Emmy Noether’s Enduring Legacy in Symmetry

Peter J. Olver
School of Mathematics
University of Minnesota
Minneapolis, MN 55455
olver@umn.edu
http://www.math.umn.edu/~olver

Abstract. A short appreciation of Emmy Noether’s original contributions to symmetry analysis of differential equations and variational problems.

The essays appearing in this thematic issue rightly celebrate the profound contributions to mathematics, physics, and beyond, made by Emmy Noether in her seminal contribution [35] published precisely one century ago. As is now well known, in this paper, she states and proves two fundamental theorems concerning symmetries of variational problems. The First Noether Theorem establishes the connection between continuous variational symmetry groups and conservation laws of their associated Euler–Lagrange equations. The Second Noether Theorem deals with the case when the variational symmetry group is infinite-dimensional, depending on one or more arbitrary functions of the independent variables, e.g., the gauge symmetry groups arising in relativity and physical field theories. The first conclusion is that the conservation laws associated with its one-parameter subgroups are all trivial, meaning that they provide no information concerning the behavior of solutions to the field equations. This result enabled Noether to fully explain the triviality of the energy conservation law in general relativity, which was perplexing Einstein, Hilbert, and Klein and was the primary reason that the latter two mathematicians had invited her, as an expert in invariant theory, to visit Göttingen during the period of Hilbert’s intense rivalry with Einstein, when they were vying to establish the foundations of general relativity. Noether explained that the triviality of the energy conservation law follows from the fact that its associated time translational symmetries form a subgroup of such an infinite-dimensional variational symmetry group of the relativistic Hilbert Lagrangian, thus resolving the relativistic energy conundrum.

Soon after Noether’s paper appeared, Bessel–Hagen, [3], implementing Noether’s own suggestion, formulated a useful extension that relaxes the infinitesimal invariance condition by allowing a divergence term, leading to what are now known as divergence symmetries.
Every divergence (and ordinary) symmetry of the variational principle is necessarily a symmetry of the associated Euler–Lagrange equations, although the converse is not always valid. However, almost all known counterexamples are (in the appropriate coordinates) scaling transformation groups; the only exception I know is the system governing three-dimensional linear isotropic elasticity which admits a first order generalized symmetry whose associated one-parameter group is, intriguingly, prescribed by the solutions to Maxwell’s electromagnetic equations, \([37]\).

Much later, in the mid 1980’s, I revisited Noether’s First Theorem, \([38]\), and proved a refined version that states that, as long as the system of Euler–Lagrange equations is “normal”, meaning that it can, in some coordinate system, be placed in Cauchy–Kovalevskaya form, then the Noether/Bessel–Hagen construction defines a one-to-one correspondence between nontrivial one-parameter variational symmetry groups and nontrivial conservation laws. “Abnormal” systems include those covered by Noether’s Second Theorem, which states that a variational problem admits an infinite-dimensional symmetry group of the aforementioned type if and only if its Euler–Lagrange equations are underdetermined, in the sense that they admit a nontrivial differential relation, i.e., some combination of their derivatives vanishes identically. In general relativity, this takes the form of a Bianchi identity for (pseudo-)Riemannian geometry. The preceding results motivate an intriguing unresolved problem, \([38]\): normal systems match nontrivial symmetries with nontrivial conservation laws; underdetermined systems match nontrivial symmetries with trivial conservation laws; it is still not known whether there are any overdetermined systems that match trivial symmetries with nontrivial conservation laws through the Noether correspondence. More recently, Hydon and Mansfield, \([17]\), have formulated an “intermediate Noether Theorem”, that produces additional identities/conservative structures for the Euler–Lagrange equations of variational problems admitting infinite-dimensional symmetry groups that do not depend upon completely arbitrary functions.

I refer to my book, \([38]\), for rigorous statements and proofs of the two basic Noether Theorems in their general form. Let me also highlight Yvette Kosmann–Schwarzbach’s magnificent monograph, \([24]\) (or \([23]\) for the French original), which contains an excellent English (respectively French) translation of Noether’s original paper, and delves into the strange and sordid history of her Theorems, exposing a profusion of misunderstandings, misattributions, and misinterpretations, that, sadly, continues to this day, \([25]\). Here, though, I would like to focus on what is the least appreciated, but nevertheless similarly profound contribution to symmetry analysis that Noether made in the same paper, and one that is particularly apropos to this journal. Namely, it is the first place in the literature that higher order generalized symmetries appear. Essentially ignored by both the mathematics and physics communities until the discovery of solitons and infinite-dimensional integrable systems in the 1960’s, generalized symmetries nowadays play a vital role in contemporary research in the latter subject. Thus, I would like to take this opportunity to give Noether her proper due by celebrating yet another of her extraordinary mathematical insights.

