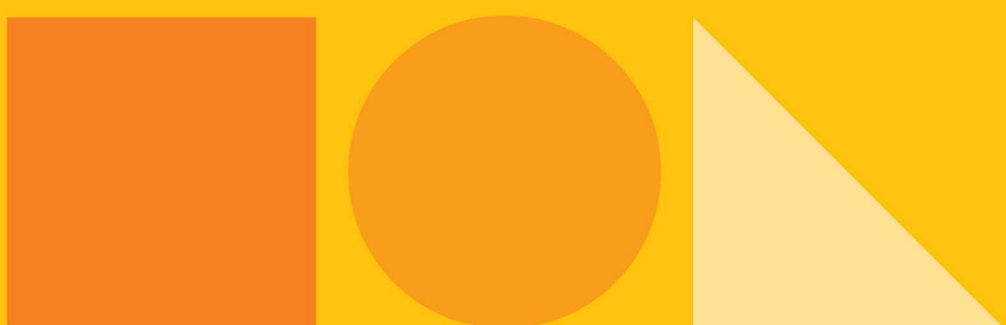
IOP Concise Physics | A Morgan & Claypool Publication

Lectures on Selected Topics in Mathematical Physics: Introduction to Lie Theory with Applications

William A Schwalm



IOP ebooks

Lectures on Selected Topics in Mathematical Physics: Introduction to Lie Theory with Applications

Lectures on Selected Topics in Mathematical Physics: Introduction to Lie Theory with Applications

William A Schwalm
Department of Physics and Astrophysics,
University of North Dakota, USA

Morgan & Claypool Publishers

Copyright © 2017 Morgan & Claypool Publishers

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher, or as expressly permitted by law or under terms agreed with the appropriate rights organization. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency, the Copyright Clearance Centre and other reproduction rights organisations.

Rights & Permissions
To obtain permission to re-use copyrighted material from Morgan & Claypool Publishers, please contact info@morganclaypool.com.

ISBN 978-1-6817-4449-0 (ebook)
ISBN 978-1-6817-4448-3 (print)
ISBN 978-1-6817-4451-3 (mobi)

DOI 10.1088/978-1-6817-4449-0

Version: 20170401

IOP Concise Physics
ISSN 2053-2571 (online)
ISSN 2054-7307 (print)

A Morgan & Claypool publication as part of IOP Concise Physics
Published by Morgan & Claypool Publishers, 40 Oak Drive, San Rafael, CA, 94903 USA

IOP Publishing, Temple Circus, Temple Way, Bristol BS1 6HG, UK

To our teacher, Dr Philip Gold.

Contents

Preface ix

Acknowledgments x

Author biography xi

Introduction xii

Bibliographic notes xiv

1 Groups 1-1

1.1 Permutations and symmetries 1-1

1.2 Subgroups and classes 1-4

1.3 Representations 1-7

1.4 Orthogonality 1-10

References 1-12

2 Lie groups 2-1

2.1 Lie groups as manifolds 2-1

2.2 Lie groups as groups of transformations or substitutions 2-5

2.3 Infinitesimal generators 2-8

2.4 Generator example: Lorentz boost 2-12

2.5 Transformations acting in three or more dimensions 2-14

2.6 Changing coordinates 2-15

2.7 Changing variables in the generator 2-16

2.8 Invariant functions, invariant curves, and groups that permute curves in a family 2-18

2.9 Canonical coordinates for a one-parameter group 2-22

References 2-24

3 Ordinary differential equations 3-1

3.1 Prolongation of the group generator and a symmetry criterion 3-4

3.2 Reformulation of symmetry in terms of partial differential operators 3-8

3.3 Symmetries in terms of A 3-12

3.4 Note on evaluating commutators 3-13

3.5 Symmetries of first-order DEs 3-16

3.6 Tabulating DEs according to groups they admit 3-17
3.7 Lie's integrating factor 3-18
3.8 Finding symmetries of a second order DE 3-21
3.9 Using a symmetry to reduce the order 3-26
3.10 Classical mechanics: Nöther's theorem 3-29
References 3-34

Preface

The idea of a continuous group that is somehow differentiable is attributed to Sophus Lie, who applied such groups to differential equations. A modern definition of a Lie group is that it is a set G which is both a group and a smooth manifold. These two aspects of G connect in the requirement that group multiplication and inversion are smooth maps. To make this definition useful to a reader who is not already familiar with them, words such as group, manifold, smooth etc, need to be addressed. The approach will be informal.

The presentation in these lectures will assume only a knowledge of calculus and familiarity with basic differential equations and physics. A framework to support a useful if less formal definition of a Lie group will develop gradually below. A Lie group is a type of group with continuously many elements. Most readers will know something about groups, but an outline of some basic ideas concerning finite groups is provided here for completeness and to fix some vocabulary or notational conventions that might be idiosyncratic. The reader who is already familiar with groups can skip these comments and refer back to them if necessary. In that case the reader should find the subsection introducing Lie groups and start there.

Acknowledgments

I am indebted first of all to my wife and colleague, Mizuho K Schwalm, without whose help this project would not have been possible. In addition to proofreading and suggesting improvements in the text, she has worked through and corrected as needed most of the calculations appearing in the book. She is responsible for video recording, editing, and production. Where the pronouns we or us appear without another obvious antecedent, they refer to her and me.

We are grateful to our teachers. The subject of one-parameter groups applied to differential equations has been of constant interest to us since 1970 when we were enrolled in an undergraduate math course on Lie groups given by Professor Philip Gold of the Mathematics Department at Portland State University. As impetus for the current volume, we owe a considerable debt to Dr Gold and his engaging lecture style. How inspiring those lectures were for us! We should like to thank as well Dr William Kinnersly and Dr Mark Peterson (professor of physics and mathematics, currently at Mount Holyoke College) and especially Dr John Hermanson of the Physics Department at Montana State University for courses on groups and other theoretical methods. Particularly, we are grateful to our dissertation advisor Dr Hermanson for many things, including a summer course in group representations and quantum mechanics that has turned out to be quite useful over the years.

Thanks to my students at the University of North Dakota and elsewhere for helping me work through the lectures and making occasional suggestions.

Finally, we must thank the University of North Dakota for use of the lecture room and the chalk boards, and for providing a stimulating work environment.

Author biography

William A Schwalm

Dr William A Schwalm received a Bachelor of Science degree in Physics from The University of New Hampshire and PhD in condensed matter theory from Montana State University. He is currently Professor of Physics in the Department of Physics and Astrophysics at The University of North Dakota, Grand Forks, where he has been employed since moving to North Dakota from Salt Lake City in 1980, where he held a postdoctoral appointment in the Department of Physics at The University of Utah. His research has specialized in mathematical analysis of physical problems. Most recent publications pertain to groups and dynamical systems and include another volume in the IOP Concise Physics series: *Lectures on Selected Topics in Mathematical Physics: Elliptic Functions and Elliptic Integrals*. He has held visiting professorships at Montana State University, The University of Minnesota, the University of Rome (Sapienza – Università di Roma) and Kyoto University.

Introduction

This volume is to accompany a set of video lectures. It is written for physics students enrolled in a first year graduate course in methods of theoretical physics, or for a physicist who in the course of their work has found it necessary to know some basic things about Lie theory. In particular, it is a summary overview of the theory of finite groups, a brief description of a manifold, and then an informal development of the theory of one-parameter Lie groups, especially as they apply to ordinary differential equations. The treatment is informal, although systematic and reasonably self-contained, as it assumes a familiarity with basic physics and applied calculus, but does not assume additional mathematical training. The presentation is concise, although there are some points on which it is more expansive. I have devoted some narrative to the points where I think the standard resources do not provide enough or may even fail to make some key observations.

The instructional intent of the book is that a motivated reader, who knows calculus but who has never heard of a group, should have a fair chance of finding symmetries of a second order differential equation encountered in the wild—provided a symmetry exists—and having found one, should be able to use it to reduce the order of the differential equation. The phrase 'in the wild' is intended to mean that the equation is not a standard textbook example but one arising from some physical problem in the context of research. This objective may seem limited in scope, but I believe it is not so limited. It prompts one to learn a number of useful things about Lie groups along the way, such as what an infinitesimal generator is and how to find one, how to construct a prolongation, how to find and use canonical coordinates, how to find symmetries, integrating factors, and so on.

There are places in the presentation where I have chosen not to move the development along on the path of least resistance by considering only the most familiar cases or a couple of deceptively simple examples, because I want the reader to encounter the general methods. For instance, at least some space is devoted to groups *not* expressed in a way that the parameters simply add under the group operation. The objective is that an interested reader should acquire a tool that is complete and that actually works to simplify or solve differential equations. At the same time, in the process, such a reader should learn enough of the practical apparatus of Lie theory to move on into other topics.

Chapter 1 is about symmetries, finite groups and representations. It is brief but contains the ideas built upon in later sections. It contains more as well. Unitary matrix representations, orthogonality, and the use of characters are outlined, primarily for applications somewhat outside the ones actually covered in this volume. It is anticipated that it might be used to support material presented in other lectures later on. The material that is built upon in later sections of this volume where Lie groups are developed is really from the first two sections of chapter 1. Readers who already have some knowledge of symmetries, group axioms, and subgroups are advised to move on to chapter 2 and refer back if needed.

As applied to differential equations, the Lie groups will occur as coordinate substitutions or coordinate transformations. They are point transformations in that they are defined completely by the group action on the dependent and independent variables. Substitution groups and infinitesimal generators, which are central to the theory, are developed in chapter 2, as is another essential concept, that of canonical coordinates of which the group transformations move only one, leaving the others fixed. The objective of applying the groups to ordinary differential equations is addressed in some detail in chapter 3. The culmination of this brief treatment of symmetry reduction is the penultimate section 3.7, which is focused on finding symmetries of second order equations and on using a symmetry to reduce the order. It follows to some extent the treatment in Hans Stephani's chapter 4 on how to find Lie point symmetries of an ordinary differential equation [12], however the presentation here is more conversational, and I have deliberately chosen to work a more complicated example, for the reasons mentioned above.

The final section is more directly related to a physical application. Nöther's theorem, namely that to each continuous symmetry of a mechanical system there corresponds a conserved quantity, is introduced and proven in the manner of Desloge and Karch in an article in the *American Journal of Physics* [16] where I first saw it. Some exercises are suggested, which include scaling symmetry of the inverse-cube law force and a case of helical symmetry.

Bibliographic notes

Background material on groups in a general physics context can be found in the popular books by Hamermesh [1] and by Tinkham [2]. Application of symmetry groups to DEs is described originally by Sophus Lie, for instance in German in [3, 4]. The content of these original works is summarized in two very readable texts, [5] and [6], both writers associated with Johns Hopkins University in a time not far removed from Lie's publications. My introduction to this subject was through a course on continuous groups and Lie theory applied to DEs by Professor Philip Gold at Portland State University, spring semester, 1970 [7]. The course was based on the Cohen's very accessible undergraduate text [6], amplified in class notes. I regard this book as a good reference for physics or engineering students. It may be read on-line via the Cornell University Library Historical Math Monographs [8]. More recently there are several important, rather complete reference for Lie theory of which I will mention, [9–11] in particular. The book by Olver [11] contains a very complete and useful bibliography. In addition to Cohen's book mentioned already, these lectures derive in some places from concise but lucid and useful treatment in a book by Hans Stephani [12], especially in his presentation of how to find symmetries of second order equations. Another similarly useful book is one by L Dresner [13]. General material on Lie groups in physics, especially semisimple groups and their algebras, is found in [14] and a descriptive outline is given in [15]. I first saw the treatment given here of Nöther's theorem in an article by Desloge and Karch [16].

References

[1] Hamermesh M 1962 *Group Theory and its Application to Physical Problems* (Reading, MA: Addison-Wesley)

[2] Tinkham M 1964 *Group Theory and Quantum Mechanics* (New York: McGraw-Hill)

[3] Lie S 1888 Klassifikation und Integration von gewönlichen Differentialgleichungen zwischen x, y die eine Gruppe von Transformationen gestatten *Math. Ann.* **32** 213–81

[4] Lie S 1891 *Vorlesungen über Differentialgleichunen mit bekannten infinitesimaln Transformationen* (Leipzig: Taubner)

[5] Page J M 1897 *Ordinary Differential Equations, with an Introduction to Lie's Theory of the Group of One Parameter* (London: MacMillan)

[6] Cohen A 1911 *An Introduction to the Lie Theory of One-parameter Groups with Applications to the Solution of Differential Equations* (New York: Heath)

[7] Gold P 1970 Continuous groups applied to differential equations *Class Notes, Mathematics, Portland State University, Fall quarter* (Unpublished)

[8] Cornell University Library 2016 Historical Math Monographs http://ebooks.library.cornell.edu/m/math/ (online repository that contains [4] and [5] in electronic form)

[9] Ovsiannikov L V 1982 *Group Analysis of Differential Equations* (New York: Academic)

[10] Ibragimov N H 1985 *Transformation Groups Applied to Mathematical Physics* (Boston, MA: Reidel)

[11] Olver P J 1986 *Application of Lie Groups to Differential Equations* (New York: Springer)
[12] Stephani H 1989 *Differential Equations: Their Solution Using Symmetry* (New York Cambridge University Press)
[13] Dresner L 1999 *Applications of Lie's Theory of Ordinary and Partial Differential Equations* (Bristol: IOP Publishing)
[14] Weybourne G 1974 *Classical Groups for Physicists* (New York: Wiley)
[15] Lipkin H J 2002 *Lie Groups for Pedestrians* (Mineola NY: Dover)
[16] Desloge E A and Karch R I 1977 Noether's theorem in classical mechanics *Am. J. Phys.* **45** 336–40

IOP Concise Physics

Lectures on Selected Topics in Mathematical Physics: Introduction to Lie Theory with Applications

William A Schwalm

Chapter 1

Groups

A group $(G, \cdot)$ is a mathematical system with a set G and one operation $\cdot$ 'dot'. The operation is a mapping from pairs of elements in G back into G. The group is closed under the operation, so that if a and b are in G, then one writes, $a \cdot b = c$, where c is some other element in G. Sometimes the operation is called a product or multiplication, and ones says 'a times b', even though the dot need not be commutative. The dot may be omitted so that as in ordinary multiplication the group operation is represented by juxtaposition, $a\,b = c$. The operation in a group has other properties as well. To wit, a group consists of a set G of elements and an operation $\cdot$ defined on them such that:

- (closure) if $a \in G$ and $b \in G$ then $a \cdot b \in G$ and also $b \cdot a \in G$;
- (associativity) if a, b and c belong to G, then $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
- (existence of identity) there is a particular element e in G such that, for each $a \in G$, $e \cdot a = a$ and $a \cdot e = a$;
- (existence of inverse elements) for each $a \in G$ there is a unique element in G denoted by a^{-1} which has the property $a^{-1} \cdot a = e$ and also $a \cdot a^{-1} = e$.

In general, a group need not be commutative or 'abelian'. An abelian group is one for which, in addition to these four properties, has the property:

- ⋄ (commutativity) for each a and b in G, $a \cdot b = b \cdot a$.

A group without this extra property may sometimes be called non-abelian just for emphasis.

1.1 Permutations and symmetries

Groups can be either finite or infinite, meaning there is either a finite or infinite number of elements in G. The order $|G|$ of a finite group $(G, \cdot)$ is the number of elements in G. For finite groups, one can write a *Cayley table* to define the operation.

doi:10.1088/978-1-6817-4449-0ch1

© Morgan & Claypool Publishers 2017

This is a group multiplication table, except that the redundant column and row for 'e times' and 'times e' are usually deleted to save space.

For example, consider the group of symmetry operations of an equilateral triangle with vertices numbered $\{a, b, c\}$. So the symbol $\{a, b, c\}$ in this context means the ordered set of vertices. These vertices are originally in positions labeled 1, 2, 3. A *symmetry operation* on the triangle is one that preserves its general shape, so that if one were to rotate it by 120° for example, it would look exactly the same as before. Two symmetry operations, a rotation R and a flip A, are illustrated in figure 1.1. We can think of the original positions as boxes. Rotating or flipping the triangle is really a matter of moving the vertices around to different boxes. Denote a cyclic permutation operator, or an *n-cycle*, by an n-tuple of integers in parentheses. For instance, the three-cycle (i, j, k) acting on the set of vertices of the triangle moves the vertex currently in box i to box j, moves the vertex currently in j to k, and moves the vertex currently in k to i, in cyclic fashion. When applying the operation, the boxes remain fixed and the corners a, b, and c move around amongst the boxes. Then, for example, the rotation could be represented as an application of the three-cycle $R = (1, 2, 3)$ by the equation

$$R\ \{a, b, c\} = (1, 2, 3)\ \{a, b, c\} = \{c, a, b\}.$$

So R is a rotation operator. The usual convention is that if the symmetry operator does not move a vertex, the vertex is not listed. The transposition (i, j) just interchanges vertices i and j leaving the other vertex fixed. Then for example, the reflection symmetry of the triangle that exchanges vertices 1 and 2 leaving 3 fixed is $(1, 2)$. From now on I will use juxtaposition, as one does for ordinary multiplication, to signify the group operation rather than the more awkward dot.

$$A\ R\ \{a, b, c\} = (1, 2)\ (1, 2, 3)\ \{a, b, c\} = (1)(2, 3)\ \{a, b, c\} = \{a, c, b\}.$$

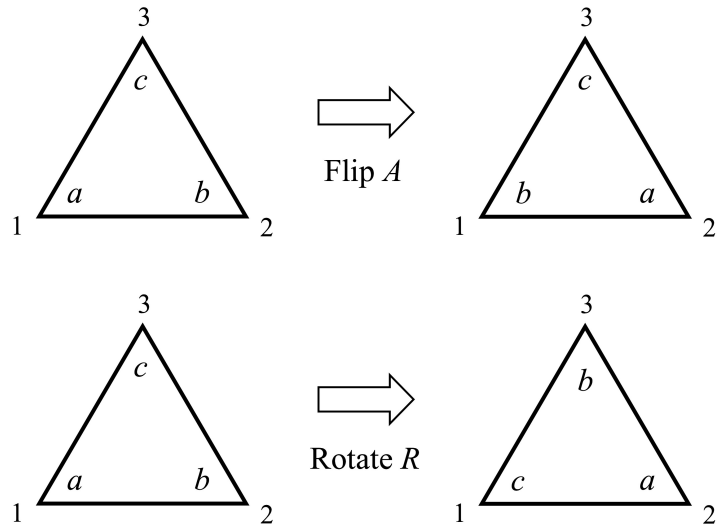


Figure 1.1. Two illustrations of cycle operators from the symmetry group of a triangle. A and R denote a flip (or reflection) and a rotation, respectively.

As was mentioned already, the group operation is sometimes referred to as a multiplication, even though it may not be commutative. So now this multiplication can be performed with no reference to pictures, or multiplication tables, since one can just perform successive substitutions. It is easy to multiply cycles as operators. The cycles give a sort of concrete representation of the operation, in that sense. For instance, bearing in mind that the cycle to the right operates first,

$$(1, 5, 3, 2)\,(1, 6, 2, 4, 5) = (1, 6)\,(2, 4, 3)\,(5).$$

Thus it is always possible to reduce a product to a sort of *standard form*. By a standard form (sometimes called a normal form) of some expression or quantity I will mean a form in which two such quantities are equal if and only if their standard forms are identical. This is important when you want to know whether the two things are equal or not. So in this case the standard form is the product of disjoint cycles, or cycles such that no index value appears in more than one of them, each cycle in lexicographical order, i.e. starting with the lowest index. The result in this case is (1, 6) (2, 4, 3) (5). The final factor (5) is dropped because it is really just the identity, meaning 'take 5 into 5 holding everything else fixed'. Thus (5) really keeps everything fixed, so (5) or (2) or any other singleton acts as the identity 1. It should be clear that one can permute entries in a given cycle cyclically, or can multiply the disjoint cycles together in any order, and still have a representation of the same group element. The order of multiplication of the cycles would be important if they were not disjoint.

The Cayley table for the triangle symmetry group with (1) representing the identity is

(1)	(1, 2, 3)	(3, 2, 1)	(1, 2)	(2, 3)	(3, 1)
(1, 2, 3)	(3, 2, 1)	(1)	(1, 3)	(1, 2)	(2, 3)
(3, 2, 1)	(1)	(1, 2, 3)	(2, 3)	(3, 1)	(1, 2)
(1, 2)	(2, 3)	(3, 1)	(1)	(1, 2, 3)	(3, 2, 1)
(2, 3)	(1, 3)	(1, 2)	(3, 2, 1)	(1)	(1, 2, 3)
(3, 1)	(1, 2)	(2, 3)	(1, 2, 3)	(3, 2, 1)	(1).

Since the Cayley table is not symmetric about its diagonal, the group of symmetries of an equilateral triangle is not abelian.

One can see an interesting thing from the cycle structure: elements of the group divide into different types according to whether the cycles look like $(\cdot)$, $(\cdot\ \cdot)$ or $(\cdot\ \cdot\ \cdot)$. These correspond to the identity, the flips (two-cycles), and the rotations (three-cycles). So the cycle structure partitions the group into elements of different types. These are different *classes* that have different cycle structure and represent different kinds of operations. More will be said about classes in the next section. Different cycle structures always correspond to different kinds of operations, but as we shall see, different kinds of operations can sometimes have similar looking cycle structures.

