary* real domain in x, and over the corresponding domain in y.[5]

On the other hand, I calculate for an arbitrary integral I, which is not necessarily invariant, the first variation δI, and I transform it, according to the rules of the

[2] Lie defines, in the "Grundlagen für die Theorie der unendlichen kontinuierlichen Transformationsgruppen" ["Basic Principles of the Theory of Infinite Continuous Transformation Groups"], Ber. d. K. Sächs. Ges. der Wissensch. 1891 (to be cited henceforth as "Grundlagen"), the infinite continuous groups as transformation groups whose elements are given by the most general solutions of a system of partial differential equations provided that these solutions do not depend exclusively on a finite number of parameters. Thus one obtains one of the above-mentioned cases distinct from that of a finite group, while, on the other hand, the limiting case of an infinite number of parameters does not necessarily satisfy a system of differential equations.

[3] I omit the indices here, and in the summations as well whenever it is possible, and I write $\frac{\partial^2 u}{\partial x^2}$ for $\frac{\partial^2 u_\alpha}{\partial x_\beta \partial x_\gamma}$, etc.

[4] I write dx, dy for $dx_1\dots dx_n$, $dy_1\dots dy_n$ for short.

[5] All the arguments x, u, ε, $p(x)$ that occur in the transformations must be assumed to be real, while the coefficients may be complex. Since the final results consist of identities among the x, the u, the parameters and the arbitrary functions, these identities are valid as well for the complex domain, once one assumes that all the functions that occur are analytic. In any event, a major part of the results can be proven without integration, so a restriction to the real domain is not necessary for the proof. However, the considerations at the end of §2 and at the beginning of §5 do not seem to be valid without integration.

caculus of variations, by integration by parts. Once one assumes that δu and all the derivatives that occur vanish on the boundary, but remain arbitrary elsewhere, one obtains the well-known result,

(2) $$\delta I = \int \cdots \int \delta f \, dx = \int \cdots \int \left(\sum \psi_i \left(x, u, \frac{\partial u}{\partial x}, \cdots \right) \delta u_i \right) dx,$$

where ψ represents the *Lagrangian expressions*, that is to say, the left-hand side of the Lagrangian equations of the associated variational problem $\delta I = 0$. To that integral relation there corresponds an *identity* without an integral in δu and its derivatives that one obtains by adding the boundary terms. As an integration by parts shows, these boundary terms are integrals of divergences, that is to say, expressions

$$\operatorname{Div} A = \frac{\partial A_1}{\partial x_1} + \cdots + \frac{\partial A_n}{\partial x_n},$$

where A is linear in δu and its derivatives. From that it follows that

(3) $$\sum \psi_i \delta u_i = \delta f + \operatorname{Div} A.$$

In particular, if f contains only the first derivatives of u, then, in the case of a simple integral, identity (3) is identical to Heun's "central Lagrangian equation,"

(4) $$\sum \psi_i \delta u_i = \delta f - \frac{d}{dx} \left(\sum \frac{\partial f}{\partial u_i'} \delta u_i \right), \qquad \left(u_i' = \frac{du_i}{dx} \right),$$

while for an n-fold integral, (3) becomes

(5) $$\sum \psi_i \delta u_i = \delta f - \frac{\partial}{\partial x_1} \left(\sum \frac{\partial f}{\partial \frac{\partial u_i}{\partial x_1}} \delta u_i \right) - \cdots - \frac{\partial}{\partial x_n} \left(\sum \frac{\partial f}{\partial \frac{\partial u_i}{\partial x_n}} \delta u_i \right).$$

For the simple integral and κ derivatives of the u, (3) yields

(6) $$\sum \psi_i \delta u_i = \delta f -$$
$$- \tfrac{d}{dx} \left\{ \Sigma \left(\binom{1}{1} \tfrac{\partial f}{\partial u_i^{(1)}} \delta u_i + \binom{2}{1} \tfrac{\partial f}{\partial u_i^{(2)}} \delta u_i^{(1)} + \cdots + \binom{\kappa}{1} \tfrac{\partial f}{\partial u_i^{(\kappa)}} \delta u_i^{(\kappa-1)} \right) \right\} +$$
$$+ \tfrac{d^2}{dx^2} \left\{ \Sigma \left(\binom{2}{2} \tfrac{\partial f}{\partial u_i^{(2)}} \delta u_i + \binom{3}{2} \tfrac{\partial f}{\partial u_i^{(3)}} \delta u_i^{(1)} + \cdots + \binom{\kappa}{2} \tfrac{\partial f}{\partial u_i^{(\kappa)}} \delta u_i^{(\kappa-2)} \right) \right\} +$$
$$+ \cdots + (-1)^\kappa \tfrac{d^\kappa}{dx^\kappa} \left\{ \Sigma \binom{\kappa}{\kappa} \tfrac{\partial f}{\partial u_i^{(\kappa)}} \delta u_i \right\},$$