Let me begin by setting the stage with some mathematical prehistory leading up to Noether’s paper. Noether wrote her Ph.D. thesis, which consisted of a complicated technical calculation of all 331 invariants of a ternary biquadratic form (homogeneous polynomial), under Paul Gordan, the pre-Hilbertian “king” of classical invariant theory.
(In her later incarnation as the founder of the modern theory of abstract algebra, she went so far as to disavow her thesis, referring to it as “Mist”, which, in English, translates as “manure” or perhaps worse.) Further, before arriving in Göttingen, Noether became au fait with Lie’s theory of continuous groups — what we nowadays know as Lie transformation groups or, in the infinite-dimensional case, Lie pseudo-groups, [7, 40]. An aside: a still unmet theoretical challenge is that the mathematical theory of Lie pseudo-groups remains, to this day, underdeveloped. Remarkably, in contrast to the comprehensive modern theory of finite-dimensional Lie groups, there is still no abstract object that adequately represents an infinite-dimensional Lie pseudo-group, which remains inextricably tied to the space it acts on.

In 1897, Lie, [28], had already explicitly formulated the infinitesimal criterion for the invariance of a variational problem under a continuous transformation group. He also noted that, for a given group, the most general invariant Lagrangian must be a function of the associated differential invariants times an invariant volume form. At this point it is worth making a digression to expound on the (surprisingly) as yet unacknowledged potential physical implications of the latter result. In modern day physics, the usual mechanism for constructing a field theory is to start with a suitable configuration space — whose coordinates consist of the independent variables, such as space, time, and other relevant physical quantities, along with the dependent variables that characterize the physical fields, and can be mathematically identified as the fiber coordinates of some bundle over the manifold coordinatized by the former. One then postulates a collection of “physical” symmetries that are to be admitted by the theory, which may include classical Galilean invariance, relativistic Poincaré invariance, gauge symmetries, supersymmetry, and/or yet more exotic species such as those arising in string theories (e.g., $E_8 \times E_8$, SO(32), etc.). The next step is to devise a variational principle over the configuration space that admits the desired symmetries. In practice, this means one constructed from the mathematically “simplest” invariant Lagrangian. The field equations are then provided by the Euler–Lagrange equations obtained by taking the first variation of the invariant functional. Their solutions are the critical points (maxima, minima, saddle points, etc.) of the underlying invariant variational functional or action principle. Moreover, since we began with an invariant Lagrangian, Noether’s Theorems guarantee the existence of corresponding conservation laws and/or differential relations satisfied by the field equations. However, what is, apparently, never addressed among the physics community is what happens if one adopts an alternative invariant Lagrangian to formulate the physical theory. More specifically, if one multiplies the “physical” invariant Lagrangian by any differential invariant admitted by the symmetry group — of which there are an infinite number of possibilities — one obtains another invariant variational principle. How does this affect the resulting physics governed by the ostensibly different field equations? A priori, the only property the two variational principles have in common is that they both admit the same variational symmetry group. So, either the underlying physics is, in some rather vague and as yet undefined sense, “the same”, in which case choosing the simplest invariant Lagrangian makes sense on purely practical grounds, or, more worryingly, the two physical theories are different, which then begs the question as to which invariant variational problem describes the correct physics — how does one decide among an infinite range of possibilities? One answer might be
that one chooses “the simplest”, e.g., the one of lowest order in the derivatives of the field variables — which may not itself be uniquely specified. But then what are the physical criteria dictating this choice? Does Nature design the physics governing the universe not just from the underlying invariance, but also from some innate notion of mathematical simplicity? This is an important and fundamental question that, to the best of my knowledge, has never been adequately addressed in the physics literature. And, to seriously investigate these issues, I am convinced that my equivariant method of moving frames and the invariant variational calculus, \([9, 21, 31, 41]\), provides the proper mathematical tools. Indeed, the moving frame calculus has been recently applied by Gonçalves and Mansfield, \([14, 15]\), to formulate a fully invariant version of Noether’s Theorems in the case of geometrical symmetry groups. On the other hand, there is as yet no moving frame theory that encompasses generalized symmetries.