Sometimes it is useful to speak of a semigroup, which is closed and associative, but may or may not contain an identity. If it does, then any particular element in the semigroup may or may not have an inverse.

Finite groups arise in physics in connection with vibrations, quantum mechanics of a few particles, etc. Thus there are certain special objectives so that within physics one often visits topics in roughly this order: subgroups, homomorphisms, factor groups, representations, orthogonality, character tables, symmetry of quantum states, selection rules, etc. We will encounter most of these topics. However, emphasis in these lectures will be on infinite rather than finite groups. In classical mechanics (outside of the theory of small oscillations), continuum mechanics, particle physics and in field theory, one often is interested in *continuous groups*, or in particular *Lie groups*, which come up naturally in connection with differential equations (DEs).

1.2 Subgroups and classes

The subgroup concept is important in the case of either finite or infinite groups. Perhaps one could guess what a subgroup is from the name. In a group $(G, \cdot)$ if there is a subset $H \subseteq G$ such that $(H, \cdot)$ forms a group under the same operation, then $(H, \cdot)$ is a subgroup of $(G, \cdot)$. Instead of using the full notation, one usually denotes the groups simply by naming the sets of elements as G and H. Then H is a subgroup of G, written $H \leqslant G$. If in addition $H \neq G$, and H contains more than just the identity, then H is said to be a proper subgroup of G, written $H < G$.

Here is a result useful for determining whether or not a subset H of a group G is actually a subgroup of G. I will put it in the form of a small theorem. There will not usually be proofs in the material presented here. These lectures are not on mathematics *per se*, and so informal proofs will be given only when they are very easy or very instructive. In this case, the proof is easy and the result is quite useful, especially for finite groups.

Theorem. If $(G, \cdot)$ is a group and H is a subset of G, then $(H, \cdot)$ is a subgroup if and only if whenever a and b belong to H, then $a \cdot b^{-1}$ also belongs to H.

The proof is as follows. The operation of $(G, \cdot)$ is associative with respect to all elements of G, and so it is still associative when restricted to elements of the subset H. Of course, the open question at this point is whether $(H, \cdot)$ is closed with respect to the dot, or whether the product of two elements in H might land outside H. Putting this aside for the moment, suppose a is any element of H. One has by hypothesis that $a \cdot a^{-1} \in H$ so that the identity e is in H. Then for any a in $H, e \cdot a^{-1} = a^{-1}$ is in H so H contains all the inverses. Finally, for any two elements a and b in H, the fact that b^{-1} is in H gives that $a \cdot (b^{-1})^{-1} = a \cdot b$ is in H so H is closed with respect to the dot. This completes the proof.

Group generators will play a role in the development of continuous groups. The concept is demonstrated most clearly in a finite group. Consider any particular element g of a finite group G, and form the set of all powers of g, namely $\{g, g^2, g^3, \ldots\}$ where $g^n = g \cdot g^{n-1}$. It follows by induction, using associativity, that $g^m g^n = g^{m+n}$. Multiplying different powers together corresponds to exponent manipulations, just as in ordinary multiplication. Since the group is finite there must be finitely many distinct powers. So, in the sequence of powers it must be that

eventually $g^n = g^k$ where $0 \leqslant k < n < |G|$. Assume this is the lowest such n. It is quite easy to show that for any power k, $(g^k)^{-1} = (g^{-1})^k$. Call this g^{-k}, and so one has

$$g^k = g^n, \quad g^{-k} \cdot g^k = g^{-k} \cdot g^n, \quad e = \text{the group identity} = g^{n-k}.$$

The set $H = \{e, g, g^2, g^3, \ldots, g^{n-k-1}\}$ comprises all distinct powers of g. If $a \in H$ then $a = g^p$ for some power $0 < p < n - k$ and $a^{-1} = g^{n-k-p}$ is also in H, I find by computation that whenever a and b are both in H, then $a \cdot b^{-1}$ is also in H so that $(H, \cdot)$ is a subgroup of $(G, \cdot)$. Also, since the computations are performed using exponent manipulation, where multiplication of group elements corresponds to addition of integers, it is clear that the subgroup H is actually an abelian (a commutative) subgroup of G. Commutativity comes in this case from the associativity. For instance,

$$g^6\, g^4 = (g^4\, g^2)\, g^4 = g^4\, (g^2\, g^4) = g^4\, g^6.$$

The group $H = \{e, g, g^2, g^3, \ldots\}$ is the *cyclic subgroup* of G generated by the single element g. One says that g is a generator of H. A cyclic group is a group generated by a single element.

An extension of this idea is to construct a subgroup starting from a subset S of group elements in G. Form all possible, distinct powers and products of the elements of S, in any order. There can only be finitely many of these, since there can be no more than $|G|$. Denote this aggregation of group elements as $\langle S \rangle$. This subset $\langle S \rangle$ must be closed under the group operation, since all distinct products are included. The inverse of any element in $\langle S \rangle$ is in $\langle S \rangle$ because we have seen that in a finite group the inverse of any element g is always some power of g. Thus, if a and b are in $\langle S \rangle$, ab^{-1} is also in $\langle S \rangle$, and so it is a subgroup of G. It is the subgroup $\langle S \rangle$ of G generated by the subset S. Since elements of S need not commute in G, $\langle S \rangle$ need not be abelian.

One can refer to a generator or the set of generators of the group G itself. A *generator* of a group is a minimal set of group elements such that the whole group can be reconstructed by taking appropriate products of the generator elements and their inverses. Usually, the elements of the generator are also called *generators* of the group. So the term generator can mean—rather ambiguously—either a single generator element, or a complete set of these elements.

As an example, for the triangle group, (1, 2, 3) and (1, 2) can be taken as generators. Thus (1, 2) is 'a generator', while 'the generator', means a set such as $S = \{(1, 2, 3), (1, 2)\}$ which generates the whole group by taking products and reciprocals. You could write $G = \langle S \rangle$ or more explicitly, $G = \langle (1, 2, 3), (1, 2) \rangle$.

Classes. It was mentioned above that elements of a particular group can be partitioned into classes, the members of which are all operations of a certain type. To see what this means it is best to start with an abstract definition: two elements a and b are said to be *in the same class* if there exists an element g in the group such that

$$b = g^{-1} a\, g.$$

When elements are in the same class, they represent the 'same kind' of symmetry operation, such as both are rotations by $2\pi/3$ about different but similar axes, or both are reflections of a certain type, etc. This is because you can view the action of g as a sort of change of coordinates. When a is a rotation in the same class as b, for instance, it is because you can (1) use some g in the group to transform the axis of rotation of b into the axis of rotation of a, then (2) perform a, and then (3) use g^{-1} to transform back. The result of this three step process is the same as operating with b. Thus rotations in the same class must be rotations by the same amount such that there is an element g in the group that transforms the axis of rotation of one into that of the other.

When you express the group operations as products of disjoint cycles, you can see most of the class structure by seeing the 'shape' or dot diagram of each element. The diagram for

$$(1, 2, 4, 6)(3, 5, 9, 8, 7)$$

would be

$$(\cdot\ \cdot\ \cdot\ \cdot)\ (\cdot\ \cdot\ \cdot\ \cdot\ \cdot)$$

for instance. There are three classes in the group of the triangle, and also three dot diagrams.

- the identity (always in a class by itself): $(\cdot)$
- the three fold rotations (two of them): $(\cdot\ \cdot\ \cdot)$
- the flips (three of them): $(\cdot\ \cdot)$.

It turns out that elements belonging to the same class must belong to the same diagram, but in general elements belonging to the same diagram can belong to different classes. The idea of the dot diagram is convenient because, when expressed as disjoint cycles, it is clear at a glance whether or not two elements have the same diagram and thus might be in the same class.

For example, in the symmetry group of the cube, if I number the corner positions in a certain way it happens that (1, 3)(2, 4)(5, 7)(6, 8) is a two-fold rotation around a four-fold axis of the cube, while (1, 5)(2, 8)(3, 7)(4, 6) is a two-fold rotation around an axis through the center of two opposing edges. They are in different classes, but have the same dot diagram,

$$(\cdot\ \cdot)(\cdot\ \cdot)(\cdot\ \cdot)(\cdot\ \cdot).$$

On the other hand, the three-fold rotations about an axis passing through diagonally opposite cube corners belong to

$$(\cdot\ \cdot\ \cdot)(\cdot\ \cdot\ \cdot).$$

It is clear at a glance that elements belonging to these two different dot diagrams are in different classes. Indeed here we see that one class comprises two-fold and the other three-fold rotations.

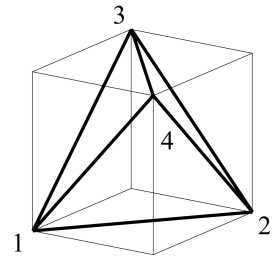


Figure 1.2. Tetrahedron vertex numbering. Notice how the tetrahedron fits within a cube.

Exercise. It is instructive to work out a minimal set of generators for the symmetry group of a regular tetrahedron using cycle notation. Work out the classes of the group and the physical meaning of the operations in each class. Use the figure provided (figure 1.2), which shows a tetrahedron embedded in a cube. Actually, two distinct cases arise depending on whether you do or do not consider reflections to be a symmetry of the figure. If you do include reflections, the group becomes S_4, the full symmetric group of all permutations of four distinct objects. In the full symmetric group S_n of n objects, each class of operation belongs uniquely to a particular dot diagram. So, what happens when you do not include any reflections in the symmetry group?

1.3 Representations

Generally, a group representation is a set of concrete objects, such as matrices, permutation cycles, or (in the case of Lie groups) transformation operators with a natural composition that reflects the group multiplication table. To make this more definite, we should define a *homomorphism*. Consider first a homomorphism between two finite groups.

A homomorphism is a mapping f from a finite group $(G, \star)$ into another finite group $(K, \circ)$ that preserves the group operation. In other words, such that if

$$g_1 \star g_2 = g_3$$

in $(G, \star)$, then f is a homomorphism if

$$f(g_1) \circ f(g_2) = f(g_3)$$

in $(K, \circ)$.

This does not require that f be either one-to-one or onto. The *image* $f(G)$ of G under f is the set of all elements of K that are equal to $f(g)$ for some $g \in G$. The image $f(G)$ of G under any homomorphism f into K must be a *subgroup* of K.

In the case that a homomorphism f is both one-to-one and onto (so that the inverse f^{-1} exists), then we call f an *isomorphism*. Two groups $(G, \star)$ and $(K, \circ)$ are called isomorphic if there is an isomorphism between them. When two groups are isomorphic, it means they are really the same group, and the elements simply have different names in the two sets G and K.

General homomorphisms are much more interesting because of the fact that information is lost in a controlled way in a homomorphism. Two or more different elements in G can go onto the same element in K. What is interesting is that, essentially, there are only a few different ways this can happen.

With this in mind, we can define a representation as follows. A *representation* Γ of group G is a homomorphism from G onto a group of $n \times n$ unitary matrices. Thus if in the group G we have

$$g_1 g_2 = g_3,$$

then in the representation Γ we have the matrix product relation,

$$\Gamma(g_1)\Gamma(g_2) = \Gamma(g_3).$$

The size n of the matrices is the *dimension* of the representation Γ. If the homomorphism Γ is an isomorphism, then it is a *faithful* representation. For example, take the symmetry group of the triangle listed above. A possible representation using the usual rotation matrices is

$$\Gamma(1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \Gamma((1,2,3)) = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and

$$\Gamma((1,2)) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is all you need to assign. The identity matrix always represents the group identity, and the triangle group is *generated* by the two elements $(1, 2, 3)$ and $(1, 2)$. This means that all the other elements are products of some combination of these two and their inverses. Thus the representation matrices for the other elements can be found by taking the corresponding matrix products.

Notice that the representation generated by the two 3×3 matrices above is faithful, since it is one-to-one. That means the matrix group is isomorphic to the original abstract group of symmetries of the triangle.

It is also easy to find a non-faithful representation. Notice that if you take the determinant of the matrices, the multiplication is always preserved, so

$$\det \Gamma(g_1)\ \det \Gamma(g_2) = \det \Gamma(g_3),$$

so if Γ is any representation, then $\det \Gamma$ is another representation (one-dimensional). Usually $\det \Gamma$ is not faithful, since usually different matrices give the same determinant. In this case, all the group operations are either rotations, including the identity, or reflections. The determinant is $+1$ for the identity and the rotations and -1 for each of the three reflections. Another one-dimensional representation exists

for all groups. This trivial representation with each element represented by the number 1 is the *identity representation*.

Two different representations Γ_1 and Γ_2 of the same group are *equivalent* if there exists one particular unitary matrix M such that, for each $g \in G$,

$$\Gamma_2(g) = M^{-1}\,\Gamma_1(g)\,M.$$

A representation Γ is said to be *reducible* if it is equivalent to a block-diagonal representation. That is to say, if it is equivalent to a representation where the matrix for *every* group element is in block-diagonal form. An example of a matrix in *block-diagonal form* is

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & 0 & 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

In this particular matrix you have a 4×4 and a 2×2 block on the diagonal. A more general example might have many different sized blocks. Keep two things in mind. One is that a reducible representation is only *equivalent* to a block-diagonal representation, it need not appear block-diagonal. The other is that the same matrix M must block diagonalize *each* matrix of the representation, i.e. each $\Gamma(g)$ for $g \in G$.

Even though reducible representations need not look block-diagonal, I will usually think of equivalent representations as the same representation. (They are obviously isomorphic as groups.)

A representation of a group G is *irreducible* if it is not reducible. If one counts equivalent representations as the same representation, then it turns out that *there is only a finite number of different, irreducible unitary representations for any finite group G.*

Thus for a given group, we can tabulate (list) the irreducible representations once and for all! Also, given any reducible representation one can in principle find a fixed matrix M that 'reduces it' to block-form in which each block corresponds to one of the possible irreducible representations.

So, suppose Γ is a reducible representation, and suppose there exists a fixed matrix M such that for each $g \in G$,

$$M^{-1}\Gamma(g)\,M = \begin{pmatrix} \Gamma^{(3)}(g) & 0 & 0 \\ 0 & \Gamma^{(5)}(g) & 0 \\ 0 & 0 & \Gamma^{(3)}(g) \end{pmatrix},$$

where $\Gamma^{(3)}$ and $\Gamma^{(5)}$ are two of the irreducible representations of G. Then we say, 'Γ contains two of $\Gamma^{(3)}$ and one of $\Gamma^{(5)}$', and we write this as

$$\Gamma = 2\Gamma^{(3)} \oplus \Gamma^{(5)}.$$

Notice, for example, that the two generator matrices shown above for a representation of the triangle group happen to be block diagonal with one two-dimensional and one one-dimensional block each. This implies that the inverses and products of these two matrices will all have the same property. The one-dimensional representation must be irreducible. (How could it reduce?) The two-dimensional block also turns out to be an irreducible representation of the triangle group. If I label these irreducible representations as $\Gamma^{(2)}$ and $\Gamma^{(3)}$, then in that example,

$$\Gamma = \Gamma^{(2)} \oplus \Gamma^{(3)}.$$

The great value of this idea of reducing representations is that one can study the much simpler irreducible representations rather than the original ones. They are simpler because (a) they are smaller matrices, and (b) they are always the same for a given group, so they can be dealt with ahead of time, tabulated once and for all. You may wonder how to find M. In most cases, one does not need to find M, as we shall see.

1.4 Orthogonality

For representations of finite groups, the orthogonality theory is the main tool used in calculations. The orthogonality theorem for group representations is stated and proved in standard texts, particularly there are presentations addressed to physicists in Hamermesh [1] and Tinkham [2]. Consider a group G with h elements in it. Then h is the *order* of G, or $h = |G|$. Let the irreducible representations of G be $\Gamma^{(1)}$ with dimensionality n_1 (meaning this representation uses $n_1 \times n_1$ matrices), $\Gamma^{(2)}$ of dimensionality n_2, etc. Then let $\Gamma^{(i)}_{\alpha\beta}(g)$ denote the α, βth entry in the matrix of representation i for the group element g. Then the orthogonality theorem says

$$\sum_{g\in G} \left(\Gamma^{(i)}_{\alpha\beta}(g)\right)^* \Gamma^{(j)}_{\gamma\delta}(g) = \frac{h}{n_i}\, \delta_{ij}\, \delta_{\alpha\gamma}\, \delta_{\beta\gamma}.$$

You can think of the orthogonality theorem as a tremendously powerful 'dot product'. The subscript being summed over is g, indicating the group elements. The other suffixes, i, α, β, and j, γ, δ, index the two vectors. There is one vector for each choice of indices. These vectors are orthogonal and normalized in a certain way, as you see. It is a powerful result.

How many vectors like

$$\Gamma^{(i)}_{\alpha\beta}$$

are there in this abstract space? For each value of i, there are n_i^2 values of the pair α, β. So there are n_i^2 vectors for representation i. The total number of representation vectors is

$$\#\text{ of vectors} = n_1^2 + n_2^2 + n_3^2 + \cdots = \sum_i n_i^2.$$

This is the dimensionality of the representation space. But the dimensionality of the representation space must also equal $h = |G|$, because each vector has h components as g sums over the elements of the group. Thus

$$\sum_i n_i^2 = h = |G|.$$

That is: *the sum of squares of the dimensions of the irreducible representations equals the order of the group.* This is a useful result. Never underestimate a theorem that counts something.

For working with finite groups, the representation matrices themselves are not the most convenient tool for reducing representations. Rather one would like to use a character table. The *character* $\chi^{(i)}(g)$ of representation i at group element g is just the trace of the representation matrix

$$\chi^{(i)}(g) = \text{tr.}\Gamma^{(i)}(g) = \sum_\alpha \Gamma^{(i)}_{\alpha\alpha}(g).$$

Thus, in the orthogonality theorem, set $\beta = \alpha$ and set $\delta = \gamma$. Then notice $(\delta_{\alpha\gamma})^2 = \delta_{\alpha\gamma}$, which is true of any function that has only the values 0 and 1. The result is

$$\sum_{g\in G} \chi^{(i)*}(g)\,\chi^{(j)}(g) = h\,\delta_{ij}.$$

A point to stress is that, in reality, the character is a class function, not a function in detail of the group element g. In other words, when g and g' are in the same class in G, which means

$$g' = f^{-1} g\, f$$

for some $f \in G$, then

$$\begin{aligned}\chi^{(i)}(g') &= \text{tr.}\Gamma^{(i)}(g') = \text{tr.}\Big((\Gamma^{(i)}(f))^{-1}\Gamma^{(i)}(g)\Gamma^{(i)}(f)\Big)\\ &= \text{tr.}\Big(\Gamma^{(i)}(f)(\Gamma^{(i)}(f))^{-1}\Gamma^{(i)}(g)\Big) = \text{tr.}\Gamma^{(i)}(g) = \chi^{(i)}(g).\end{aligned}$$

In commuting matrices in the middle step I am using cyclical invariance of the trace. Thus one can view the characters as functions of the class of g. Let $\chi_k^{(i)}$ stand for the common value of the character of representation i on class k. Then if there are p_k members of class k,

$$\sum_k p_k \chi_k^{(i)*} \chi_k^{(j)} = h\,\delta_{ij}.$$

Thus, think of

$$\sqrt{\frac{p_k}{h}}\,\chi_k^{(i)}$$

as the kth component of a complex unit vector $\hat{e}_i$. We have

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij}.$$

The number of components in each of these vectors is the number of classes. The total number of vectors is the number of representations. Since they are mutually orthogonal, the number of vectors cannot exceed the number of components (dimension). In fact, one finds that *the number M of classes equals the number of irreducible representations*.

Thus, consider the matrix formed with the unit vectors $\hat{e}_i$ as columns

$$E = \begin{pmatrix} \hat{e}_1 & \hat{e}_2 & \cdots & \hat{e}_M \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix}.$$

The character orthogonality can be written as

$$E^\dagger E = 1,$$

so the matrix E is unitary. Thus also

$$E\,E^\dagger = 1,$$

so

$$\sum_i \chi_k^{(i)*}\,\chi_m^{(i)} = \frac{h}{p_k}\,\delta_{km}.$$

In this second orthogonality relation, the sum is over representations.

Homework. Use orthogonality and other information to construct a character table for T, the symmetry group of the tetrahedron, not including reflections as symmetries.

References

[1] Hamermesh M 1962 *Group Theory and its Application to Physical Problems* (Reading, MA: Addison-Wesley)

[2] Tinkham M 1964 *Group Theory and Quantum Mechanics* (New York: McGraw-Hill)

IOP Concise Physics

Lectures on Selected Topics in Mathematical Physics: Introduction to Lie Theory with Applications

William A Schwalm

Chapter 2

Lie groups

The previous chapter addressed some properties of groups. Now, in contrast to the finite groups mostly dealt with so far, consider a group with infinitely many elements. A Lie group is a set G that is both a group and a smooth manifold.