and there is a corresponding identity for an n-fold integral; in particular, A contains δu and its derivatives up to order $\kappa - 1$. That the Lagrangian expressions ψ_i are actually defined by (4), (5) and (6) is a result of the fact that, by the combinations of the right-hand sides, all the higher derivatives of the δu are eliminated, while, on the other hand, relation (2), which one clearly obtains by an integration by parts, is satisfied.

In what follows we shall examine the following two theorems:

I. *If the integral I is invariant under a* [*group*] $\mathfrak{G}_\rho$*, then there are* ρ *linearly independent combinations among the Lagrangian expressions which become divergences—and conversely, that implies the invariance of I under a* [*group*] $\mathfrak{G}_\rho$*. The theorem remains valid in the limiting case of an infinite number of parameters.*

II. *If the integral I is invariant under a* [*group*] $\mathfrak{G}_{\infty\rho}$ *depending on arbitrary functions and their derivatives up to order* σ*, then there are* ρ *identities among the Lagrangian expressions and their derivatives up to order* σ*. Here as well the converse is valid.*[6]

For mixed groups, the statements of these theorems remain valid; thus one obtains identities[B] as well as divergence relations independent of them.

If we pass from these identity relations to the associated variational problem, that is to say, if we set $\psi = 0$,[7] then Theorem I states in the one-dimensional case—where the divergence coincides with a total differential—the existence of ρ first integrals among which, however, there may still be nonlinear identities;[8] in higher dimensions one obtains the divergence equations that, recently, have often been referred to as "conservation laws." Theorem II states that ρ Lagrangian equations are a consequence of the others.[C]

The simplest example for Theorem II—without its converse—is Weierstrass's parametric representation; here, as is well known, the integral is invariant in the case of homogeneity of the first order when one replaces the independent variable x by an arbitrary function of x which leaves u unchanged $(y = p(x);\ v_i(y) = u_i(x))$. Thus an arbitrary function occurs though none of its derivatives occurs, and to this corresponds the well-known linear relation among the Lagrangian expressions themselves, $\sum \psi_i \dfrac{du_i}{dx} = 0$. Another example is offered by the physicists' "general theory of relativity"; in this case the group is the group of *all* the transformations of the $x : y_i = p_i(x)$, while the u (called $g_{\mu\nu}$ and q) are thus subjected to the transformations induced on the coefficients of a quadratic and of a linear differential form, respectively transformations which contain the first derivatives of the arbitrary functions $p(x)$. To that there correspond the n known identities among the Lagrangian expressions and their first derivatives.[9]

[6] For some trivial exceptions, see §2, note 13.

[B] *Abhängigkeit*, dependence, has been translated by "identity." *Identität* has been translated by "identity" or "identity relation." Both *Relation* and *Beziehung* have been translated by "relation" and *Verbindung* by "combination" (Translator's note).

[7] More generally, one can also set $\psi_i = T_i$; see §3, note 15.

[8] See the end of §3.

[C] I.e., among the Lagrangian equations, ρ equations are consequences of the remaining ones (Translator's note).

[9] For this, see Klein's presentation.

If, in particular, one considers a group such that there is no derivative of the $u(x)$ in the transformations, and that furthermore the transformed independent quantities depend only on the x and not on the u, then (as is proven in §5) from the invariance of I, the relative invariance of $\sum \psi_i \delta u_i$[10] follows, and also that of the divergences that appear in Theorem I, once the parameters are subjected to appropriate transformations. From that it follows as well that the first integrals mentioned above are also invariant under the group. For Theorem II, the relative invariance of the left-hand sides of the identities, expressed in terms of the arbitrary functions, follows, and consequently another function whose divergence vanishes identically and which is invariant under the group—which, in the physicists' theory of relativity, establishes the link between identities and law[D] of energy.[11] Theorem II ultimately yields, in terms of group theory, the proof of a related assertion of Hilbert concerning the lack of a proper law of energy in "general relativity." As a result of these additional remarks, Theorem I includes all the known theorems in mechanics, etc., concerning first integrals, while Theorem II can be described as the maximal generalization in group theory of "general relativity."