Returning to our historical development, intimations of the general Noether correspondence between variational symmetry groups and conservation laws began to appear in the nineteenth and early twentieth century literature. The starting point might be ascribed to Lagrange — see \([24]\) for details — who viewed the known physical conservation laws as consequences of the equations of mechanics, and observed correlations with basic symmetry principles. Later, Jacobi, \([18]\), noted that translational and rotational symmetries of space give rise to conservation of linear and angular momentum. In 1897, Schütz, \([44]\), showed how time translational symmetry induces conservation of energy. In 1911, Herglotz, \([16]\), employed Poincaré invariance to construct ten independent conservation laws for special relativistic mechanics on Minkowski space-time, followed by Engel, \([8]\), who went to the non-relativistic limit, demonstrating, in particular that Galilean invariance corresponds to linear motion of the center of mass. But Noether was the first to realize these results were all, in fact, special cases of a completely general correspondence between conservation laws and symmetry properties in the presence of a variational formulation.

So how did this inspire Noether to generalize Lie’s concept of a (variational) symmetry group? Let me attempt to reconstruct her thinking, but please read this with the appropriate grain of salt. In the previously established instances of the correspondence, the symmetry groups that lead to physical conservation laws are all geometrical, meaning they are transformation groups acting on the underlying configuration space (bundle). Noether’s fundamental identity, obtained through a clever integration by parts argument, demonstrated that Lie’s infinitesimal criterion for the invariance of the Lagrangian always produces a corresponding conservation law, which, in the aforementioned contexts, is precisely the apposite physical quantity — momentum, energy, etc. Thus, her initial conclusion was that every one-parameter geometrical variational symmetry group produces a conservation law. But then she must have wondered whether the converse is also true — does every conservation law arise in this way? And here is where her insightful generalization appears. Again, with her fundamental identity in hand, the answer to the latter question was clearly in the affirmative provided one allows the infinitesimal generator to depend not just on the independent and dependent variables but also on the derivatives of the dependent variables. Thus, to formulate the converse result, Noether had to radically extend Lie’s theory of continuous transformation groups, giving birth to the concept of generalized or higher order symmetries.
It should be noted that, in yet another misunderstanding of Noether’s historical role, a significant fraction of the modern literature, e.g., [1], refers to generalized symmetries as “Lie–Bäcklund transformations” in honor of Lie and his contemporary Bäcklund, known for his discovery of Bäcklund transformations. However, these are not symmetry groups, and first arose through an ingenious geometrical construction for producing new constant curvature surfaces in three-dimensional space from known ones. They were then reinterpreted as a means of generating new solutions of an integrable partial differential equation by superimposing an additional soliton on a known solution profile, the prototypical example being the sine–Gordon equation that governs said constant curvature surfaces. Now, it is true that Lie and Bäcklund extended the aforementioned geometrical group actions to include contact transformations, [29, 39], which are geometrical transformations on the space of derivatives of the field variables (now known, following Ehresmann, [7], as a jet bundle), that preserve the innate contact structure. For example, the canonical transformations of Hamiltonian mechanics can be interpreted as contact transformations on the physical configuration space. Lie’s tremendous enthusiasm for this construction was tempered after Bäcklund proved that all contact transformations were merely prolongations of point transformations or, in the single dependent variable case, first order ones, [1, 2, 39]. Indeed, Lie never brought himself to mention this disappointing result in his “definitive” treatise, [29], on the subject! And, while one can interpret contact transformation groups as a rather special type of first order generalized symmetry, [39], nowhere in Lie or Bäcklund’s work does a true generalized symmetry appear. Thus, the aforementioned terminology is ahistorical, and fails to do justice to their true originator. One might thus be tempted to adopt the term “Noether symmetry” to designate a generalized symmetry; however, this would be impossibly confusing since this nomenclature is commonly used to refer to variational symmetries, sometimes of a restricted type, and hence fails to cover the full range of generalized symmetries. Hence my preference for the neutral descriptive term “generalized symmetry”, with proper acknowledgment of Noether’s fundamental role in their origination.