2.1 Lie groups as manifolds

Roughly speaking, a manifold is a set of points to which coordinates can be assigned, where in this case the points are group elements. These two aspects of G come together in the requirement that the group multiplication and inversion have to be smooth maps. A Lie group is an infinite group where the group elements are indexed by one or more continuous parameters, α. So these parameters are coordinates on the group manifold. This gives some meaning to the idea of closeness of group elements because distance between points can be computed using the coordinates. The term continuous will apply to Lie groups in more than one way. For one thing, to say that an index or a group parameter is continuous is to say that it can take on a continuum of values. Another implication of continuity is that group elements g_α and g_β are close to a certain extent when the parameter sets α and β differ by only a small amount.

A Lie group can depend on several parameters, rather than just one. In such a group the index α becomes a vector in a parameter space then could be written as $\vec{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ for an r-parameter group. I shall not use this arrow notation. Then the distance between group elements is computed in some way from the manifold coordinates. The multi-parameter case leads to the idea of a *Lie algebra*, which will come up somewhat further on. First, in order to give meaning to calculus operations applied to groups, one would like at least a rough idea of the group parameter space and how it works. A longer, more complete narrative on groups as manifolds is to be found for instance in Olver [3].

doi:10.1088/978-1-6817-4449-0ch2 © Morgan & Claypool Publishers 2017

In more detail, then, an n-dimensional manifold is a set X of points and a collection of open subsets or neighborhoods such that each point p of X belongs to at least one open subset $U \subseteq X$ which is homeomorphic to an open subset $V \subseteq R^n$. Points in R^n are n-tuples of real numbers, $\{\alpha_1, \alpha_2, \ldots\}$ that serve as coordinates. In the case of the group manifold G, these coordinates are the group parameters. To say U is homeomorphic to V means there is mapping φ: $U \to V$ such that φ has an inverse φ^{-1}: $V \to U$ and that both φ and φ^{-1} are continuous.The homeomorphisms φ: $U \to V$ are *charts*. In the same way as I have been referring to a group $(G, \cdot)$ that comprises a set and an operation as just G—a nickname, as it were—I will usually refer to the manifold as just X.

For instance, consider the example illustrated in figure 2.1. Imagine reading a road atlas during a trip by car through the mid-western United States. The manifold X (meaning the space or point set of the manifold) is the world. In the world there are people, places, and things, but not coordinates. However, there is a travel atlas containing charts. Chart

$$\varphi_k\colon U_k \to V_k$$

shows Kansas and its vicinity, while chart

$$\varphi_n\colon U_n \to V_n$$

shows Nebraska and surrounding regions. These charts appear on different pages, but the regions covered overlap. The actual places for which they overlap are in $U_k \bigcap U_n$ on X. A town near the Kansas–Nebraska border will appear on at least two maps on different pages of the road atlas, once on a road map of Kansas, and once on a road map of Nebraska. The two different maps give location coordinates of this area either in Kansas coordinates on the Kansas page $\varphi_k(U_k \bigcap U_n)$ or in Nebraska coordinates on the Nebraska page $\varphi_n(U_k \bigcap U_n)$. Where the maps overlap, they should agree. If you look on the map of Kansas you may find a town with certain coordinates. The name of the town is a map to X. With the name you can turn to the Nebraska map, locate the town, and then find its Nebraska coordinates.

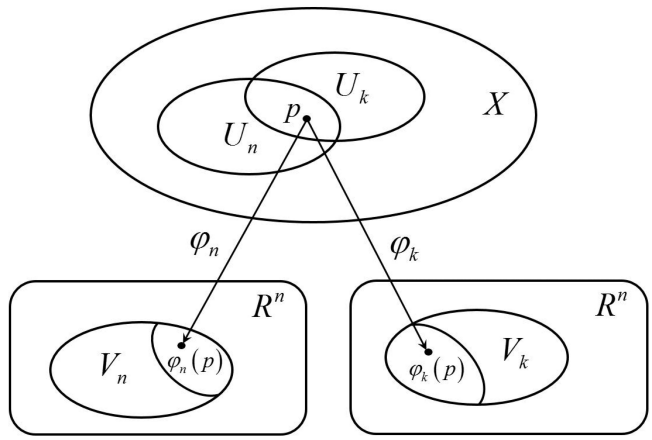


Figure 2.1. Point p belongs to the intersection of the two open sets U_k and U_n on manifold X. The maps φ_k and φ_n take p into coordinates represented in V_k and V_n, respectively.

A *smooth manifold* is a manifold X together with charts $\varphi_k\colon\ U_k \to V_k$ that cover all of X and are *smoothly related* wherever they overlap.

To say that the maps φ_k and φ_n are smoothly related is to say that the composite map $\varphi_{k,n} = \varphi_n \circ \varphi_k^{-1}$

$$\varphi_{k,n}\colon\ \varphi_k(U_k \bigcap U_n) \to\ \varphi_n(U_k \bigcap U_n)$$

is a smooth map from $\varphi_k(U_k \bigcap U_n)$ to $\varphi_n(U_k \bigcap U_n)$. I use the term smooth to mean that all the derivatives to all orders exist for all the functions such as $\varphi_n \circ \varphi_k^{-1}$ involving φ_n and φ_k. So to say the group G is a Lie group implies that, in addition to having algebraic structure of a group, G is a smooth manifold equipped with a complete set of charts, overlapping as necessary in this compatible way. This 'coordinatization' of G gives a practical meaning to the question of whether or not g_α is close to g_β. The answer can be found using the group-parameter coordinates α and β. This is represented in figure 2.2.

But why have multiple, overlapping charts? Why would a single set of coordinates not be enough? The answer is that it is often not possible to treat the group manifold as flat. The parameters may need to be cyclic, as in the case of rotation where $0 \leqslant \theta < 2\pi$ requires identifying $\theta = 2\pi$ with $\theta = 0$. So θ refers to a circle rather than a line segment. A two-dimensional manifold, for example, may be like a torus or a sphere. It is not possible to give a non-singular coordinate representation on a sphere using only one chart. The Mercator latitude–longitude coordinates are locally Euclidean but become singular at the north and south poles. So one needs other charts, at least very close to the poles.

The group operation and the inverse of $(G, \cdot)$ have to be defined by smooth maps also, if G is a Lie group. So for instance, for a particular one-parameter group parameterized in a certain way α might take on values of any real number in the open interval from $-\infty$ to ∞. Then, since clearly there can be no Cayley table, the result of 'multiplying' group elements together must be specified by a formula relating values of the index parameters. If I denote a group element as g_α, where α is

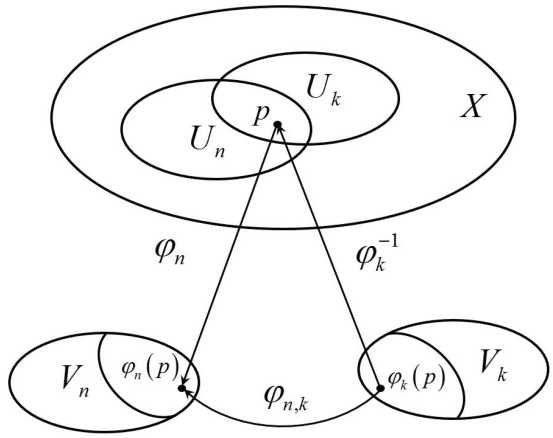


Figure 2.2. Compatibility. The coordinate transformation $\varphi_{n,k}$ and its inverse are smooth maps between coordinates in two different charts.

the parameter value (or set of parameter values in the multi-parameter case) and hence the α value specifies which group element, then there has to be a formula

$$\alpha_3 = \phi(\alpha_1, \alpha_2)$$

to define the product

$$g_{\alpha_3} = g_{\alpha_1} g_{\alpha_2} \equiv g_{\alpha_1} \cdot g_{\alpha_2}.$$

When we say a Lie group is a continuous group, it also implies that the product and the inverse are described by continuous functions. So the function ϕ is continuous in each of its arguments. Used in this way, the term continuous means roughly that when either of the arguments varies by a small amount, the function value changes by a small amount. There has to be a notion of closeness in the parameter space, so that 'small amount', is defined. In the case of a one-parameter group, the product is continuous if for general α and β such that $g_\alpha g_\beta = g_\gamma$, when one makes small variations $\delta\alpha$ and $\delta\beta$ in the parameters,

$$\phi(\alpha + \delta\alpha, \beta + \delta\beta) = \gamma + \delta\gamma,$$

then for any small positive ϵ, there is a δ such that $|\delta\gamma| < \epsilon$ whenever $\sqrt{\delta\alpha^2 + \delta\beta^2} < \delta$. This is just the usual meaning of a continuity when the term is applied to a function.

Similarly, the parameter value α' of the inverse $(g_\alpha)^{-1}$ of a group element must also be connected to α by a continuous function

$$\alpha' = \psi(\alpha).$$

Moreover, the functions $\phi(\alpha, \beta)$ and $\psi(\alpha)$ are assumed to have derivatives of all orders near points with general values of α and β, and thus one can write a Taylor series for these functions. Furthermore, for the groups of interest these series give convergent representations of the functions ϕ and ψ themselves in some neighborhood of a general point in parameter space. For a Lie group, the product and the inverse are described by (real) analytic functions. They have convergent power series representations. This assertion will have strong consequences, and it is an essential point. In keeping with this informal presentation I will make no attempt to justify the assertion, but as a result of it the transformation formulas and other expressions that evolve in specific examples treated in sections below will often be assumed to have series expansions. The exceptional cases, and the qualification of a 'general point' in the parameter space, will be taken up where they arise.

The requirements of closure and associativity put restrictions on the multiplication function ϕ. Therefore, as we shall see, a one-parameter Lie group is always abelian and so in fact one can usually parameterize it in such a way that

$$g_{\alpha_1} g_{\alpha_2} = g_{\alpha_1 + \alpha_2}.$$

When parameterized in this way,

$$g_0 = e,$$

so that $\alpha = 0$ is the identity. The identity is at the origin in parameter space, $\alpha = 0$.

2.2 Lie groups as groups of transformations or substitutions

Lie groups of interest to us will often be groups of transformations. This is analogous to what has been seen already in the example of the finite group of symmetry transformations of a triangle. It means that in addition to the group parameters such as α, β, γ, …, which are coordinates on the group manifold, the group elements act as transformations on a space which also has coordinates, such as x, y, z, …. The most natural way to represent the transformations might be as matrices (if they are linear) or perhaps more generally as groups of substitutions. Following Lie's development, we could think of substitution groups. A nice introduction to the subject is the straightforward one written for undergraduate students by Abraham Cohen [1], for which there is on-line access via the Cornell monograph collection [2]. We will view the group elements as transformations that act on functions of x, y, and z, in three dimensions, or in a general problem on functions of all the dynamical variables. In classical mechanics, for instance, the coordinates might be denoted as $q = \{q_1, q_2, \ldots, q_n\}$, although in these lectures I will usually refer to coordinates in the space where the transformations act as $x = \{x^1, x^2, x^3, \ldots, x^n\}$ with superscripts.

The group acts on functions of the coordinates through function composition, or in other words by a substitution recipe. Consider first a two-dimensional reference system in the plane with coordinates x and y. If the coordinates (x, y) of a point are represented in terms of another set of generalized coordinates by means of a substitution or change of variables, a function $F(x, y)$ of the coordinates becomes another function by composition.

Suppose for instance that (x, y) is replaced by $(\tilde{x}, \tilde{y})$ where $\tilde{x} = \tilde{x}(x, y)$, $\tilde{y} = \tilde{y}(x, y)$ then $F(x, y)$ is changed by function composition to $F(\tilde{x}(x, y), \tilde{y}(x, y)) = \tilde{F}(x, y)$. So the transformation is defined by substitutions. In a substitution group each such transformation is indexed by a certain value of the group parameter. If g_α takes (x, y) into $(\tilde{x}, \tilde{y})$ then the tilde symbols have two meanings. First, $\tilde{x}$ and $\tilde{y}$ are coordinates of the image point. So I could write

$$g_\alpha(x, y) = (\tilde{x}, \tilde{y}),$$

but the action of g_α is really defined by substitution formulas, one for each coordinate variable. In two dimensions,

$$g_\alpha\, x = \tilde{x}(x, y, \alpha),$$
$$g_\alpha\, y = \tilde{y}(x, y, \alpha),$$

since the coordinates of the image point must each be functions of the group parameter as well the original coordinates x and y. The tildes signify first that $(\tilde{x}, \tilde{y})$ is the image point, and then also that $\tilde{x}(x, y, \alpha)$ and $\tilde{y}(x, y, \alpha)$ are functions defining substitutions for x and y.

Several points come up. The group product $g_\alpha\, g_\beta$ is defined by composition of the transformations, or in fact successive substitutions, where g_β is assumed to act first:

$$(g_\alpha\, g_\beta)x = g_\alpha(g_\beta\, x) = \tilde{x}(\tilde{x}(x, y, \beta), \tilde{y}(x, y, \beta), \alpha),$$
$$(g_\alpha\, g_\beta)y = g_\alpha(g_\beta\, y) = \tilde{y}(\tilde{x}(x, y, \beta), \tilde{y}(x, y, \beta), \alpha).$$

There has to be a parameter value α_o corresponding to the identity. For a given g_α there has to be an α' corresponding to the inverse transformation, and most fundamentally the group operation needs to be closed and associative. These requirements can fail in various ways, although when the group is viewed as a group of transformations, and the operation is defined by successive substitutions, associativity follows somewhat naturally from associativity of function composition. Therefore, one will not usually need to worry about associativity for groups of substitutions or transformations.

It will be worthwhile to observe some things about one-parameter Lie groups by means of examples before moving on. Mostly these will be transformations acting in the plane. It will be instructive at first to see group parameterizations that are not additive, but in which the function ϕ defining the operation takes on a more general form. Consider first a very simple case, namely the group of dilations or contractions about the origin in the x–y plane:

$$g_\alpha\, x = \tilde{x}(x, y, \alpha) = \alpha\, x, \qquad g_\alpha\, y = \tilde{y}(x, y, \alpha) = \alpha\, y,$$

where $\alpha \geqslant 0$. In this case, the identity is where $\alpha = 1$. Closure would seem to follow from the fact that

$$g_\alpha\, g_\beta\, x = \alpha\, \beta\, x, \qquad g_\alpha\, g_\beta\, y = \alpha\, \beta\, y,$$

and the latter is a transformation of the same kind, governed by the same formula,

$$g_\alpha\, g_\beta\, x = \tilde{x}(\tilde{x}(x, y, \beta), \tilde{y}(x, y, \beta), \alpha) = \gamma\, x,$$
$$g_\alpha\, g_\beta\, y = \tilde{y}(\tilde{x}(x, y, \beta), \tilde{y}(x, y, \beta), \alpha) = \gamma\, y.$$

In this case, $g_\alpha\, g_\beta = g_\gamma$ where $\gamma = \phi(\alpha, \beta) = \alpha\, \beta$. The associative group property is seen to hold because ordinary multiplication is associative. One sees also, as predicted, that the group is abelian since ordinary multiplication is commutative. The inverse of g_α is $g_{1/\alpha}$ except in the case where $\alpha = 0$. The transformation is singular then, as it takes every point in the plane into the origin $(0, 0)$. So as a matter of fact, the parameter range has to be restricted to $0 < \alpha < \infty$. The identity element is g_1 or at the point $\alpha = 1$. So $\alpha_o = 1$ in this example.

It is easy to see how to make the parametrization additive in this simple example. Instead of α one can use the parameter $s = \ln \alpha$ and then $g_s\, g_t = g_u$ where $u = s + t$, the identity is at $s = 0$ and the parameter range is $-\infty < s < \infty$.

Another, more involved example is

$$g_\alpha\, x = \tilde{x}(x, y, \alpha) = x + \frac{1}{2}\alpha(x^2 + x - y),$$
$$g_\alpha\, y = \tilde{y}(x, y, \alpha) = y + \frac{1}{2}\alpha(2x - 1)(x^2 + x - y) + \frac{1}{4}\alpha^2(x^2 + x - y)^2.$$

This time the group elements are not linear transformations of the coordinates, so they are not equivalent to matrices acting on the coordinates (x, y). Notice that in this case the identity transformation corresponds to $\alpha = 0$. Some algebra shows that

$$g_\alpha\, g_\beta\, x = x + \frac{1}{2}(\alpha + \beta + \alpha\,\beta)(x^2 + x - y),$$

$$g_\alpha\, g_\beta\, y = y + \frac{1}{2}(\alpha + \beta + \alpha\,\beta)(2x - 1)(x^2 + x - y) + \frac{1}{4}(\alpha + \beta + \alpha\,\beta)^2(x^2 + x - y)^2.$$

So in this case $\phi(\alpha, \beta) = \alpha + \beta + \alpha\,\beta$. The symmetrical form of this expression shows that once again the group is abelian. To demonstrate the associativity,

$$\phi(\alpha, \phi(\beta, \gamma)) = \alpha + \beta + \gamma + \alpha\,\beta + \beta\,\gamma + \gamma\,\alpha + \alpha\,\beta\,\gamma = \phi(\phi(\alpha, \beta), \gamma).$$

One observes incidentally that the result is a combination of elementary symmetric functions of α, β, and γ. Setting $\phi(\alpha, \alpha') = 0$ one can find the inverse $g_{\alpha'}$ of g_α,

$$(g_\alpha)^{-1} = g_{\alpha'} \qquad \text{with} \quad \alpha' = -\frac{\alpha}{1 + \alpha}.$$

Once again it is not very difficult to change to an additive parametrization. Notice that $(1 + \alpha)(1 + \beta) = 1 + \alpha + \beta + \alpha\,\beta$ so

$$\ln(1 + \alpha) + \ln(1 + \beta) = \ln(1 + \alpha + \beta + \alpha\,\beta).$$

The substitution $s = \ln(1 + \alpha)$ will serve to introduce the new parameter. The parameter range is $-\infty < s < \infty$ and the parametrization is additive.

Another example is the family of linear transformations

$$g_\alpha\, x = \frac{1 - \alpha^2}{1 + \alpha^2}x - \frac{2\,\alpha}{1 + \alpha^2}y,$$

$$g_\alpha\, y = \frac{1 - \alpha^2}{1 + \alpha^2}y + \frac{2\,\alpha}{1 + \alpha^2}x.$$

The identity is $\alpha = 0$. Then if I apply two transformations, first g_β and then g_α,

$$g_\alpha\, g_\beta\, x = \frac{(\alpha\,\beta - \alpha - \beta - 1)(\alpha\,\beta + \alpha + \beta - 1)}{(\alpha^2 + 1)(\beta^2 + 1)}x + \frac{2\,(\alpha + \beta)(\alpha\beta - 1)}{(\alpha^2 + 1)(\beta^2 + 1)}y,$$

$$g_\alpha\, g_\beta\, y = \frac{(\alpha\,\beta - \alpha - \beta - 1)(\alpha\,\beta + \alpha + \beta - 1)}{(\alpha^2 + 1)(\beta^2 + 1)}y - \frac{2\,(\alpha + \beta)(\alpha\beta - 1)}{(\alpha^2 + 1)(\beta^2 + 1)}x.$$

it turns out to be a transformation of the same form indeed, where

$$\phi(\alpha, \beta) = \frac{\alpha + \beta}{1 - \alpha\,\beta}.$$

Regarding associativity one finds

$$\phi(\alpha, \phi(\beta, \gamma)) = \frac{\alpha + \beta + \gamma - \alpha\beta\gamma}{1 - \alpha\beta - \alpha\gamma - \beta\gamma} = \phi(\phi(\alpha, \beta), \gamma).$$

I notice once again that the middle expression is a combination of symmetric functions of α, β and γ. An alternative parametrization that is additive derives from the fact that

$$\tan(s + t) = \frac{\tan s + \tan t}{1 - \tan s \tan t}.$$

The additive parameter is s such that $\alpha = \tan s$ for $-\pi < s \leqslant \pi$.

The examples above have demonstrated the curious fact that in each of these cases the group has an additive parametrization such that $g_\alpha\, g_\beta = g_{\alpha+\beta}$ so that $\alpha = 0$ is the identity and the parameter range is either $-\infty < \alpha < \infty$ or $-\pi < \alpha \leqslant \pi$. The fact that this is true generally was mentioned above and it will be derived informally below. It means the group itself is isomorphic to (essentially the same as) either the group of real numbers under addition or the group of rotations in the plane. The latter is the same as the group of complex numbers on the unit circle $e^{i\theta}$ under multiplication. So the purely algebraic properties of a one-parameter Lie group are not complicated. The group is abelian. We shall see that, as in the case of the cyclic subgroups of a finite group, the uncomplicated, commutative structure comes from the fact that the group elements are all powers of a single generator element, in a certain sense. The more interesting structure in the examples presented, and in general for a one-parameter transformation group, has to do with the way the group transformations act on the x–y plane or wherever they act.