2 Divergence Relations and Identities

Let $\mathfrak{G}$ be a continuous group—finite or infinite; one can always assume that the identity transformation corresponds to the vanishing of the parameters ε, or to the vanishing of the arbitrary functions $p(x)$,[12] respectively. The most general transformation is then of the form

$$y_i = A_i\left(x, u, \frac{\partial u}{\partial x}, \cdots\right) = x_i + \Delta x_i + \cdots$$
$$v_i(y) = B_i\left(x, u, \frac{\partial u}{\partial x}, \cdots\right) = u_i + \Delta u_i + \cdots,$$

where Δx_i, Δu_i are the terms of lowest degree in ε, or in $p(x)$ and its derivatives, respectively, and we shall assume that in fact they are *linear*. As we shall show further on, this does not restrict the generality.

---

[10] This is to say that $\sum \psi_i \delta u_i$ is invariant under the transformation up to a multiplicative factor.

[D] *Energiesatz* has been translated literally as "law of energy," in the sense of "law of conservation of energy," just as, *infra*, in §6, *eigentlich Energiesatz*, has been translated as "proper law of energy," in the sense of "proper law of conservation of energy" (Translator's note).

[11] See Klein's second note.

[12] Cf. Lie, "Grundlagen," p. 331. When dealing with arbitrary functions, it is necessary to replace the special values a^σ of the parameters by fixed functions $p^\sigma, \frac{\partial p^\sigma}{\partial x}, \cdots$; and correspondingly the values $a^\sigma + \varepsilon$ by $p^\sigma + p(x), \frac{\partial p^\sigma}{\partial x} + \frac{\partial p}{\partial x}$, etc.

Now let the integral I be invariant under $\mathfrak{G}$; then relation (1) is satisfied. In particular, I is also invariant under the infinitesimal transformations contained in $\mathfrak{G}$,

$$y_i = x_i + \Delta x_i; \qquad v_i(y) = u_i + \Delta u_i,$$

and therefore relation (1) becomes

$$(7) \qquad 0 = \Delta I = \int \cdots \int f\left(y, v(y), \frac{\partial v}{\partial y}, \cdots\right) dy - \int \cdots \int f\left(x, u(x), \frac{\partial u}{\partial x}, \cdots\right) dx,$$

where the first integral is defined on a domain in $x + \Delta x$ corresponding to the domain in x. But this integration can be replaced by an integration on the domain in x by means of the transformation

$$(8) \qquad \int \cdots \int f\left(y, v(y), \frac{\partial v}{\partial y}, \cdots\right) dy = \int \cdots \int f\left(x, v(x), \frac{\partial v}{\partial x}, \cdots\right) dx + \int \cdots \int \mathrm{Div}(f.\ \Delta x)\, dx,$$

which is valid for infinitesimal Δx, If, instead of the infinitesimal transformation Δu, one introduces the variation

$$(9) \qquad \bar{\delta} u_i = v_i(x) - u_i(x) = \Delta u_i - \sum \frac{\partial u_i}{\partial x_\lambda} \Delta x_\lambda,$$

(7) and (8) thus become

$$(10) \qquad 0 = \int \cdots \int \{\bar{\delta} f + \mathrm{Div}(f.\ \Delta x)\} dx.$$

The right-hand side is the classical formula for the simultaneous variation of the dependent and independent variables. Since relation (10) is satisfied by integration on an *arbitrary* domain, the integrand must vanish identically; Lie's differential equations for the invariance of I thus become the relation

$$(11) \qquad \bar{\delta} f + \mathrm{Div}(f.\ \Delta x) = 0.$$

If, using (3), one expresses $\bar{\delta} f$ here in terms of the Lagrangian expressions, one obtains

$$(12) \qquad \sum \psi_i \bar{\delta} u_i = \mathrm{Div}\, B \qquad (B = A - f.\ \Delta x),$$