Sadly, for several decades after Noether, generalized symmetries were, for all intents and purposes, ignored by both mathematicians and physicists. As best I can tell, the next time they appear in the literature is in a 1962 paper by Steudel, [46], on the calculus of variations, followed independently by a pair of 1964 papers by H.H. Johnson, [19, 20], who pointed out that the generalized symmetries of a differential equation form a Lie algebra under the natural commutator bracket. But their modern revival was primarily inspired by the amazing numerical discovery of solitons in 1964 by Zabusky and Kruskal, [48], who were trying to understand the unexpected non-ergodic behavior of a Fermi–Pasta–Ulam (FPU) system consisting of masses coupled by nonlinear springs, [10]. They replaced the discrete FPU system by a continuum model, which turned out to be the Korteweg–deVries equation, a nonlinear partial differential equation governing the propagation of long waves over shallow water, [22], first written down, along with its first three conservation laws and one-soliton solution, by Boussinesq in 1877, [4, 6]. Inspired by this initial numerical evidence, the first confirmation of the integrability of the Korteweg–deVries equation was a proof that it admits an infinite hierarchy of higher order conservation laws, [34]. It was gradually understood that these were associated with the commuting flows provided
by the higher order Korteweg–deVries equations discovered by Lax, [27] — see also [13] — which, in fact, can be reinterpreted as generalized symmetries of the original equation, and linked through Noether’s First Theorem via either a Lagrangian structure or, more directly, through the Hamiltonian framework, [38]. With Zakharov and Shabat’s subsequent discovery of the integrability of the nonlinear Schrödinger equation, a model for propagation of optical pulses in fibers and also modulation of periodic wave trains, [49], the floodgates opened and integrable partial differential equations became a major mathematical industry that continues to this day.

Generalized symmetry analysis turned out to be the most productive means for constructing additional interesting examples of integrable systems. Inspired by the recursion relation for the higher order Korteweg–deVries symmetries discovered by Andrew Lenard, [13, 42], along with subsequent investigations of Kumei into the connections between generalized symmetries and Bäcklund transformations, [26], I introduced, in 1977, [36], the notion of a recursion operator as a general mechanism for generating infinite hierarchies of higher order symmetries. In particular, every geometrical symmetry of a linear partial differential equation produces a recursion operator that generates an infinite hierarchy of higher order symmetries, showing that all linear equations with symmetry, which includes all constant coefficient partial differential equations since they admit translational symmetries, automatically possess infinitely many. Linearizable equations such as Burgers’ equation, a simple model of viscous fluid mechanics, also admit recursion operators even in the absence of a variational principle. My construction was quickly followed by the fundamental theorem of Magri, [30], which showed that an evolution equation possessing two compatible Hamiltonian structures — a biHamiltonian system — automatically possesses a recursion operator and an infinite hierarchy of generalized symmetries and associated conservation laws. Yet another method for generating infinite hierarchies of higher order symmetries are the master symmetries, based on a refinement of recursion operators known as hereditary operators, due to Fuchssteiner, [11, 12]. Such recursive schemes have proved to be remarkably effective for establishing the integrability of partial differential equations. In the hands of Mikhailov et. al, [32, 33], the existence of higher order generalized symmetries, reformulated and extended through the notion of a formal symmetry, was turned into a powerful computational tool for explicitly generating and classifying an impressive variety of new integrable systems, many of which remain incompletely investigated. Later, Sanders and Wang, [43], used a deep combination of number theory and symmetry analysis to produce complete classifications of integrable systems of a specified form, e.g., polynomial evolution equations in one space variable. There is now a vast literature on generalized symmetries, recursion operators, bi- and multi-Hamiltonian structures, and related features of integrable systems. In my perhaps biased assessment, the best introduction to the subject remains my first book, [38].

Following her all too brief foray into symmetry and the calculus of variations, Noether soon turned her attention towards establishing the foundations of modern abstract algebra, the endeavor that occupied the remaining years of her tragically shortened career and the one for which she is most celebrated among the majority of mathematicians, many of whom remain un- or under-informed concerning her remarkable contributions to physics, the calculus of variations, and symmetry analysis. Indeed, the extremely influential algebra
text by van der Waerden, [47], who was one of her disciples, collectively known as the “Noether boys”, was in essence a transcription of her lectures establishing the subject. Appreciations of her mathematical work can be found in the commemorative volumes [5, 45], although, in these collections, neither of the short articles concerning her work in the calculus of variations do it justice, nor are her contributions to generalized symmetry analysis even mentioned. I hope this short article helps set the record straight.

Acknowledgments: Thanks to Peter Hydon, Elizabeth Mansfield, Yvette Kosmann–Schwarzbach, and the referees for helpful remarks on earlier versions of this paper.

References