One more example is needed before moving on. It is important to realize that not every set of transformations of the form

$$x \;\rightarrow\; \tilde{x}(x, y, \alpha), \qquad y \;\rightarrow\; \tilde{y}(x, y, \alpha)$$

defines a one-parameter continuous group. For instance, consider

$$\tilde{x}(x, y, \alpha) = x + \alpha\, y, \qquad \tilde{y}(x, y, \alpha) = y + \alpha\, x^2.$$

If I compose two transformations of this family, the result is

$$\tilde{x}(\tilde{x}(x, y, \alpha), \tilde{y}(x, y, \alpha), \beta) = x + \alpha\beta x^2 + y(\alpha + \beta),$$

$$\tilde{y}(\tilde{x}(x, y, \alpha), \tilde{y}(x, y, \alpha), \beta) = y + x^2(\alpha + \beta) + 2\alpha\beta xy + \alpha^2\beta y^2.$$

This is not a transformation in the same family, because new powers and products of the coordinates show up. So the one-parameter family of transformations as I have defined it is not closed under composition. The closure requirement fails and so it is not a group.

2.3 Infinitesimal generators

An interesting fact concerning Lie groups is that the groups (or the subgroups connected continuously to the identity element in parameter space) can be reconstructed from only the linear terms in a Taylor series. To augment the description in this section, the early sections in the books by Cohen ([1], p 6) and Stephani [4] are highly recommended.

Now consider a one-parameter transformation group and a general parametrization with the composition rule $g_\alpha\, g_\beta = g_\gamma$ such that $\gamma = \phi(\alpha, \beta)$. In a Lie group I assume this function to be smooth so that it has derivatives of all orders. The basic notion of continuity is that a small change in either of the two factors will produce a small change in the product of two group elements. The smallness of the changes refers to the changes in the group parameters or, where the transformations act, to the smallness of changes in the image points.

Consider the action of the group on a particular point (x, y) in the plane. The resulting $(\tilde{x}(x, y, \alpha), \tilde{y}(x, y, \alpha))$ as a function of α is the parametric form of a curve passing through the point (x, y). For every point (x', y') on the curve there is an α value that takes the initial point (x, y) into (x', y'). There is also a parameter value β that would move (x', y') along the curve back to (x, y), namely the value β such that $g_\beta = (g_\alpha)^{-1}$. This curve is the *orbit* of (x, y) or of (x', y') or of any point on the curve. An orbit curve can be open or closed. An example of an orbit that is a closed curve would be the one consisting of points on a circle of fixed radius, the points of which are permuted by the action of the group of rotations of the plane about the center of the circle. Except for the end points, each pair of points on the same orbit can be transformed back and forth into one another by the action of the group. Two different orbits cannot intersect at finite values of the parameter, because if they did, the group transformation would fail to be one-to-one in some neighborhood of the cross point. So the x–y plane is partitioned into orbits, each point being on one and only one orbit, except possibly for singularities that might happen at the open ends where the group parameter α tends to $\pm\infty$.

The transformation law is assumed to be analytic in the parameter so that one can expand in a series about the group identity. For simplicity, suppose it is possible to select $\alpha = 0$ as the identity. Since $\tilde{x}(x, y, 0) = x$ and $\tilde{y}(x, y, 0) = y$, one has

$$g_\alpha\, x = \tilde{x}(x, y, \alpha) = x + \alpha \frac{\partial \tilde{x}(x, y, \alpha)}{\partial \alpha}\bigg|_{\alpha\to 0} + \cdots$$

$$g_\alpha\, y = \tilde{y}(x, y, \alpha) = y + \alpha \frac{\partial \tilde{y}(x, y, \alpha)}{\partial \alpha}\bigg|_{\alpha\to 0} + \cdots$$

For the moment, let us say the two linear terms are not both zero. Assuming this is so, since they are functions only of x and y, I will define, as one usually does, the two very significant functions of x and y,

$$\xi(x, y) = \frac{\partial \tilde{x}(x, y, \alpha)}{\partial \alpha}\bigg|_{\alpha\to 0}, \qquad \eta(x, y) = \frac{\partial \tilde{y}(x, y, \alpha)}{\partial \alpha}\bigg|_{\alpha\to 0}.$$

These two functions $\xi(x, y)$ and $\eta(x, y)$ specify the group action completely. An application of the transformation g_α assigns a displacement to each point (x, y) in the plane from (x, y) to $(\tilde{x}, \tilde{y})$, which, when α is small is a point nearby. If α were time, one could think of the group action as motion of a fluid. The velocity vector of a fluid element near (x, y) would be $(\xi(x, y), \eta(x, y))$. In physical terms $\xi(x, y)$ and

$\eta(x, y)$ are like the components of a velocity field of a steady fluid flow. The functions $\tilde{x}(x, y, \alpha)$ and $\tilde{y}(x, y, \alpha)$ can be found, at least for some valid parametrization, by integrating the autonomous system

$$\frac{d\tilde{x}}{d\alpha} = \xi(\tilde{x}, \tilde{y}), \qquad \frac{d\tilde{y}}{d\alpha} = \eta(\tilde{x}, \tilde{y}),$$

subject to the initial conditions $\tilde{x}(x, y, 0) = x$, $\tilde{y}(x, y, 0) = y$. This means a one-parameter group can be reconstructed from the field of tangent vectors to the orbits—or in other words tangent vectors to the flow lines—having components $(\xi(x, y), \eta(x, y))$.

Now consider some function $f(x, y)$. The group transformations act on this through a substitution recipe,

$$g_\alpha f(x, y) = f(\tilde{x}(x, y, \alpha), \tilde{y}(x, y, \alpha)) \equiv F(x, y, \alpha),$$

where f is arbitrary, except that F is assumed to have a convergent Taylor expansion in α, so that

$$F(x, y, \alpha) = F(x, y, 0) + \alpha \frac{\partial F(x, y, \alpha)}{\partial \alpha}\bigg|_{\alpha\to 0} + \frac{1}{2!}\alpha^2 \frac{\partial^2 F(x, y, \alpha)}{\partial \alpha^2}\bigg|_{\alpha\to 0} + \cdots$$

Then notice

$$\begin{aligned}\frac{\partial F(x, y, \alpha)}{\partial \alpha}\bigg|_{\alpha\to 0} &= \frac{\partial f(\tilde{x}, \tilde{y})}{\partial \tilde{x}}\bigg|_{\alpha\to 0} \frac{\partial \tilde{x}}{\partial \alpha}\bigg|_{\alpha\to 0} + \frac{\partial f(\tilde{x}, \tilde{y})}{\partial \tilde{y}}\bigg|_{\alpha\to 0} \frac{\partial \tilde{y}}{\partial \alpha}\bigg|_{\alpha\to 0} \\ &= \left(\xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y}\right) f(x, y),\end{aligned}$$

so that in fact for any $f(x, y)$ transformed by the group,

$$\left(\xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y}\right) f(x, y) = \frac{d}{d\alpha} f(\tilde{x}, \tilde{y})\bigg|_{\alpha\to 0}.$$

The derivative $d/d\alpha$ on the right-hand side is the partial with respect to α holding the initial coordinates x and y fixed. Since this relation is supposed to hold for any sufficiently well behaved function, i.e. one equal to its convergent power series expansion, the same sequence of steps can be carried out again to obtain second and higher derivatives of $f(\tilde{x}, \tilde{y})$ evaluated at $\alpha = 0$. Let

$$U = \xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y},$$

then

$$U^n f(x, y) = \frac{d^n}{d\alpha^n} f(\tilde{x}, \tilde{y})\bigg|_{\alpha\to 0}.$$

The differential operator U, a sort of directional derivative, has several common names. Since $\xi(x, y)$ and $\eta(x, y)$ are components of a tangent vector along the orbit passing through (x, y), U is sometimes called a tangent vector. It represents the field of flow due to the group action. When applied to scalar functions such as $f(x, y)$, U is also a Lie derivative. Perhaps the most insightful name for U is the one used by Lie who called U an *infinitesimal generator* of the group. Why is that? The answer lies in the Taylor series expansion, which simplifies now as follows:

$$f(\tilde{x}, \tilde{y}) = \left(1 + \alpha U + \frac{1}{2!}\alpha^2 U^2 + \cdots\right) f(x, y) = e^{\alpha U} f(x, y).$$

This exhibits $e^{\alpha U}$ as a rather concrete representation of g_α. It could be defined formally by the operator summation, which makes sense for a domain of power-series expandable functions. It could be defined alternatively as the group action in terms of the solution to the autonomous system of DEs written previously, namely

$$\frac{d}{d\alpha}\tilde{x} = \xi(\tilde{x}, \tilde{y}), \qquad \frac{d}{d\alpha}\tilde{y} = \eta(\tilde{x}, \tilde{y}),$$

subject to $(\tilde{x}, \tilde{y}) = (x, y)$ at $\alpha = 0$.

From the identification

$$g_\alpha = e^{\alpha U}$$

it follows at once that the group is abelian, since

$$g_\alpha\, g_\beta = e^{\alpha U}\, e^{\beta U} = e^{(\alpha+\beta)U} = e^{\beta U}\, e^{\alpha U} = g_\beta\, g_\alpha.$$

Then clearly

$$g_\alpha\, g_\beta = g_{\alpha+\beta},$$

and the parametrization is additive, $\phi(\alpha, \beta) = \alpha + \beta$. Then notice that

$$g_\alpha = e^{\alpha U} = (e^U)^\alpha,$$

and so in direct analogy with the case of the finite groups, one can see that the one-parameter Lie groups are in a sense cyclic groups generated by the element e^U. Accordingly, the operator U is the infinitesimal generator of the group. It is rather like a logarithm of a group generator. Multi-parameter groups have sets of infinitesimal generators, one for each parameter. We consider only groups with finitely many parameters which hence are *finitely generated*. The generators need not commute as operators, and so as seen for finite groups the Lie groups need not be abelian.

It was mentioned previously that one-parameter groups can be isomorphic to the group of addition of real numbers or the group of addition modulo 2π of angles of rotation in the plane. This is an interesting difference between the one-parameter groups and the finite cyclic groups. Each derives from a single generating element. However, for the finite case, there must be some power p of a generator g such that $g^p = e$, the identity. This argument is based on finiteness of $|G|$ and it does not work

when the group is infinite. Starting with some g_α there may or may not be a power $p \neq 0$ such that g_α^p gives the identity. Of course, in any case the identity in the set $S = \{g_\alpha^p\}$ is supplied by including $\alpha \to 0$ or $p = 0$. This is an important distinction. When there is an element such that $g_\alpha^p = e$ for some finite p it implies a sort of periodicity—periodic boundary conditions if you will—and as in the case of the Fourier series expansion it implies the possibility of expanding in a set of orthogonal functions. The distinction is in whether the range of the parameter α can be made finite or not, or more accurately, whether or not it corresponds to a closed and bounded set on the group manifold. This will come up again later.

Exercise. Find the Lie generator corresponding to the special conformal transformation

$$\tilde{x} = \frac{x + \alpha(x^2 + y^2)}{(1 + \alpha x)^2 + \alpha^2 y^2}, \quad \tilde{y} = \frac{y}{(1 + \alpha x)^2 + \alpha^2 y^2}.$$

2.4 Generator example: Lorentz boost

One utility of the generator is that one has a recipe to re-parameterize a group to simply $\phi(\alpha, \beta) = \alpha + \beta$. For example, consider the homogeneous Lorentz transformation, or a boost in the x direction. This set of transformations describes a change of reference frame in the theory of relativity from one inertial frame with space and time coordinates (x, t) to another moving past the first toward the right (positive x direction) at constant speed v in which the new coordinates are $(\tilde{x}, \tilde{t})$. So in this case t takes the place of y. The transformation is linear and is made homogeneous by setting clocks in such a way that $t = 0$ and $\tilde{t} = 0$ at the moment in either frame when the two frames of reference coincide. Of course this means that the 'event' of coincidence has coordinates $(x, t) = (0, 0) = (\tilde{x}, \tilde{t})$. Then by requiring that the speed of light c be the same in either frame one finds, following Einstein and others, that

$$\tilde{x} = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \qquad \tilde{t} = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}}.$$

A natural dimensionless parameter is the ratio $\beta = v/c$ of the relative speed between the reference frames to light speed. So

$$g_\beta\, x = \tilde{x}(x, t, \beta) = \frac{x - \beta ct}{\sqrt{1 - \beta^2}}, \qquad g_\beta\, t = \tilde{t}(x, t, \beta) = \frac{1}{c}\frac{ct - \beta\, x}{\sqrt{1 - \beta^2}}.$$

By substituting and a bit of simplifying, the composition of two such transformations for different parameter values α and β gives the familiar velocity transformation law,

$$g_\alpha\, g_\beta = g_\gamma, \qquad \text{where} \qquad \gamma = \frac{\alpha + \beta}{1 + \alpha\, \beta}.$$

One time through the derivation of this composition law will convince the reader that this parametrization is not the most convenient one. So next I compute the infinitesimal generator,

$$U = \xi(x,\, t)\frac{\partial}{\partial x} + \tau(x,\, t)\frac{\partial}{\partial t}, \qquad \text{where}$$

$$\xi(x,\, t) = \frac{\partial}{\partial \beta}\tilde{x}(x,\, t,\, \beta)\bigg|_o = -ct, \quad \tau(x,\, t) = \frac{\partial}{\partial \beta}\tilde{t}(x,\, t,\, \beta)\bigg|_o = -\frac{x}{c},$$

and then reconstruct the group by exponentiating it. I will use θ as the new group parameter. Observe that it cannot be identified with β since it will be additive by construction, where the original β is not (namely the previous combination rule),

$$\frac{d\tilde{x}}{d\theta} = -c\tilde{t}, \qquad \frac{d\tilde{t}}{d\theta} = -\frac{1}{c}\,\tilde{x},$$

$$\frac{d^2\tilde{x}}{d\theta^2} - \tilde{x} = 0 \qquad \tilde{x}(x,\, t,\, \theta) = A(x,\, t)\cosh(\theta) + B(x,\, t)\sinh(\theta).$$

From the initial condition on $\tilde{x}$, which is to say the limit at the identity, one sees that $A(x,\, t) = x$. Then

$$\frac{d\tilde{x}}{d\theta} = \frac{d}{d\theta}(x\cosh(\theta) + B(x,\, t)\sinh(\theta)) = -c\tilde{t}$$

and so the other condition at the identity gives $B(x,\, t) = -ct$, and I arrive at a different parametrization:

$$g_\theta\, x = \tilde{x}(x,\, t,\, \theta) = x\cosh\theta - ct\sinh\theta,$$
$$g_\theta\, t = \tilde{t}(x,\, t,\, \theta) = t\cosh\theta - \frac{1}{c}\,x\sinh\theta.$$

The additive group parameter θ is the *rapidity*. To connect it to the original β, consider an object sitting at rest at the origin of the frame with coordinates $(\tilde{x},\, \tilde{t}\,)$. So the object is moving to the right with speed v when seen from the other frame of reference. So $\tilde{x} = 0$ and

$$0 = x\cosh\theta - ct\sinh\theta, \qquad \text{with} \qquad x = vt.$$

Thus,

$$v/c = \beta = \frac{\sinh\theta}{\cosh\theta} = \tanh\theta.$$

The original composition law is recovered from the addition in θ via

$$\theta_\gamma = \theta_\alpha + \theta_\beta \quad \Rightarrow \quad \gamma = \tanh\big(\theta_\alpha + \theta_\beta\big) = \frac{\sinh\big(\theta_\alpha + \theta_\beta\big)}{\cosh\big(\theta_\alpha + \theta_\beta\big)},$$

and hence

$$\gamma = \frac{\sinh(\theta_\alpha)\cosh(\theta_\beta) + \cosh(\theta_\alpha)\sinh(\theta_\beta)}{\cosh(\theta_\alpha)\cosh(\theta_\beta) + \sinh(\theta_\alpha)\sinh(\theta_\beta)} = \frac{\alpha + \beta}{1 + \alpha\beta}.$$

where

$$\alpha = \tanh(\theta_\alpha), \qquad \beta = \tanh(\theta_\beta).$$

2.5 Transformations acting in three or more dimensions

When the transformations act on more than the two variables x and y, then I will sometimes use the general notations $x = \{x^1, x^2, \ldots, x_n\}$ and $g_\alpha\, x^k = \tilde{x}^k(x, \alpha) = x^k + \alpha\xi^k(x) + \cdots$ so that

$$U = \xi^k(x)\frac{\partial}{\partial x^k},$$

where the repeated index k is assumed to be summed on from 1 to n. Or in particular, for three dimensions I might write simply

$$U = \xi(x, y, z)\frac{\partial}{\partial x} + \eta(x, y, z)\frac{\partial}{\partial y} + \zeta(x, y, z)\frac{\partial}{\partial z}.$$

Here are some standard examples of one-parameter groups of transformations acting on functions of the three-dimensional coordinates are x, y, and z. In each case the parametrization is additive.

1. Translation in the x direction: $g_\alpha f(x, y, z) = f(x + \alpha, yz)$, where

$$U = \frac{\partial}{\partial x}.$$

2. Rotation about y axis: $g_\alpha f(x, y, z) = f(x\cos\alpha - z\sin\alpha, y, z\cos\alpha + x\sin\alpha)$ so that

$$\xi = \frac{\partial\tilde{x}}{\partial\alpha}\bigg|_{\alpha=0} = -z, \qquad \eta = 0, \qquad \zeta = \frac{\partial\tilde{z}}{\partial\alpha}\bigg|_{\alpha=0} = x.$$

Thus, since

$$\tilde{x} = x - \alpha z + \cdots \qquad \tilde{y} = y, \qquad \tilde{z} = z + \alpha x + \cdots,$$

one finds

$$U = -z\,\frac{\partial}{\partial x} + x\,\frac{\partial}{\partial z}.$$

3. Dilation: $g_\alpha f(x, y, z) = f(e^\alpha x, e^\alpha y, e^\alpha z)$. Thus in brief notation

$$\xi = x, \qquad \eta = y, \qquad \zeta = z,$$

$$U = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + z \frac{\partial}{\partial z}.$$

2.6 Changing coordinates

How a transformation group acts on a space in a particular case is generally described using two sets of coordinates. The transformation formulas contain the group parameters, $\{\alpha, \beta, \ldots\}$ and the space coordinates $\{x, y, \ldots\}$. It was seen above that a group can be re-parameterized. The same one-parameter group can be defined using a new parameter that is a monotone function of the original one. It is possible also, of course, to change coordinates in the space where the transformations act.

Suppose the two coordinates (x, y) are transformed in terms of new coordinates (u, v) by substitution formulas

$$x = x(u, v), \qquad y = y(u, v),$$

and the inverse

$$u = u(x, y), \qquad v = v(x, y).$$

The action of a group element g_α is defined on any function by substitution.

$$g_\alpha u = g_\alpha u(x, y) = u(\tilde{x}(x, y, \alpha), \tilde{y}(x, y, \alpha))$$

so that one finds the action in terms of (u, v) is

$$g_\alpha u = \tilde{u}(x, y, \alpha) = u(\tilde{x}(x(u, v), y(u, v), \alpha), \tilde{y}(x(u, v), y(u, v), \alpha)),$$
$$g_\alpha v = \tilde{v}(x, y, \alpha) = v(\tilde{x}(x(u, v), y(u, v), \alpha), \tilde{y}(x(u, v), y(u, v), \alpha)).$$

This can represent a bit of algebra, but the idea is simple enough.

Example. Consider the transformation

$$u = xy, \qquad v = \frac{x}{y}.$$

If x and y are both positive, u and v are also. Inverting, I choose signs so as to return to the first quadrant.

$$x = \sqrt{u}\ \sqrt{v}, \qquad y = \frac{\sqrt{u}}{\sqrt{v}}.$$

For a proper demonstration with this change of coordinates, the group action should be one that permutes points in the first quadrant. Each transformation should take the first quadrant onto itself. So consider uniform dilation,

$$g_\alpha(x, y) = (e^\alpha x, e^\alpha y).$$

So, start with (u, v), write these in terms of (x, y), apply the group and then transform back to u and v:

$$(\tilde{u}, \tilde{v}) = g_\alpha(u, v) = g_\alpha\left(xy, \frac{x}{y}\right) = \left(e^{2\alpha} xy, \frac{x}{y}\right) = (e^{2\alpha}u, v).$$

Notice that the final result need no longer be restricted to the first quadrant in (x, y). This concludes the example.

2.7 Changing variables in the generator

It is easy to transform the infinitesimal group generator from one set of coordinates to another through the chain rule. Suppose ξ and η are given such that

$$U = \xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y}.$$

In the new coordinates

$$\bar{U} = \bar{\xi}(u, v)\frac{\partial}{\partial u} + \bar{\eta}(u, v)\frac{\partial}{\partial v},$$

where

$$\bar{\xi}(u, v) = \frac{\partial}{\partial \alpha}\tilde{u}(u, v, \alpha)\Big|_{\alpha\to 0},$$

$$\bar{\eta}(u, v) = \frac{\partial}{\partial \alpha}\tilde{v}(u, v, \alpha)\Big|_{\alpha\to 0}.$$

Using the chain rule and somewhat abbreviated notation

$$\bar{\xi}(u, v) = \frac{\partial}{\partial \tilde{x}}u(\tilde{x}, \tilde{y})\frac{\partial \tilde{x}}{\partial \alpha}\Big|_{\alpha\to 0} + \frac{\partial}{\partial \tilde{y}}u(\tilde{x}, \tilde{y})\frac{\partial \tilde{y}}{\partial \alpha}\Big|_{\alpha\to 0} = \left(\xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y}\right)u(x, y),$$

or finally

$$\bar{\xi}(u, v) = (U\ u(x, y))_{x\to x(u,v), y\to y(u,v)},$$

$$\bar{\eta}(u, v) = (U\ v(x, y))_{x\to x(u,v), y\to y(u,v)}.$$

This construction for the uniform dilation generator in the previous example in which the new coordinates are $u = xy$ and $v = \frac{x}{y}$ becomes

$$\bar{\xi} = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)xy = 2xy = 2u,$$

$$\bar{\eta} = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)\frac{x}{y} = \frac{x}{y} - \frac{x}{y} = 0,$$

and thus

$$\bar{U} = 2u\frac{\partial}{\partial u}.$$

So the generator is simplified sometimes by an appropriate coordinate change. In fact, had I selected coordinates $t = x/y$, $s = \frac{1}{2}\ln(xy)$ the generator would have become

$$\bar{U} = \frac{\partial}{\partial s}.$$

The general pattern is summarized in one line in the following way. In order to convert from U expressed in variables (x, y) to a generator $\bar{U}$ expressing the same group action in a different set of coordinates (u, v),

$$\bar{U} = \left((U\ u(x, y))\frac{\partial}{\partial u} + (U\ v(x, y))\frac{\partial}{\partial v}\right)\Bigg|_{x\to x(u,v), y\to y(u,v)}.$$

Rotations. As mentioned earlier with the introduction of transformation groups, the action of a group can be seen from two different points of view. These are sometimes described as active or passive. There is a duality between them. In the first place, the coordinates are moved by the group transformation along the orbits as the parameter increases. But since the group is viewed as acting on functions via substitution, the more natural way to think of this is in terms of level curves of the function values. The level curves, when graphed with respect to the original coordinates, move in the opposite direction. Consider the simplest example:

$$\tilde{x} = x + \alpha.$$

The points are transformed into new points that move to the right with increasing α. But with the substitution

$$f(x) \to f(\tilde{x}) = f(x + \alpha),$$

points of constant f move to the *left* as α increases continuously. The first description focuses on how the coordinates change and the second focuses on how the function values change. Accordingly, rotations can be described by how the coordinates move, or by how the level curves of the functions graphed with respect to the original coordinates move under substitution. They rotate in contrary directions.

In Cartesian coordinates

$$U = x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x},$$

rotates function level contours clockwise around the positive z axis. This is easy to see in the polar coordinates where with $x = r\cos\theta$, $y = r\sin\theta$ this becomes

$$\bar{U} = \frac{\partial}{\partial \theta}.$$

Thus

$$e^{\alpha \bar{U}} f(r, \theta) = f(r, \theta + \alpha).$$

Again this is a very simple form. Polar coordinates are convenient for treating rotations.

2.8 Invariant functions, invariant curves, and groups that permute curves in a family

The examples of transformations in three dimensions, and particularly the case of rotation, suggest there may be a preferred coordinate system in which the action of a given group simplifies. These are canonical or normal coordinates, which will be the subject of the section following this one. The current section will introduce the general notion of groups that either leave the curves or surfaces of a given family invariant, or else take one member of the family into another.

A function $F(x, y)$ is said to be invariant under the group G if it is not changed by any group transformation g_α. So if G is a one-parameter Lie group, and if U is an infinitesimal generator of G, then $F(x, y)$ is invariant under G if $g_\alpha F(x, y) = F(x, y)$ or, in more detail, if

$$g_\alpha F(x, y) = e^{\alpha U} F(x, y) = \left(1 + \alpha U + \frac{1}{2}\alpha^2 U^2 + \cdots\right) F(x, y) = F(x, y).$$

If $U F(x, y)$ is zero, then $F(x, y)$ is invariant under G. The converse is also true, since I assume analyticity and there is no α on the right-hand side. Thus $F(x, y)$ is invariant under G if and only if

$$U F(x, y) = 0.$$

For example, suppose

$$U = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}.$$

It is fairly clear that $U xy = 0$. What group does this U generate?

$$\frac{d\tilde{x}}{d\alpha} = \tilde{x}, \qquad \frac{d\tilde{y}}{d\alpha} = -\tilde{y}, \qquad \Rightarrow \quad \tilde{x} = xe^{\alpha}, \quad \tilde{y} = ye^{-\alpha}.$$

To solve

$$U F = 0,$$

I observe that any function of the product xy is invariant under the action of the group. In other words, $F(x, y) = f(xy)$ for any differentiable f is invariant under G.

In general, when a one-parameter group acts on an n-dimensional space, it moves the points along orbits. The orbits are one-dimensional subspaces. At any point of a particular orbit in an n-dimensional space there is one direction along the orbit, and it would seem there should be $n - 1$ directions perpendicular to it at each point.

Consider a curve defined implicitly by

$$F(x, y) = k.$$

where k is some particular constant. If the curve itself is invariant with respect to the action of G, then this expression has to be the same before and after any transformation acts.

Figure 2.3 shows two curves from the one-parameter family of ellipses

$$\frac{x^2}{9} + y^2 = k.$$

For some fixed value of α the points a, b, and c on the inner curve $k = k_1$ are mapped by dilation to the corresponding points a', b', and c' on the outer curve $k = k_2$, where dilation means

$$g_\alpha\, x = \tilde{x} = x\, e^\alpha, \qquad g_\alpha\, y = \tilde{y} = y\, e^\alpha,$$

So what does dilation do, exactly?

$$g_\alpha\left(\frac{x^2}{9} + y^2\right) = \left(\frac{x^2}{9} + y^2\right)e^{2\alpha}, \qquad \Rightarrow \left(\frac{x^2}{9} + y^2\right) = e^{-2\alpha}k_1 = k_2.$$

The effect is to take the ellipse labeled by k_1 into another labeled by k_2. One can say the dilation group *permutes* the curves of the family, meaning that each transformation of the group takes any of the ellipses into some other member of the family.

To find a group that transforms a point such as a on one curve of the family into another point on the *same* curve, the group action has to take $F(x, y) = k_1$ into

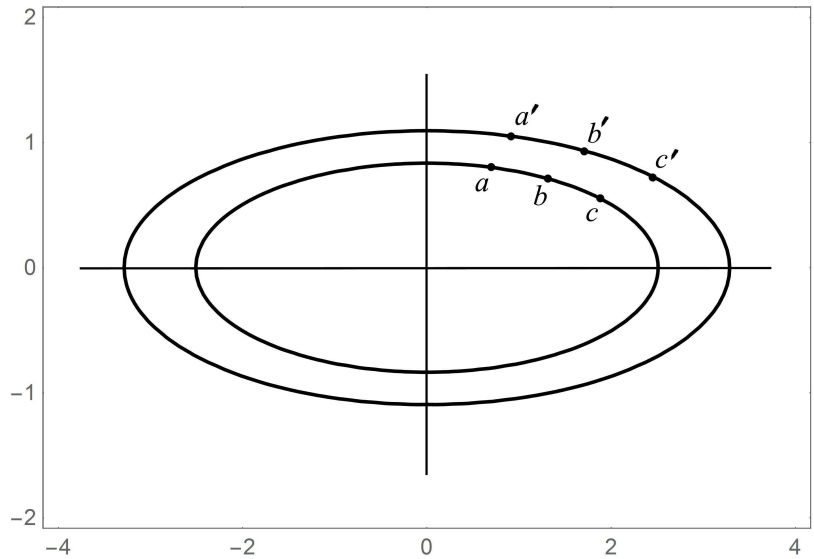


Figure 2.3. Two ellipses of the family $\frac{x^2}{9} + y^2 = k_1$. The points a, b, c on the inner ellipse $k = k_1$ map under a particular dilation onto a', b', c' on the outer ellipse $k = k_2$.

$F(\tilde{x}, \tilde{y}) = k_1$. The function has to go back into itself. That means either both $\xi(x, y)$ and $\eta(x, y)$ vanish on the curve, so that in fact the group does not move any of the points on the curve at all, or else the functions $\xi(x, y)$ and $\eta(x, y)$ do not vanish, but the group leaves $F(x, y)$ invariant. So, in either case, the curve $F(x, y) = k_1$ is invariant if

$$g_\alpha F(x, y) = k_1, \qquad \text{or} \qquad e^{\alpha U} F(x, y) = F(\tilde{x}, \tilde{y}) = F(x, y).$$

Accordingly the curve is invariant if and only if the function is invariant under the action of the group in a region containing the curve.

$$U F(x, y) = 0.$$

So it would seem that an invariant curve must be made up of group orbits.

Example. Consider the family of curves in the example of figure 2.3. To find a group that takes a curve of the family into itself in this example, one needs

$$\left(\xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y}\right)\left(\frac{x^2}{9} + y^2\right) = 0 = 2\,\xi(x, y)\frac{x}{9} + 2\,\eta(x, y)y.$$

A simple choice would be $\xi = 9\,y$, $\eta = -x$ so that

$$U = 9\,y\,\frac{\partial}{\partial x} - x\,\frac{\partial}{\partial y}.$$

This generates the group transformation

$$e^{\alpha U}\,x = x\cos 3\alpha + 3y\sin 3\alpha, \qquad e^{\alpha U}\,y = y\cos 3\alpha - \frac{1}{3}\,x\sin 3\alpha.$$

There is some value for α such that $e^{\alpha U}$ takes point a into point b, there is a value of α (probably different) that takes b' into c'. The curves of the family are the orbits of the group. In a higher dimensional space the family would be a family of hypersurfaces and the group flow generated by U would be *within* the hypersurfaces. In higher dimensions, there would be several different generators of this type.

In general, in two dimensions, a group G acting on two coordinates permutes the curves of a family if it takes one curve into another of the same family. (In higher dimensions, it would move one hypersurface into another.) Two curves of a family

$$F(x, y) = k$$

correspond to two different choices of the constant k, as was just seen for the example shown in figure 2.3. Suppose I did not know that dilation permutes this family of curves, and that I wanted to find a group that permutes them. Let V be the generator of this unknown group. This would mean that if point a and point b are on one curve of the family, such as $F(x, y) = k_1$, and for some particular parameter value β, the transformation $e^{\beta V}$ would take a into a' on another curve of the family,

and, since a and b are points on the first curve and since the group takes curves into curves, the same parameter value β must take b into another point b' on the same curve as a'. The basic idea must be that $e^{\beta V}$ permutes curves of $F(x, y) = k$ if

$$F(x, y) = k, \qquad \Rightarrow \qquad e^{\beta V} F(x, y) = K(k, \beta),$$

where $K(k, \beta)$ is some other constant that depends on the original constant and the value of the group parameter. Since $\beta = 0$ is the identity, $K(k, 0) = k$. Differentiate the $K(k, \beta)$ with respect to β and evaluate at $\beta = 0$:

$$\frac{\partial}{\partial \beta}\left(e^{\beta V} F(x, y)\right) = \frac{\partial K(k, \beta)}{\partial \beta},$$

$$V\left(e^{\beta V} F(x, y)\right) = \frac{\partial K(k, \beta)}{\partial \beta},$$

$$V F(x, y) = \frac{\partial K(k, \beta)}{\partial \beta}\bigg|_{\beta=0}.$$

The right-hand side must now be some function H only of the original value k. But that means, since $F(x, y) = k$, it must be that

$$V F(x, y) = H(F(x, y)).$$

The criterion that the group action should permute the curves of the family is that V should take $F(x, y)$ into some function of $F(x, y)$.

What function H actually is does not matter very much, since the original curve could be re-written using any arbitrary function of $F(x, y)$ rather than F itself, say

$$M(F(x, y)) = M(k_1) = k_2.$$

Then using the chain rule,

$$V M(F(x, y)) = M'(F(x, y))V F(x, y) = M'(F(x, y))H(F(x, y)).$$

Since one is free to choose the function M, suppose

$$M'(x) = \frac{1}{H(x)}.$$

So

$$V M(F(x, y)) = 1.$$

This means if I define $s(x, y) = M(F(x, y))$, then s is moved by the group generated by V in such a way that $\tilde{s} = s + \alpha$. (Of course, $F(x, y)$ is a variable left constant by another group generated by U.) However, it is not necessary to find M and to transform from $F(x, y)$ to $s(x, y)$. All that is required is to find a V such that $V F(x, y)$ gives back some function of $F(x, y)$. This guarantees that V generates a group that permutes the curves of the family $F(x, y) = k$.

2.9 Canonical coordinates for a one-parameter group

In the cases of rotation and dilation it was seen that in certain convenient coordinates, the group generator had a simple form $\partial/\partial s$. Referring in general to the discussion at the end of the previous section, as applied to the example of figure 2.3, when one applies the uniform dilation generator,

$$V\ F(x,\,y) = \left(x\,\frac{\partial}{\partial x} + y\,\frac{\partial}{\partial y}\right)\left(\frac{x^2}{9} + y^2\right) = 2\left(\frac{x^2}{9} + y^2\right),$$

it turns out that indeed the dilation group permutes the ellipses in this family. It takes any one of the ellipses into another. The variable s in this case is

$$s(x,\,y) = \frac{1}{2}\ln\left(\frac{x^2}{9} + y^2\right).$$

So a question comes up: when a one-parameter group of transformations acts on an m-dimensional space, is it always possible to change to a set of coordinates $\{s,\,t^1,\,t^2,\,\ldots,\,t^{n-1}\}$, where

$$U = \frac{\partial}{\partial s},$$

holding the other $n-1$ coordinates $\{t^1,\,t^2,\,\ldots,\,t^{n-1}\}$ constant? If so, the action of the group on any function of the coordinates would become very simple. To wit,

$$g_\alpha f(s,\,t^1,\,t^2,\,\ldots) = e^{\alpha U} f(s,\,t^1,\,t^2,\,\ldots) = f(s+\alpha,\,t^1,\,t^2,\,\ldots).$$

The group, 'moves', or translates the first coordinate while leaving the others unchanged. Is it always possible to choose such coordinates? The answer is yes. These are *canonical coordinates*, or canonical variables—sometimes called normal coordinates for the group. To construct them involves solving a set of partial DEs (PDEs). Because in such coordinates $\bar{U} = \frac{\partial}{\partial s}$, one finds

$$\bar{U}\ s = 1, \qquad \bar{U}\ t^k = 0 \quad \text{for} \quad k = 1,\,2,\,\ldots\,n-1.$$

Accordingly, these same equations when written in terms of the original coordinates,

$$U\ s(x^1,\,x^2,\,\ldots,\,x^n) = 1, \qquad U\ t^k(x^1,\,x^2,\,\ldots,\,x^n) = 0, \quad k = 1,\,2,\,\ldots\,n-1,$$

become a set of n PDEs for the canonical variables. One follows a standard procedure to solve them.

In the case of two dimensions, for instance, one has the two PDEs

$$U\ s = \xi(x,\,y)\frac{\partial s}{\partial x} + \eta(x,\,y)\frac{\partial s}{\partial y} = 1, \qquad U\ t = \xi(x,\,y)\frac{\partial t}{\partial x} + \eta(x,\,y)\frac{\partial t}{\partial y} = 0.$$

Consider the second equation first:

$$\xi\,\frac{\partial t}{\partial x} + \eta\,\frac{\partial t}{\partial y} = 0, \qquad \Rightarrow \qquad (\xi,\ \eta)\cdot(\partial t/\partial x,\ \partial t/\partial y) = 0,$$

and simultaneously, if I consider a level curve along which t = some constant,

$$dt = \frac{\partial t}{\partial x}\,dx + \frac{\partial t}{\partial y}\,dy = 0, \qquad \Rightarrow \qquad (dx,\, dy)\cdot(\partial t/\partial x,\, \partial t/\partial y) = 0.$$

Therefore, along a level curve of $t(x, y)$, the two two-dimensional vectors (ξ, η) and (dx, dy) are both perpendicular to the vector $(\partial t/\partial x, \partial t/\partial y)$ and so they must be parallel. This gives the ordinary DE

$$\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)}.$$

In principle this integrates to give $t(x, y)$. One is interested primarily in the level curves $t(x, y) = \text{const}$.

Now consider the first PDE, the one for $s(x, y)$, the coordinate along which the group action takes place. Here I find

$$\xi\frac{\partial s}{\partial x} + \eta\,\frac{\partial s}{\partial y} = 1, \quad \Rightarrow (\xi, \eta, 1)\cdot(\partial s/\partial x,\, \partial s/\partial y,\, -1) = 0,$$

$$ds = \frac{\partial s}{\partial x}\,dx + \frac{\partial s}{\partial y}\,dy, \quad \Rightarrow (dx, dy, ds)\cdot(\partial s/\partial x,\, \partial s/\partial y,\, -1) = 0.$$

It is sufficient that (dx, dy, ds) be parallel to $(\xi, \eta, 1)$. So there are two ordinary DEs,

$$\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)} = ds.$$

The first equation $dx/\xi = dy/\eta$ is satisfied along contours of constant t. Constant t means staying at the same group orbit. Thus along such an orbit $t = t(x, y)$ can be solved for x if $\partial t/\partial x \neq 0$, or for y if $\partial t/\partial y \neq 0$. Let us say it can be solved to give $y = y(x, t)$. Then

$$s(x, y) = s(x_o, y_o) + \left(\int_{x_o}^{x} \frac{dx}{\xi(x, y(x, t))}\right)\Bigg|_{t\to t(x,y)}.$$

Both s and t have a great deal of ambiguity, as whenever $U\ t(x, y) = 0$, one has that $U\ f(t(x, t)) = 0$ also, so if f and h are any smooth functions,

$$s_1 = s + f(t) \qquad \text{and} \qquad t_1 = h(t)$$

are also suitable canonical coordinates for the same group.

To summarize: in the space where a one-parameter group of transformations acts, canonical coordinates may be chosen such that the group action moves only one of the coordinates s while holding the others fixed. The coordinate s may be scaled in such a way that $U = \partial/\partial s$. So the group action is quite simple when expressed in these coordinates.

References

[1] Cohen A 1911 *An Introduction to the Lie Theory of One-parameter Groups with Applications to the Solution of Differential Equations* (New York: Heath)

[2] Cornell University Library 2016 *Historical Math Monographs* http://ebooks.library.cornell.edu/m/math/

[3] Olver P J 1986 *Application of Lie Groups to Differential Equations* (New York: Springer)

[4] Stephani H 1989 *Differential Equations: Their Solution Using Symmetry* (New York: Cambridge University Press)

IOP Concise Physics

Lectures on Selected Topics in Mathematical Physics: Introduction to Lie Theory with Applications

William A Schwalm

Chapter 3

Ordinary differential equations

The original use Sophus Lie made of Lie groups was for integrating DEs. Again, in order to fix ideas and notations, a few brief comments are given on the general topic of DEs. I shall use quite often the older notations:

$$p \equiv y'(x) = \frac{dy}{dx},$$
$$q \equiv y''(x) = \frac{d^2y}{dx^2} = \frac{dp}{dx}.$$

For the most part, the development below will not go beyond equations of second order, both because there is not much new to be learned by doing so, and because the chances of encountering equations in physics with derivatives of order three or higher are relatively small. However, one will always have in mind the general case, and where it is appropriate the extension to DEs of any order will be indicated.

A DE is an equation containing derivatives of an unknown function. For instance,

$$y\frac{dy}{dx} + x^2y^2 = 0$$

or in other notation

$$yp + x^2y^2 = 0,$$

is a first order DE for $y(x)$. One calls x the independent and y the dependent variable. It should be clear that the roles of dependent and independent variable can interchange depending on circumstance. A particular solution to this DE is

$$y(x) = e^{-\frac{1}{3}x^3},$$

doi:10.1088/978-1-6817-4449-0ch3 © Morgan & Claypool Publishers 2017

since substituting this for y reduces the equation to an identity. The general solution, however, contains an arbitrary parameter, to wit

$$y(x) = ae^{-\frac{1}{3}x^3}.$$

To see how it happens that the general solution to an nth order DE has n parameters, consider the equation

$$\phi(x, y, a, b) = ax^2y - by^2x + a^2 + 1 = 0.$$

This defines a two-parameter family of functions in x and y. The parameters are a and b. To derive the DE, eliminate the parameters a and b by differentiating. So I differentiate with respect to x holding a and b fixed:

$$\begin{aligned}\frac{d}{dx}\phi(x, y, a, b) &= 2axy + ax^2p - 2byx - by^2 \\ &= x(ax - 2by)p + y(2ax - by) = 0, \\ \frac{d^2}{dx^2}\phi(x, y, a, b) &= x(ax - 2by)q - 2bxp^2 \\ &\quad + 4(ax - by)p + 2ay = 0.\end{aligned}$$

Eliminating a and b yields the DE

$$\psi(x, y, p, q) = 3x^2y^2q + 2x^3p^3 - 2y^3 = 0.$$

The n-parameter family of functions $\phi(x, y, c_1, c_2, \ldots) = 0$ is the general solution of an nth order DE $\psi(x, y, p, q, \ldots) = 0$. One also calls $\phi(x, y, c_1, c_2, \ldots) = 0$ the *complete primitive* of the DE. A particular solution results from choosing particular values for one or more of the constants in ϕ. Faced with any DE, an essential first step in finding a solution is classification. This really means classification of the DE or system of DEs by symmetry. Introduction of group theoretic ideas permits a more organized treatment of the subject.

Definition Transformation T is a *symmetry* of the DE $\psi(x, y, p, q, \ldots) = 0$ if

$$T\psi(x, y, p, q, \ldots) = \psi(\tilde{x},\tilde{y},\tilde{p},\tilde{q}, \ldots) = 0 \quad \text{whenever} \quad \psi(x, y, p, q, \ldots) = 0.$$

Transformations g_α that belong to Lie groups are the most interesting, since these are the ones that are easier to find using the machinery of Lie theory. Once found, a symmetry can be used to reduce the order of the DE. If the equation is first order, this means the group gives a way to solve the DE.

A comment is in order here. You may wonder what the magic of Lie's theory actually is. What in fact is the importance of there being a continuous group of transformations at all? The most basic thing is that in this case when a Lie group is involved, for the most part, the whole story is contained in the the linear terms of the Taylor expansions with respect to the group parameter. In all of the following, this

idea will be developed and demonstrated bit by bit. It may be instructive to start with an example.

Example of a linear homogeneous DE. This section is intended to illustrate how a group acts on the DE space, containing the derivative variables p, q etc, as well as x and y, and what it means for the group to be a symmetry of a particular DE.

The terms linear and homogeneous are used in several different ways to classify particular DEs. In each case they refer to scaling properties. When a DE $\psi(x, y, p, q, \ldots) = 0$ is said to be linear homogeneous, or homogeneously linear, it usually means that

$$\psi(x, \lambda y, \lambda p, \lambda q, \ldots) = \lambda\psi(x, y, p, q, \ldots).$$

In other words, ψ scales linearly in y and in the x derivatives of y. So in this expression, one should think of d/dx as a partial derivative with respect to x holding λ fixed. I could describe this situation by writing

$$\frac{d}{dx} = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + q\frac{\partial}{\partial p} + \cdots,$$

where in each case the partial derivatives are taken, 'holding everything else fixed'.

The condition 'holding everything else fixed' is a desirable simplification of the partial derivatives. A transformation of the group in this example is a scaling of the dependent variable y with $e^\alpha = \lambda$:

$$g_\alpha x = x, \qquad g_\alpha y = e^\alpha y.$$

So if the group were to act on functions only of x and y, the generator could be written $U = y\partial/\partial y$. However this will not do, because it gives no instruction for what is to be done with functions of the derivative coordinates p, q, and so on. If the group transformations act on DEs there must be an extended formula for U that takes these derivative coordinates into account also. The general way of doing this will be taken up presently, but in this particular case the scaling should be such that d/dx can be taken either before or after the transformation and the result should be the same. So the form of the generator should be such that $Ux = 0$, $Uy = y$ and when the group acts on the general space including derivative coordinates, if I write

$$\tilde{p} = \lambda p = e^\alpha p, \qquad \tilde{q} = e^\alpha q, \qquad \text{etc},$$

then

$$\eta_1 = p, \qquad \eta_2 = q, \qquad \cdots.$$

Therefore, this kind of scaling corresponds to a generated by

$$U = y\frac{\partial}{\partial y} + p\frac{\partial}{\partial p} + q\frac{\partial}{\partial q} + \cdots$$

when the group acts on DE space. When one calculates the commutator $[U, d/dx]$ the result is zero because terms cancel in pairs to all orders of differentiation. So indeed the x derivative can be taken before or after the group acts and the result will be the same. The group action is of course just a uniform scaling of all the derivatives. The classification of a DE as being homogeneously linear in y then means

$$e^{\alpha U}\psi(x, y, p, q, \ldots) = \psi(x, e^{\alpha}y, e^{\alpha}p, e^{\alpha}q, \ldots) = e^{\alpha}\psi(x, y, p, q, \ldots).$$

The essential point in the context of a DE is that *if* function ψ defining the DE under consideration, namely the DE

$$\psi(x, y, p, q, \ldots) = 0,$$

is scaled by the group in this simple way, then the equation is preserved by the action of the group, so that

$$e^{\alpha U}\psi(x, y, p, q, \ldots) = e^{\alpha}\psi(x, y, p, q, \ldots) = 0.$$

This shows that the resulting, transformed expression vanishes whenever the original DE is satisfied. The new expression is 'zero modulo ψ' in the same sense that 25 is 'zero modulo 5', because ψ divides it evenly. One says that a DE which is homogeneously linear in y in this sense *admits* the group generated by U. Or one might say alternatively that U generates a 'symmetry' of the DE, meaning that the group action transforms it back into the same DE, just as symmetry operations bring the equilateral triangle back into itself. The group is a symmetry of equation if each transformation of the group is a symmetry or, in other words, if each transformation preserves the truth of the equation. But to see that this is so in this example, I had to extend the generator U so as to see how it acts on the larger space including the derivatives p, q, and so on, as well as the original x and y. So now this idea of extending the group action to include transformation of the derivative variables will need to be developed and applied to more general transformation groups.

3.1 Prolongation of the group generator and a symmetry criterion

To be specific, consider the general second order DE,

$$\psi(x, y, p, q) = 0,$$

and suppose that the complete primitive (general solution)

$$\phi(x, y, a, b) = 0$$

is also known somehow. If U generates a symmetry $g_\alpha = e^{\alpha U}$ of the DE, then the symmetry must act on the complete primitive, or general solution, which is a two-parameter family of solutions, in such a way as to take one solution into another. That is to say, since g_α leaves the DE invariant, meaning

$$e^{\alpha U}\psi(x, y, p, q) = 0 \qquad \text{modulo} \quad \psi(x, y, p, q),$$

then the group must 'permute the solutions', so that

$$e^{\alpha U}\phi(x, y, a, b) = 0 \qquad \text{modulo} \quad \phi(x, y, a_1, b_1),$$

the latter meaning that for some $a_1 = A(a, b)$, $b_1 = B(a, b)$,

$$e^{\alpha U}\phi(x, y, a, b) = (\cdots\text{something}\cdots)\phi(x, y, A(a, b), B(a, b)),$$

and so the group takes the solution with parameter values a and b into another (possibly the same) solution having parameter values a_1 and b_1. The group therefore permutes the solutions.

Now the two outstanding questions concern how to find such groups for a given DE, and then how to use the groups once they are found.

Finding groups admitted by a given DE is partly an art. A less ambitious objective is to find a test to determine whether a given DE admits a particular group. The project undertaken by Lie was to enumerate one-parameter groups, starting with the simplest, and for each to determine the most general DEs for which the group is a symmetry.

The first step is to see how to extend a generator U to act on p and q. In the general case, the group action does not commute with d/dx because the group acts in general on both x and y:

$$g_\alpha p = g_\alpha \frac{dy}{dx} = \frac{d\tilde{y}}{d\tilde{x}} = \frac{dy + \alpha d\eta + \cdots}{dx + \alpha d\xi + \cdots},$$

$$g_\alpha p = \frac{p + \alpha\dfrac{d\eta}{dx} + \cdots}{1 + \alpha\dfrac{d\xi}{dx} + \cdots} = \left(p + \alpha\frac{d\eta}{dx} + \cdots\right)\left(1 - \alpha\frac{d\xi}{dx} + \cdots\right),$$

$$g_\alpha p = \tilde{p} = p + \alpha\left(\frac{d\eta}{dx} - p\frac{d\xi}{dx}\right) + \cdots.$$

Thus,

$$g_\alpha = \tilde{p}(x, y, p, \alpha) = p + \eta_1(x, y, p)\alpha + \cdots$$

where

$$\eta_1(x, y, p) = \left(\frac{\partial\tilde{p}}{\partial\alpha}\right)_\alpha = 0 = \frac{d\eta}{dx} - p\frac{d\xi}{dx}.$$

Since

$$\frac{d}{dx} = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + \cdots,$$

this is also

$$\eta_1 = \frac{\partial\eta}{\partial x} + p\left(\frac{\partial\eta}{\partial y} - \frac{\partial\xi}{\partial x}\right) - p^2\frac{\partial\xi}{\partial y}.$$

Therefore, when the group acts on the space of variables including x, y, and p, the generator is

$$U = \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \eta_1\frac{\partial}{\partial p}$$

Sometimes this is denoted by $U^{(1)}$ and is referred to as the *first prolongation* of U.

One would like to be able to transform DEs of any order. There are second and third, etc, prolongations of the group generator, since in a general DE space

$$U = \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \eta_1\frac{\partial}{\partial p} + \eta_2\frac{\partial}{\partial q} + \cdots.$$

In general, if $y^{(n)} = d^n y/dx^n$ and

$$g_\alpha y^{(n)} = y^{(n)} + \alpha\eta_n(x, y, p, q, \ldots y^{(n)}) + \cdots,$$

then proceeding just as in the case of the first derivative

$$g_\alpha y^{(n+1)} = g_\alpha\frac{dy^{(n)}}{dx} = \frac{dy^{(n)} + \alpha d\eta_n + \cdots}{dx + \alpha d\xi + \cdots}$$

$$g_\alpha y^{(n+1)} = y^{(n+1)} + \alpha\eta_{n+1} + \cdots,$$

where

$$\eta_{n+1} = \frac{d\eta_n}{dx} - y^{(n+1)}\frac{d\xi}{dx}.$$

From here on I will normally write U to mean the prolongation of the generator to whatever order is appropriate, unless it becomes necessary to restrict attention to some particular order. One can show by induction [1] that another formula for the extended coefficient η_n is

$$\eta_n(x, y, p, q, \ldots, y^{(n)}) = \frac{d^n}{dx^n}(\eta(x, y) - p\xi(x, y)) + \xi(x, y)y^{(n+1)}.$$

The last term cancels. The explicit forms become complicated very quickly with increasing n. The first two are

$$\eta_1 = \eta_{,x} + (\eta_{,y} - \xi_{,x})p - \xi_{,y}\,p^2,$$

$$\eta_2 = \eta_{,xx} + (2\eta_{,xy} - \xi_{,xx})p + (\eta_{,yy} - 2\xi_{,xy})p^2 - \xi_{,yy}\,p^3 + (\eta_{,y} - 2\xi_{,x} - 3p\xi_{,y})q.$$

I am using a standard notation for the partial derivatives such that subscripts after a comma indicate differentiation, so for instance

$$\xi_{,xy} = \frac{\partial^2\xi}{\partial y\partial x}, \qquad \xi_{,xx} = \frac{\partial^2\xi}{\partial x^2}, \text{ etc.}$$

One more comment on prolongations is the following. Each one-parameter group has sets of canonical coordinates in which the action of the generator is just $\bar{U}s = 1$ and $\bar{U}t^k = 0$, and therefore when expressed in these coordinates the prolongation of $\bar{U}$ to any order is just

$$\bar{U} = \partial/\partial s,$$

so that in canonical variables $\xi = 1$; $\eta = 0$ and all the other η_k coefficients are zero also. All additional terms in the prolongations vanish and the prolongation of U to any order is only $\partial/\partial s$.

A *criterion* that a DE admits a group can be written now, starting from the DE

$$\psi(x, y, p, q, \ldots) = 0.$$

Thus, if

$$g_\alpha \psi(x, y, p, q, \ldots) = \psi(\tilde{x},\tilde{y},\tilde{p},\tilde{q},\ldots) = 0 \quad \text{modulo} \quad \psi = 0$$

then g_α is a symmetry of the DE, or the DE admits the group. Differentiating with respect to α and setting $\alpha = 0$,

$$U\psi(x, y, p, q, \ldots) = \xi\frac{\partial\psi}{\partial x} + \eta\frac{\partial\psi}{\partial y} + \eta_1\frac{\partial\psi}{\partial p} + \eta_2\frac{\partial\psi}{\partial q} + \cdots = 0, \qquad \text{modulo} \quad \psi = 0.$$

The equation has to hold as an identity in all the variables x, y, p, …. Practically, to evaluate modulo $\psi = 0$ really means to use the equation $\psi = 0$ to eliminate the highest derivative. So $U\psi = 0$ mod. $\psi = 0$ is a necessary circumstance when the DE admits the transformation generated by U, or in other words when U generates a symmetry of $\psi = 0$. It is also sufficient, but only after one detail is taken into account.

Consider the following difficulty. If a DE is represented by

$$\psi(x, y, p, q, \ldots) = 0,$$

then the same DE can be represented by

$$\chi(x, y, p, q, \ldots) = (\psi(x, y, p, q, \ldots))^2 = 0.$$

However, when the DE is represented in this way, since U is a differential operator,

$$U\chi = 2\psi U\psi.$$

The latter expression is always zero modulo ψ, whether U is a symmetry or not. This comes about due to the extra factor of ψ. In fact each partial derivative of χ vanishes mod ψ. So to avoid this problem, one must ensure that ψ *is chosen in such a way that not all of its partial derivatives vanish modulo* ψ.

Example. Consider the DE

$$2x^2p^2 - p^2 - 4xyp + 2y^2 - 1 = 0.$$

This admits rotation. To see that it does, the rotation generator

$$U_{\text{rot}} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}$$

can be extended to operate on functions of x, y, and also p. Accordingly the first prolongation is

$$U^{(1)}_{\text{rot}} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} + (1 + p^2)\frac{\partial}{\partial p}.$$

If I operate on the left-hand side of the DE,

$$U^{(1)}_{\text{rot}}(2x^2p^2 - p^2 - 4xyp + 2y^2 - 1) = 2p(2x^2p^2 - p^2 - 4xyp + 2y^2 - 1).$$

Since the original expression appears as a factor on the right-hand side it is clear that the result is zero modulo the DE. So rotation is a symmetry.

3.2 Reformulation of symmetry in terms of partial differential operators

Since the group generator U is a partial differential operator acting on functions of the coordinates $x = \{x^i\}$ with $i \in \{1, 2, \ldots, n\}$, a more uniform treatment of symmetry is possible if the DEs are also written in terms of partial differential operators.

It is always possible to change from $\{x^1, x^2, \ldots, x^n\}$ to canonical coordinates $\{s, t^1, t^2, \ldots, t^{n-1}\}$ where $Us = 1$, $Ut^i = 0$ and so U is simply $\partial/\partial s$. The canonical coordinates are not unique, as has been noted already. Given s and $t = \{t^k\}$, there are infinitely many other sets,

$$\bar{t}^k = f^k(t), \qquad \bar{s} = s + h(t),$$

that will work as well, where $f^k(t)$ and $h(t)$ are smooth functions of the t coordinates. It was noticed above that in canonical coordinates every prolongation of U is just $U = \partial/\partial s$.

To illustrate the operator form, consider for example a first order DE, $\psi(x, y, p) = 0$. Suppose it is solved algebraically for p, the derivative. Then

$$p = \frac{dy}{dx} = w(x, y).$$

Assume too that the primitive $\phi(x, y, c) = 0$ can be solved for the parameter c,

$$c = u(x, y).$$

From the DE one has

$$w(x, y)dx - dy = 0.$$

From the solution

$$dc = 0 = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy.$$

In matrix form, this is

$$\begin{pmatrix} w(x, y) & -1 \\ \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The matrix must be singular, so

$$\det\begin{pmatrix} w(x, y) & -1 \\ \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \end{pmatrix} = w(x, y)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial x} = 0.$$

The DE is indeed equivalent to the PDE

$$Au(x, y) = 0, \qquad \text{where} \quad A = \frac{\partial}{\partial x} + w(x, y)\frac{\partial}{\partial y}.$$

The vanishing of the determinant indicates that the differentials dx and dy are not independent, since given dx, dy must be such as to satisfy the DE or to ensure the solution function $u(x, y)$ remains at the same value c. To appreciate the meaning of this, it is instructive perhaps to work back again from the PDE to the original DE.

If u were any function of x and y and dx and dy any small displacements,

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy.$$

But if the displacement needs to be such as to keep $u(x, y) = c$, then dx and dy are no longer independent. At a particular point (x, y) the partial derivatives become components of a vector that must be perpendicular to (dx, dy),

$$\frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy = 0 \qquad \Rightarrow \qquad \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \cdot (dx, dy) = 0,$$

but the values of these coefficients evaluated at (x,y) need to satisfy the PDE, so

$$w(x, y)\frac{\partial u}{\partial y} + \frac{\partial u}{\partial x} = 0 \qquad \Rightarrow \qquad \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \cdot (1, w(x, y)) = 0.$$

In this two-dimensional space, it means the vectors (dx, dy) and $(1, w(x, y))$ must be parallel, so one arrives at the usual characteristic system,

$$\frac{dx}{1} = \frac{dy}{w(x, y)}, \qquad \Rightarrow \qquad p = \frac{dy}{dx} = w(x, y),$$

which is the original DE.

Now extend the development in terms of partial differential operators to the general DE of order n. In terms of the original coordinates x and y the complete primitive,

$$\phi(x, y, a, b, \ldots) = 0,$$

determines a DE for which it is the general solution,

$$\psi(x, y, p, q, \ldots) = 0.$$

Suppose the DE is given. Now if the DE can be solved algebraically for its highest order derivative $y^{(n)}$,

$$y^{(n)} = w(x, y, p, q, \ldots),$$

then it can be rewritten in terms of the differential operator

$$A = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + q\frac{\partial}{\partial p} + \cdots + w(x, y, p, q, \ldots)\frac{\partial}{\partial y^{(n-1)}}.$$

Notice that $A = d/dx$ modulo the DE so that when applied to functions that satisfy the DE, $e^{\alpha A}$ is a translation along a solution curve. A *first integral* of the DE is a function $u_k(x, y, p, \ldots, y^{(n-1)})$ that is constant $u_k = c_k$ along a solution curve. So

$$e^{\alpha A}u_k(x, y, p, \ldots, y^{(n-1)}) = u_k(x, y, p, \ldots, y^{(n-1)}).$$

This means

$$Au_k(x, y, p, \ldots, y^{(n-1)}) = 0.$$

Therefore, a first integral is annihilated by A and has a constant value along any solution curve. With a complete set of n first integrals, one can in principle solve the equations

$$u_k(x, y, p, \ldots, y^{(n-1)}) = c_k \qquad \text{for} \quad k \in \{1, 2, \ldots, n\}$$

that hold along a solution curve to obtain

$$y = v_1(c_1, c_2, \ldots, x),\ p = v_2(c_1, c_2, \ldots, x),\ y^{(k)}(x) = v_{k+1}(c_1, c_2, \ldots, x).$$

Notice the constants for any particular solution curve are evaluated from initial conditions at any point on the curve,

$$c_k = u_k\left(x_o, y_o, p_o, \ldots, y_o^{(n-1)}\right) \qquad \text{for} \quad k \in \{1, 2, \ldots, n\}.$$

A complete set of first integrals for the original DE is equivalent to its general solution, or to a complete set of solutions to $Au = 0$.

Example. Here is an example of a second order DE:

$$\psi(x, y, p, q) = 2y(x^2 + y^2)q + 2xyp - 2(2x^2 - 3y^2)p^2 - 6xyp^3 + 2y^2 = 0.$$

The equation is complicated, but the primitive it derives from is relatively simple, namely

$$\phi(x, y, a, b) = ax + by(x^2 + y^2) + y = 0.$$

Solving this first for b and then for a, and differentiating each, produces two first integrals,

$$u_1(x, y, p) = \frac{2y^2(yp + x)}{y(y^2 - x^2) - (x^3 + 3xy^2)p} = a,$$

$$u_2(x, y, p) = \frac{y - xp}{x^3p + 3xy^2p + x^2y - y^3} = b.$$

By eliminating p between $u_1 = a$ and $u_2 = b$ one arrives back at the general solution. On the other hand, the partial differential operator found from solving the DE for q is

$$A = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + \left(\frac{(xp - y)(3yp^2 + 2xp + y)}{y(x^2 + y^2)}\right)\frac{\partial}{\partial p}.$$

This operator annihilates both u_1 and u_2. This concludes the example.

Finally, notice in general that the characteristic system for the PDE

$$\frac{\partial \phi}{\partial x} + p\frac{\partial \phi}{\partial y} + q\frac{\partial \phi}{\partial p} + \cdots + w(x, y, p, q, \ldots, y^{(n-1)})\frac{\partial \phi}{\partial y^{(n-1)}} = 0,$$

which is

$$\frac{dx}{1} = \frac{dy}{p} = \frac{dp}{q} = \cdots = \frac{dy^{(n-1)}}{w(x, y, \ldots)},$$

so that for $0 < k < n - 2$,

$$\frac{dx}{1} = \frac{dy^{(k)}}{y^{(k+1)}}, \qquad \Rightarrow \qquad \frac{dy^{(k)}}{dx} = y^{(k+1)},$$

is a system of ordinary DEs equivalent to the original one,

$$p = \frac{dy}{dx},\ q = \frac{dp}{dx},\ \cdots \frac{dy^{(n-1)}}{dx} = w\left(x, y, \ldots, y^{(n-1)}\right).$$

So to summarize, if the original nth order ordinary DE can be solved algebraically for its highest derivative, then

$$y^{(n)} - w(x, y, p, q, \ldots, y^{(n-1)}) = 0$$

is equivalent to

$$Au = 0,$$

where

$$A = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + q\frac{\partial}{\partial p} + \cdots + w(x, y, p, \ldots, y^{(n-1)})\frac{\partial}{\partial y^{(n-1)}}.$$

Integrating either one leads to an equivalent form of the n-parameter family of functions defined by

$$\phi(x, y, c_1, \ldots, c_n) = 0,$$

which is the general solution of $y^{(n)} = w(x, y, p, q, \ldots)$.

3.3 Symmetries in terms of A

A symmetry of an equation is a transformation that permutes its solutions. A symmetry group permutes solutions of the DE, which permutes members of the family of curves

$$\phi(x, y, c_1, c_2, \ldots, c_n) = 0$$

and so it permutes solutions of the equivalent PDE. The group is a symmetry of the DE if whenever

$$Au(x, y, p, \ldots) = 0$$

then for each transformation g_α in the group,

$$Ag_\alpha u(x, y, p, \ldots) = 0$$

also. Thus $Ag_\alpha u$ must vanish modulo Au. The group takes u, which must be some function of the complete set $u = \{u_k\}$ of first integrals, into another function of the first integrals. In particular

$$Uu_j = H_j(\{u_k\}) = H_j(u).$$

In order to eliminate reference to the unknown functions, it suffices to form the commutator

$$[U, A] \equiv UA - AU.$$

Then

$$[U, A]u = 0,$$

if u is any solution of $Au = 0$ and U generates a symmetry. Since A and $[U, A]$ annihilate the same set of functions there will exist some function $\lambda(x, y, p, \ldots, y^{(n-1)})$ such that

$$[U, A] = \lambda(x, y, p, \ldots, y^{(n-1)})A.$$

This is the second form of the symmetry condition.

Example. The DE of a previous example, namely

$$2x^2p^2 - p^2 - 4xyp + 2y^2 - 1 = 0,$$

is quadratic in p and so solving for p there are two roots. These two formulas give two different slope fields, and really two different DEs, each of which has a unique solution through a given point, provided the expression for p evaluates to a real number. I choose to study one of the roots. So consider the following DE deriving from a single root or factor in this example,

$$p = w(x, y) = \frac{2xy + \sqrt{2x^2 + 2y^2 - 1}}{2x^2 - 1}.$$

Then the partial differential operator is

$$A = \frac{\partial}{\partial x} + \left(\frac{2xy + \sqrt{2x^2 + 2y^2 - 1}}{2x^2 - 1} \right) \frac{\partial}{\partial y}.$$

The commutator of this with the rotation generator gives

$$\begin{aligned}
[U_{\text{rot}}, A] = & \left(\frac{\sqrt{2x^2 + 2y^2 - 1} + 2xy}{2x^2 - 1} \right) \frac{\partial}{\partial x} \\
& + \left(\frac{2x\left(2y\sqrt{2x^2 + 2y^2 - 1} + 2xy^2 + x\right) + 2y^2 - 1}{(1 - 2x^2)^2} \right) \frac{\partial}{\partial y} \\
= & \left(\frac{\sqrt{2x^2 + 2y^2 - 1} + 2xy}{2x^2 - 1} \right) A.
\end{aligned}$$

The occurrence of A on the right-hand side shows again that rotation is a symmetry of the DE, or that this factor of the original DE also admits rotation. It illustrates the second method of checking to see that a particular group generated by U is a symmetry of a DE corresponding to a given operator A.

3.4 Note on evaluating commutators

It is useful to derive some general results about commutators in order to facilitate calculation. A commutator is antisymmetric, apparently, meaning that $[B, A] = -[A, B]$. If a and b are constants,

$$[aA + bB, C] = a[A, C] + b[B, C].$$

When a and b are not constant things are more complicated. A basic property of the commutator $[A, BC]$ involving any three operators follows immediately from the definition, namely

$$[A, BC] = B[A, C] + [A, B]C.$$

Thus in some ways, applying $[A, \ldots]$, or in other words taking a commutator of some expression with respect to A, is like differentiating with respect to A. Thus the identity is a sort of Leibnitz rule for differentiating a product of operators. So I will call it the derivation property of the commutator. Similarly,

$$[AB, C] = A[B, C] + [A, C]B.$$

The partial differential operators of interest in this treatment of symmetry are of the general form

$$A = a^j(x)\partial_j, \qquad B = b^j(x)\partial_j \qquad \text{etc},$$

where $x = \{x^1, x^2, \ldots, x^n\}$ is some set of independent coordinates,

$$\partial_j = \frac{\partial}{\partial x^j}$$

and I intend to sum over the repeated index. Thus to find an explicit formula for commutators in this case, compute using the derivation property

$$\begin{aligned}
[A, B] &= [a^j(x)\partial_j, b^k(x)\partial_k] \\
&= a^j(x)[\partial_j, b^k(x)\partial_k] + [a^j(x), b^k(x)\partial_k]\partial_j \\
&= a^j(x)\Big(b^k(x)[\partial_j, \partial_k] + [\partial_j, b^k(x)]\partial_k\Big) \\
&\quad + \Big(b^k(x)[a^j(x), \partial_k] + [a^j(x), b^k(x)]\partial_k\Big)\partial_j.
\end{aligned}$$

In the first term, $[\partial_j, \partial_k]$ will vanish, assuming the functions are such that the cross partial derivatives are equal. In the last term, $[a^j(x), b^k(x)] = 0$

$$[A, B] = a^j(x)\Big([\partial_j, b^k(x)]\partial_k\Big) + \Big(b^k(x)[a^j(x), \partial_k]\Big)\partial_j.$$

Then notice that, if $f(x)$ is any function and $F(x)$ is any test function,

$$\begin{aligned}
[\partial_k, f(x)]F(x) &= \partial_k(f(x)F(x)) - f(x)\partial_k F(x) \\
&= (\partial_k f(x))F(x) = f_{,k}\ (x)F(x).
\end{aligned}$$

Thus, thinking of $f(x)$ as the operator, the recipe for which is to 'multiply from the left by $f(x)$',

$$[\partial_k, f(x)] = (\partial_k f(x)).$$

So we have

$$\begin{aligned}
[A, B] &= \Big(a^j(x)\Big(\partial_j b^k(x)\Big)\Big)\partial_k - (b^k(x)(\partial_k a^j(x)))\partial_j, \\
&= \Big(a^j(x)\Big(\partial_j b^k(x)\Big) - b^j(x)\Big(\partial_j a^k(x)\Big)\Big)\partial_k.
\end{aligned}$$

This can be written briefly as

$$[A, B] = \Big(\Big(Ab^k(x)\Big) - \Big(Ba^k(x)\Big)\Big)\partial_k.$$

It is the general result for coefficients of a commutator.

Now for applying the general result to the symmetry condition $[U, A] = \lambda A$, where

$$U = \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \eta_1\frac{\partial}{\partial p} + \cdots$$
$$A = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + q\frac{\partial}{\partial p} + \cdots + w(x, y, p, \ldots)\frac{\partial}{\partial y^{(n-1)}},$$

the first step is to form the commutator $[U, A]$. So introducing coefficients

$$[U, A] = C^k\partial_k,$$

one finds

$$C^1 = (U1) - (A\xi),\ C^2 = (Up) - (A\eta),\ C^3 = (Uq) - (A\eta_1),\ \text{etc.}$$

It should be kept in mind that $A = d/dx$ modulo $y^{(n)} = w$. Thus in terms of these coefficients, the symmetry condition $[U, A] = \lambda A$ is

$$C^1\partial_x + C^2\partial_y + C^3\partial_p + \cdots + C^k\partial_{y^{(k-2)}}\cdots = \lambda\left(\partial_x + p\partial_y + q\partial_p + \cdots\right),$$

which yields

$$C^1 = -\frac{d\xi}{dx} = \lambda.$$

This defines λ. Then

$$C^2 = \eta_1 - \frac{d\eta}{dx} = -\frac{d\xi}{dx}p \qquad \text{or} \qquad \eta_1 = \frac{d\eta}{dx} - p\frac{d\xi}{dx},$$
$$C^3 = \eta_2 - \frac{d\eta_1}{dx} = -\frac{d\xi}{dx}q \qquad \text{or} \qquad \eta_2 = \frac{d\eta_1}{dx} - q\frac{d\xi}{dx},$$
$$\cdots$$
$$C^{k+1} = \eta_k - \frac{d\eta_{k-1}}{dx} = -\frac{d\xi}{dx}y^{(k)} \qquad \text{or} \qquad \eta_k = \frac{d\eta_{k-1}}{dx} - y^{(k)}\frac{d\xi}{dx},$$

and so on. These equations are tautologies in that they reiterate the definitions of the functions η_n. The commutator symmetry criterion in fact provides an alternative way to define the prolongations. This is true, up to the point where

$$C^{n+1} = \eta_n - \frac{d\eta_{n-1}}{dx} = -\frac{d\xi}{dx}w \qquad \text{modulo} \qquad y^{(n)} = w\left(x, y, p, q, \ldots, y^{(n-1)}\right).$$

This is

$$Uw = \eta_n \qquad \text{modulo} \qquad y^{(n)} = w\left(x, y, p, q, \ldots, y^{(n-1)}\right),$$

the same as would result from the previous symmetry criterion. To wit,

$$U\left(y^{(n)} - w(x, y, p, \ldots)\right) = 0 \quad \text{mod.} \quad y^{(n)} = w.$$

Now given a DE, there are two steps to solving it using groups: find a group that permutes the solutions and then use it to reduce the order of the DE.

3.5 Symmetries of first-order DEs

For the general first order DE,

$$p = \frac{dy}{dx} = w(x, y),$$

the operator

$$A = \frac{\partial}{\partial x} + w(x, y)\frac{\partial}{\partial y}$$

gives the symmetry condition

$$\begin{aligned}[U, A] &= \left[\xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y}, \frac{\partial}{\partial x} + w(x, y)\frac{\partial}{\partial y}\right] \\ &= -\left(w\xi_{,y} + \xi_{,x}\right)\frac{\partial}{\partial x} + \left(w_{,y}\,\eta + w_{,x}\,\xi - w\eta_{,y} - \eta_{,x}\right)\frac{\partial}{\partial y} \\ &= \lambda\frac{\partial}{\partial x} + \lambda w\frac{\partial}{\partial y} = \lambda A.\end{aligned}$$

So, as in general

$$\lambda = -\frac{d\xi}{dx} = -\left(\xi_{,x} + w\xi_{,y}\right),$$

and using this result to eliminate λ from the coefficients of $\partial/\partial y$ gives

$$\xi w_{,x} + \xi_{,x}\,w + \xi_{,y}\,w^2 = \eta_{,x} + \eta_{,y}\,w - \eta w_{,y}\ .$$

So this is the general symmetry condition for a first order DE that can be solved algebraically for p and put into the form $p = dy/dx = w(x, y)$.

The current result is quite general, and there are always infinitely many solutions for ξ and η for any given function $w(x, y)$, so that a first order DE always has infinitely many symmetries. In fact, since $f(x, y)U$ annihilates the same functions as U, it is possible to choose $\xi(x, y)$ and attempt to solve for $\eta(x, y)$ or vice versa. However the frustrating fact, as remarked by Stephani [1] is that one has no definite strategy for finding the solutions for the components of U. It would almost seem that, since one has substituted an ordinary DE for a PDE such that in either case the general solution method is not known, then perhaps not much progress has been made. In fact it often happens that a symmetry is suggested by the physical meaning or by the physical origin of the problem. Knowing from where the DE originally came, i.e. knowing the physical origin of it, is a major hint toward finding a symmetry for it.

The situation is better for DEs of second or higher order where in most cases one will have a definite way to proceed and can often find a group admitted by a DE by means of which the order can be reduced.

3.6 Tabulating DEs according to groups they admit

As noted above, one approach taken by Lie's was to turn the problem around. Arranging groups in order of increasing complexity he began tabulating for each one the most general form of the DEs that admit the group. This process is described in several places, including Cohen's section 18, [2] and Dresner's section 2.6, [3]. A tabulation of results from such a project is found for instance in Cohen [2] section 19 and in the tables I and II in the back of that book.

If a function of x, y, p, and so on is annihilated by the extended group generator U,

$$U\psi(x, y, p, q, \ldots) = \left(\xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \eta_1\frac{\partial}{\partial p} + \cdots\right) f(x, y, p, \ldots) = 0$$

then

$$f(x, y, p, \ldots) = c$$

gives a DE, or a family of DEs, invariant under the group. It is enough to set c to zero, since the constant can be included in the general function f. So far we have used this fact to seek groups admitted by a given DE. Now the idea is to turn the process around. For a given group, one needs to solve the characteristic system,

$$\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)} = \frac{dp}{\eta_1(x, y, p)} = \cdots$$

up to some fixed order.

Example. Consider for instance the rotation group generated by

$$U_{\text{rot}} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

Then either the symmetry condition derived above or more directly $Uw = 0$ with U the extended form of the rotation generator U_{rot} becomes a simple PDE for $w(x, y)$

$$-yw_{,x} + xw_{,y} = 1 + w^2.$$

The characteristic system is

$$\frac{dx}{-y} = \frac{dy}{x} = \frac{dw}{1 + w^2}.$$

The first equation gives the characteristic

$$u(x, y) = x^2 + y^2$$

that is annihilated by U_{rot} and is rotation invariant, and the second gives

$$\pm\int \frac{dx}{\sqrt{u - x^2}} = c(u) + \arctan w,$$

where u is treated as constant in the integration, so

$$\pm \arctan\left(\frac{x}{y}\right) = c(x^2 + y^2) + \arctan w,$$

where c is an arbitrary function. Let

$$c(x^2 + y^2) = -\arctan\left(f(x^2 + y^2)\right).$$

Then

$$\arctan w = \arctan\left(f(x^2 + y^2)\right) \pm \arctan\left(\frac{x}{y}\right),$$

$$w(x, y) = \frac{f(x^2 + y^2) \pm x/y}{1 \mp (y/x) f(x^2 + y^2)}.$$

The most general first order DE invariant under rotation is thus

$$\frac{yp - x}{xp + y} = f(x^2 + y^2),$$

where f is an arbitrary function.

In this way, one can develop a table of one-parameter groups with generators of increasingly complicated form, showing which sorts of DEs admit each group. This was done by Lie. The problem of finding a group admitted by the DE arising in a certain problem is best solved by deciding what symmetries might exist for physical reasons, and then consulting the table. The next question concerns how to use a symmetry, once it is found, to integrate the equations that admit it as a symmetry.

3.7 Lie's integrating factor

Perhaps the most famous result of the theory is Lie's method of obtaining an integrating factor for a first order DE that admits a group. This topic is presented in some detail in chapter II of Cohen's book [2], especially section 12. The subject is also taken up at length by Page [4] in chapters IV and V. Dresner introduces the idea in the following way in section 2.1 of [3]. Suppose the general solution is

$$\phi(x, y, a) = 0.$$

Then, suppose I can solve this for a,

$$\varphi(x, y) = a,$$
$$d\varphi = \varphi_{,x}\ dx + \varphi_{,y}\ dy = 0.$$

This is a first order DE that is very easy to solve because of course it is an exact differential. However, it would typically arise in the form,

$$p - w(x, y) = 0 \qquad \Rightarrow \qquad N(x, y)dx - M(x, y)dy = 0,$$

where $w = N/M$ and

$$d\varphi(x, y) = \mu(x, y)(N(x, y)dx - M(x, y)dy),$$

and $\mu(x, y)$ is some integrating factor. Without knowing $\mu(x, y)$, since the expression $N\,dx - M\,dy$ is not an exact differential, the problem is not as easy to solve. So one way to proceed is to seek an integrating factor first, and then to use the integrating factor to integrate the DE. Lie was able to show that if the DE

$$N(x, y)dx - M(x, y)dy = 0$$

admits a group with generator

$$U = \xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y}$$

then an integrating factor is

$$\mu = \frac{1}{N\xi - M\eta}.$$

Here is how it works.

If U generates a symmetry of the DE, then $g_\alpha\varphi = c'$ whenever $\varphi = c$, so that, as was seen previously,

$$U\varphi = \xi\varphi_{,x} + \eta\varphi_{,y} = f(\varphi).$$

Also, rewriting the DE as a PDE,

$$A\varphi = M\varphi_{,x} + N\varphi_{,y} = 0.$$

So,

$$\begin{pmatrix} \xi & \eta \\ M & N \end{pmatrix}\begin{pmatrix} \varphi_{,x} \\ \varphi_{,y} \end{pmatrix} = \begin{pmatrix} f(\varphi) \\ 0 \end{pmatrix}.$$

Inverting,

$$\begin{pmatrix} \varphi_{,x} \\ \varphi_{,y} \end{pmatrix} = \frac{1}{N\xi - M\eta}\begin{pmatrix} N & -\eta \\ -M & \xi \end{pmatrix}\begin{pmatrix} f(\varphi) \\ 0 \end{pmatrix}.$$

Then, since

$$d\varphi = \begin{pmatrix} dx & dy \end{pmatrix}\begin{pmatrix} \varphi_{,x} \\ \varphi_{,y} \end{pmatrix} = \varphi_{,x}\,dx + \varphi_{,y}\,dy,$$

one has

$$d\varphi = \frac{f(\varphi)}{N\xi - M\eta}(Ndx - Mdy).$$

The presence of the factor $f(\varphi)$ seems at first bothersome, but keep in mind that a solution of the DE is of the form $\chi(x, y) = c$ for some constant, and that the function χ need not be φ but can be any composite function $g(\varphi)$. Thus, let

$$g(t) = \int \frac{dt}{f(t)}.$$

So the solution is $g(\varphi(x, y)) = c$ where

$$dg(\varphi) = \frac{d\varphi}{f(\varphi)} = \frac{1}{M\eta - N\xi}(Ndx - Mdy).$$

When a symmetry is known, the components $\xi(x, y)$ and $\eta(x, y)$ provide an integrating factor,

$$\mu(x, y) = \frac{1}{M\eta - N\xi}.$$

Example. Consider again the DE

$$p = w(x, y) = \frac{2xy + \sqrt{2x^2 + 2y^2 - 1}}{2x^2 - 1},$$

which was seen to admit rotations generated by

$$U = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y},$$

where $\xi = -y$ and $\eta = x$. So the first order DE

$$\frac{\left(\sqrt{2x^2 + 2y^2 - 1} + 2xy\right)}{2x^2 - 1}dx - dy = 0$$

should have as an integrating factor

$$\mu(x, y) = \left(\frac{y\left(2xy + \sqrt{2x^2 + 2y^2 - 1}\right)}{2x^2 - 1} + x\right)^{-1}.$$

Thus, $\varphi = c$ is the general solution, where

$$\varphi_{,x} = \frac{\sqrt{2x^2 + 2y^2 - 1} + 2xy}{(2x^2 - 1)\left(\frac{y\left(\sqrt{2x^2 + 2y^2 - 1} + 2xy\right)}{2x^2 - 1} + x\right)},$$

$$\varphi_{,y} = -\frac{1}{\frac{y\left(\sqrt{2x^2 + 2y^2 - 1} + 2xy\right)}{2x^2 - 1} + x}.$$

One may check that

$$\varphi_{,xy} = \varphi_{,yx}$$

is satisfied. A complete primitive in simplified form is found to be

$$a^2 + 1 - (1 - a^2 - 2a)x - (1 - a^2 + 2a)y = 0,$$

but rotation symmetry is most easily seen by letting $a = \tan\theta/2$

$$(x + y)\cos\theta - (x - y)\sin\theta = 1,$$

where here θ is the solution parameter.

3.8 Finding symmetries of a second order DE

The second order DEs are of particular interest in physics. In this section I present a general algorithm for finding symmetries of a second order DE, and then illustrate the method by carrying out the procedure, in rather a lot of detail, for a particular second order equation as an example. The method is presented in the book by Stephani [1], where a different, somewhat shorter example is worked out in some detail. In order to show an example that is a little bit more complex, and thus illustrate a few additional points, Mathematica [5] has been used to perform or to check some of the calculations in this and the previous section.

First I summarize the results of the previous section for a DE of any order. The DE corresponding to the operator

$$A = \frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + \sum_{k=1}^{n-2} y^{(k+1)}\frac{\partial}{\partial y^{(k)}} + w\frac{\partial}{\partial y^{(n-1)}},$$

and a candidate symmetry group generated by

$$U = \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \sum_{k=1}^{n-2} \eta_k\frac{\partial}{\partial y^{(k)}} + \eta_{n-1}\frac{\partial}{\partial y^{(n-1)}},$$

should be such that

$$[U, A] = \lambda A,$$

if indeed U is to generate a symmetry of the DE represented by A. Thus

$$\begin{aligned}(0 - A(\xi))\frac{\partial}{\partial x} + (U(p) - A(\eta))\frac{\partial}{\partial y} + \sum_{k=1}^{n-2}\Big(U(y^{(k+1)}) - A(\eta_k)\Big)\frac{\partial}{\partial y^{(k)}} \\ + \Big(U(w) - A(\eta_{n-1})\Big)\frac{\partial}{\partial y^{(n-1)}} \\ = \lambda\left(\frac{\partial}{\partial x} + p\frac{\partial}{\partial y} + \sum_{k=1}^{n-2} y^{(k+1)}\frac{\partial}{\partial y^{(k)}} + w\frac{\partial}{\partial y^{(n-1)}}\right).\end{aligned}$$

Since η_k depends on $x, y, p, \ldots, y^{(k)}$, and since A is d/dx modulo $y^{(n)} = w$, by comparing the coefficients of $\partial/\partial x$ it turns out quite generally that

$$\lambda = -\frac{d}{dx}\xi = -\big(\xi_{,x} + p\xi_{,y}\big),$$

and also that

$$A(\eta_k) = \frac{d}{dx}\eta_k, \qquad k = 1, 2, \ldots, n-1.$$

Further, since by definition

$$\eta_{k+1} = \frac{d}{dx}\eta_k - y^{(k+1)}\frac{d}{dx}\xi,$$

equating coefficients for most of the partial derivatives produces an identity. To wit, it results in

$$U(y^{(k+1)}) - A(\eta_k) = \eta_{k+1} - \frac{d}{dx}\eta_k = -y^{(k+1)}\frac{d}{dx}\xi.$$

Then, finally, the non-identical part of the criterion is

$$U(w) - A(\eta_{n-1}) = -\big(\xi_{,x} + p\xi_{,y}\big)w.$$

The latter could be written also as

$$U(w) = \eta_n \qquad \text{modulo } y^{(n)} - w = 0.$$

Case $n = 2$, second order DE. When $n = 2$,

$$\eta_1 = \frac{d}{dx}\eta - p\frac{d}{dx}\xi = \eta_{,x} + \big(\eta_{,y} - \xi_{,x}\big)p - \xi_{,y}\,p^2,$$

and then

$$\eta_2 = \frac{d}{dx}\eta_1 - q\frac{d}{dx}\xi,$$

yields

$$\eta_2 = \eta_{,xx} + \big(2\eta_{,xy} - \xi_{,xx}\big)p + \big(\eta_{,yy} - 2\xi_{,xy}\big)p^2 - \xi_{,yy}\,p^3 + \big(\eta_{,y} - 2\xi_{,x}\big)q - 3\xi_{,y}\,pq.$$

Then

$$U(w) = \left(\xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \eta_1\frac{\partial}{\partial p}\right)w(x, y, p),$$

and so for the symmetry condition in the second order case one has

$$\begin{aligned}\xi w_{,x} + \eta w_{,y} + \eta_1 w_{,p} &= \xi w_{,x} + \eta w_{,y} + \Big(\eta_{,x} + \big(\eta_{,y} - \xi_{,x}\big)p - \xi_{,y}\,p^2\Big)\\ w_{,p} &= \eta_{,xx} + \big(2\eta_{,xy} - \xi_{,xx}\big)p + \big(\eta_{,yy} - 2\xi_{,xy}\big)p^2\\ &\quad - \xi_{,yy}\,p^3 + \big(\eta_{,y} - 2\xi_{,x}\big)w - 3\xi_{,y}\,pw.\end{aligned}$$

This is the general equation from which to start looking for symmetries of a given second order DE. The reader should remember that for any random DE of second order there may not be any symmetries at all. However, there often are, especially since the physical origin of the equations should have symmetries buried in it, one usually assumes. But to see how this result is really used, an example is quite important. I have chosen an example here that is not trivial in order to encounter a few of the details that often come up.

Example. For an example, consider the second-order DE,

$$y(x-y)q + 2xp^3 - 2(x+y)p^2 + 2yp = 0.$$

From this,

$$w = \frac{2(1-p)p(xp-y)}{y(x-y)}.$$

With this substitution, the symmetry condition becomes a bit of a mess:

$$\begin{aligned}
&p^3\xi_{,yy} - p^2\big(\eta_{,yy} - 2\xi_{,xy}\big) + \left(\frac{2(p-1)p(xp-y)}{y(x-y)^2} - \frac{2(p-1)p^2}{y(x-y)}\right)\xi \\
&+ \left(\frac{2(p-1)p(xp-y)}{y^2(x-y)} - \frac{2(p-1)p(xp-y)}{y(x-y)^2} + \frac{2(p-1)p}{y(x-y)}\right)\eta \\
&+ \frac{2(1-p)p(xp-y)\big(3p\xi_{,y} - \eta_{,y} + 2\xi_{,x}\big)}{y(x-y)} - p\big(2\eta_{,xy} - \xi_{,xx}\big) \\
&- \frac{2}{y(x-y)}\big(3xp^2 - 2(x+y)p + y\big)\big(p\eta_{,y} - p\big(p\xi_{,y} + \xi_{,x}\big) + \eta_{,x}\big) - \eta_{,xx} = 0.
\end{aligned}$$

The symmetry relation is to be solved for $\xi(x, y)$ and $\eta(x, y)$. Notice that in terms of the variable p it is a polynomial equation. It has to be satisfied identically in x, y, and p. But since $\xi(x, y)$ and $\eta(x, y)$ do not depend on p, then the coefficient of each different power of p has to vanish separately, so the relation breaks down into several separate DEs that have to be satisfied simultaneously. To wit, the coefficient equations arising from p^3, p^2, p, and 1, respectively, are:

$$\begin{gathered}
2\eta_{,x} + (x-y)\eta_{,xx} = 0, \\
2x^2y\eta_{,xy} - 4x^2\eta_{,x} - x^2y\xi_{,xx} + 2y^3\eta_{,xy} - y^3\xi_{,xx} + 4y^2\eta_{,x} \\
- 4xy^2\eta_{,xy} - 2y^2\xi_{,x} + 2xy^2\xi_{,xx} + 2y\eta + 2xy\xi_{,x} - 2y\xi = 0, \\
x^2y^2\eta_{,yy} - 2x^2y^2\xi_{,xy} - 2x^2y\eta_{,y} + 6x^2y\eta_{,x} + 2x^2\eta + y^4\eta_{,yy} \\
- 2y^4\xi_{,xy} + 2y^3\eta_{,y} - 2xy^3\eta_{,yy} - 4y^3\xi_{,y} + 4xy^3\xi_{,xy} \\
- 6xy^2\eta_{,x} - 2y^2\eta + 4xy^2\xi_{,y} + 4y^2\xi - 4xy\eta = 0, \\
x^2y^2\xi_{,yy} - 4x^2y\eta_{,y} + 2x^2\eta + 2x^2y\xi_{,y} + 2x^2y\xi_{,x} \\
+ y^4\xi_{,yy} - 2y^3\xi_{,y} - 2xy^3\xi_{,yy} + 4xy^2\eta_{,y} - 2xy^2\xi_{,x} \\
+ 2y^2\xi - 4xy\eta = 0.
\end{gathered}$$

These four PDEs seem formidable, however they solve systematically. The first one, which corresponds to p^3,

$$2\eta_{,x} + (x - y)\eta_{,xx} = 0,$$

gives

$$\eta(x, y) = g(y) - \frac{f(y)}{x - y},$$

where at this point $f(y)$ and $g(y)$ are arbitrary functions. Making this substitution into the second PDE, factoring, and taking the numerator, yields

$$\begin{gathered} x^3y\xi_{,xx} - 3x^2y^2\xi_{,xx} - 2x^2y\xi_{,x} - y^4\xi_{,xx} - 2y^3\xi_{,x} + 3xy^3\xi_{,xx} + 4xy^2\xi_{,x} \\ - 2y^2\xi + 2xy\xi - 2xyf'(y) + 2y^2f'(y) + 4xf(y) + 2yf(y) \\ - 2xyg(y) + 2y^2g(y) = 0. \end{gathered}$$

The equation for ξ has separated in the sense that it can be solved treating x as the independent variable and y as a parameter. The solution is

$$\begin{aligned} \xi(x, y) = {} & \frac{xyf'(y) - y^2f'(y) - 2xf(y) + yf(y) + xyg(y) - y^2g(y)}{y(x - y)} \\ & + h(y)(x - y) + k(y)(x - y)^2. \end{aligned}$$

Two more arbitrary functions $h(y)$ and $k(y)$ appear, which will be constrained by the other two PDEs. Making substitutions for both ξ and η in these remaining two, factoring, and taking numerators results in

$$\begin{gathered} 3y^2f''(y) - 6yf'(y) + 6f(y) - y^3g''(y) + 2y^2g'(y) \\ + x\big(y^2g''(y) - 2yg'(y) + 2g(y) + 2y^2h'(y)\big) - 2yg(y) - 2y^3h'(y) = 0, \\ x\Big(y^2f^{(3)}(y) + y^2g''(y) - 2yg'(y) + 2g(y) \\ - 2y^3h''(y) - 2y^2h'(y) + 3y^4k''(y) + 6y^3k'(y)\Big) \\ - y^3f^{(3)}(y) + 3y^2f''(y) - 6yf'(y) + 6f(y) - y^3g''(y) + 2y^2g'(y) - 2yg(y) + y^4h''(y) \\ + x^2y\big(yh''(y) + 2h'(y) - 3y^2k''(y) - 6yk'(y)\big) \\ - y^5k''(y) - 2y^4k'(y) + x^3y\big(yk''(y) + 2k'(y)\big) = 0. \end{gathered}$$

Now, again these two DEs break apart. Since they have to be satisfied identically, and since the argument of each of the unknown functions is y, they are polynomial equations in the parameter x. As was the case with the variable p, the coefficient of each separate power of x must vanish identically. So the last term in the second equation shows that, since the coefficient of x^3 must vanish,

$$yk''(y) + 2k'(y) = 0.$$

Hence,

$$k(y) = \beta - \frac{\alpha}{y}.$$

Then, on substituting for $k(y)$ with this result I see that

$$yh''(y) + 2h'(y) = 0,$$

so that

$$h(y) = \delta - \frac{\gamma}{y}.$$

Continuing on, when I make this substitution, I find that

$$2\gamma + y^2 g''(y) - 2yg'(y) + 2g(y) = 0,$$

which gives

$$g(y) = -\gamma + y^2\epsilon + \delta y.$$

When this substitution is made, the following conditions remain.

$$3y\big(yf''(y) - 2f'(y)\big) + 6f(y) = 0, \qquad f^{(3)}(y) = 0,$$

$$6f(y) - y\Big(6f'(y) + y\big(yf^{(3)}(y) - 3f''(y)\big)\Big) = 0,$$

and the most general simultaneous solution is

$$f(y) = \sigma y + \tau y^2.$$

Putting these results together the general form of the coefficients $\xi(x, y)$ and $\eta(x, y)$ we obtain

$$\xi(x, y) = -\frac{\alpha(x-y)^2}{y} + \beta(x-y)^2 - \frac{\gamma x}{y} + \delta x + y^2\epsilon - \frac{\sigma x}{x-y} - \frac{\tau y^2}{x-y},$$

$$\eta(x, y) = -\gamma - \frac{\tau y^2}{x-y} - \frac{\sigma y}{x-y} + y^2\epsilon + \delta y.$$

Each letter of the set $\{\alpha, \beta, \gamma, \delta, \epsilon, \sigma, \tau\}$ can be regarded as a group parameter, so the result is really a set of seven group generators.

$$U_1 = -\frac{(x-y)^2}{y}\frac{\partial}{\partial x}, \qquad U_2 = (x-y)^2\frac{\partial}{\partial x},$$

$$U_3 = -\frac{x}{y}\frac{\partial}{\partial x} - \frac{\partial}{\partial y}, \qquad U_4 = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y},$$

$$U_5 = y^2\frac{\partial}{\partial x} + y^2\frac{\partial}{\partial y}, \qquad U_6 = -\frac{x}{x-y}\frac{\partial}{\partial x} - \frac{y}{x-y}\frac{\partial}{\partial y},$$

$$U_7 = -\frac{y^2}{x-y}\frac{\partial}{\partial x} - \frac{y^2}{x-y}\frac{\partial}{\partial y}.$$

This concludes the example.

3.9 Using a symmetry to reduce the order

There may be several independent symmetries, typically, and the full symmetry of the DE comprises the multiparameter group in which there is one parameter for each of the generators. The group generated by an individual U_k is a subgroup. It will be of interest to understand how these different subgroups interact, and hence to study the commutator relations between the generators. This leads to the topic of Lie algebras. Given this set of generators, one can choose a subgroup and use it to reduce the order of the DE from n to $n-1$. After so doing, a question that arises naturally is whether or not some of the other symmetries will remain, so that these might be used to reduce the order further. This suggests one possible reason to become interested in the algebra.

For now, I will just show a way to use a symmetry to reduce the order. If there is a DE

$$\psi(x, y, p, q, \ldots, y^{(n)}) = 0$$

of order n, and I know a symmetry of it—let us say this symmetry is generated by U—then recall that it is possible to change coordinates from (x,y) to the canonical coordinates (s,t) of U, defined by

$$Us = 1, \qquad Ut = 0$$

so that

$$U = \frac{\partial}{\partial s}$$

and every prolongation of this generator is really just the same expression; no other terms appear. This is because $\tilde{\xi} = 0$ and $\tilde{\eta} = 1$. Consequently one has that for $k \geqslant 1$

$$\tilde{\eta}_k = \frac{d}{dt}\tilde{\eta}_{k-1} - y^{(k)}\frac{d}{dt}\tilde{\xi} = 0.$$

Consider what happens if I change coordinates form (x, y) to (s, t) with t as the independent variable. The form would be

$$\chi(t, s, u, v, \ldots, s^{(n)}) = 0,$$

where $u = ds/dt$ and $v = du/dt$. Since U generates a symmetry,

$$e^{\alpha U}\chi(t, s, u, v, \ldots, s^{(n)}) = \chi(t, s + \alpha, u, v, \ldots, s^{(n)}) = 0.$$

That means the transformed DE does not contain the dependent variable s. (Observe for instance the case when the parameter $\alpha = -s$.) The result is actually a DE of order $n-1$ where the first derivative of u can be taken as the new dependent variable. One has succeeded in reducing the order.

It may be worthwhile to outline a way to change variables. Suppose one has found the canonical variables

$$s = s(x, y), \qquad t = t(x, y),$$

and inverted

$$x = x(s, t), \qquad y = y(s, t).$$

Then

$$p = \frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{y_{,t} + uy_{,s}}{x_{,t} + ux_{,s}}.$$

Similarly,

$$q = \frac{dp}{dx} = \frac{dp/dt}{dx/dt} = \frac{p_{,t} + up_{,s} + vp_{,u}}{x_{,t} + ux_{,s}},$$

and so on. So substituting in the original nth order DE

$$\psi\left(x, y, p, q, \ldots, y^{(n)}\right) = 0$$

factoring and taking the numerator yields the reduced DE

$$\chi\left(t, u, v, \ldots, u^{(n-1)}\right) = 0,$$

now seen to be independent of u and thus of order $n - 1$, and the relation

$$s(t) = c + \int u(t)dt.$$

To illustrate the reduction of order using symmetry, consider again the previous example. The DE

$$y(x - y)q + 2xp^3 - 2(x + y)p^2 + 2yp = 0$$

admits several transformation groups, some of which are elementary. I see that one of the symmetries is a homogeneous scaling of x and y that might be noticed from inspection of the DE. But for instance, consider

$$U_6 = -\frac{x}{x - y}\frac{\partial}{\partial x} - \frac{y}{x - y}\frac{\partial}{\partial y}.$$

I am choosing a more complicated generator in order to encounter some of the more typical steps in the reduction that become trivial in the simpler cases. To find the canonical coordinates for U_6, solve

$$U_6 t(x, y) = 0.$$

Then it turns out that t can be any function of y/x. Take $t = y/x$. Then

$$U_6 s(x, y) = 1$$

has a particular solution $s = y - x$. So

$$s(x, y) = y - x, \qquad t(x, y) = y/x,$$

and the inverse is

$$x(s, t) = \frac{s}{t - 1}, \qquad y(s, t) = \frac{st}{t - 1}.$$

Using the chain rule

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial s} + v\frac{\partial}{\partial u} + \cdots$$

I construct

$$p(s, t, u) = \frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{y_{,t} + uy_{,s}}{x_{,t} + ux_{,s}} = \frac{s + (1 - t)tu}{s + (1 - t)u},$$

and then

$$\begin{aligned} q(s, t, u, v) &= \frac{dp}{dx} = \frac{dp/dt}{dx/dt} = \frac{p_{,t} + up_{,s} + vp_{,u}}{x_{,t} + ux_{,s}} \\ &= \frac{(1 - t)^3(2u(s(1 - t)v - 2u(s + (1 - t)u)}{(s + (1 - t)u)^3}. \end{aligned}$$

Now I am in a position to substitute these results into the original DE. On simplifying I obtain

$$\frac{s^3(1 - t)^3}{(s + (1 - t)u)^3}(tv + 2u) = 0.$$

It appears that the reduced DE

$$tv + 2u = 0, \qquad \text{or} \qquad t\frac{du}{dt} + 2u = 0$$

is easy to solve in a number of ways, as it is linear, separable, and homogeneous. One should notice that the fact that it is homogeneous, or admits a further scaling symmetry, is reminiscent of the fact that the original DE in x and y also had such a symmetry, to wit U_4. The general solution is

$$u = at^{-2},$$

so

$$s = -\frac{a}{t} + b.$$

In the original variables,

$$y - x + a\frac{x}{y} - b = 0.$$

This is the general solution of the original DE

$$y(x - y)q + 2xp^3 - 2(x + y)p^2 + 2yp = 0,$$

which concludes the example.

3.10 Classical mechanics: Nöther's theorem

The example above has illustrated several things, one being that the DE considered admits several symmetries. A natural next step would be to begin an overview of multi-parameter groups and the associated Lie algebras. However this will need to wait for another lecture sequence. Instead, this volume will conclude with an application to classical mechanics, namely Nöther's theorem concerning the connection between symmetries and conserved quantities. A standard textbook for an introduction to classical mechanics is the one by Goldstein [6]. The treatment of Nöther's theorem appearing here follows to some degree that of Desloge and Karch [7].

Some notation in this section will differ from the notation used previously, especially as regards the letters p and q, which in most cases above were second and first derivatives. In this section, time t is the independent variable. The configuration of a mechanical system in space is assumed to be specified completely by a set of generalized coordinates

$$q = \{q_1, q_2, q_3, \ldots, q_m\},$$

where m would be the number of degrees of freedom. A system of two particles, each free to move in three-dimensional space, would require six coordinates, so m would be 6. If the distance between the particles were held fixed, then the number of coordinates would be reduced to $m = 5$ and so on. These generalized coordinates vary in time according to Newton's laws of motion, or equivalently the time evolution is obtained in the following way in the Lagrange formulation of mechanics.

I assume the reader may have experience with classical mechanics, but for completeness I summarize the algorithms without proof. One supposes there is a *Lagrangian function* $L(q, \dot{q}, t)$ that corresponds to the mechanical system, and in fact defines it mathematically. So L is a function of the coordinate set q, the time derivatives

$$\dot{q} = \{\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_m\},$$

where

$$\dot{q}_k = \frac{dq_k}{dt}$$

is a generalized velocity, and also perhaps L is explicitly a function of time t. Then the algorithm that gives the DEs of motion is that for each degree of freedom k,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0.$$

In this way L leads to a set of m second order DEs the simultaneous solution of which gives a complete description of the motion. This set of equations can be deduced from Newtonian mechanics. The quantity $p_k = \partial L/\partial \dot{q}_k$ is the *canonical momentum*. Sometimes it is of the form $p_k = m\dot{q}_k$ and sometimes not.

The reader is apt to recognize also that these DEs are the Euler–Lagrange equations resulting from a variational problem. Consider the set q of coordinates as a point in m-dimensional space that moves as a function of time. So physically $q = q(t)$, meaning that the coordinates are functions of time. Hamilton's variational principle says that the trajectory $q(t)$ from a given $q(t_A)$ to another given $q(t_B)$, along which the system configuration actually evolves, is the one that minimizes the functional

$$A_L[q] = \int_{t_A}^{t_B} L(q, \dot{q}, t)dt$$

with respect to all neighboring trajectories, holding the two end points fixed. Thus the Euler variational equations that result are the Lagrange mechanical equations of motion. So Newtonian mechanics is embedded in a variational formalism. All of this may be very familiar.

For a given mechanical system the Lagrangian function, or *Lagrangian*, is not unique. Two Lagrangians are equivalent if they are related by a *gauge transformation*. In other words, if L_1 and L_2 are related by

$$L_2(q, \dot{q}, t) = L_1(q, \dot{q}, t) + \frac{d}{dt}F(q, t),$$

where F is any differentiable function. To show that these two Lagrangians lead to the same DEs, I just compute

$$\frac{d}{dt}F(q, t) = \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q_j}\dot{q}_j,$$
$$\frac{\partial L_2}{\partial \dot{q}_k} = \frac{\partial L_1}{\partial \dot{q}_k} + \frac{\partial F}{\partial q_k}.$$

Then

$$\frac{d}{dt}\left(\frac{\partial L_2}{\partial \dot{q}_k}\right) = \frac{d}{dt}\left(\frac{\partial L_1}{\partial \dot{q}_k}\right) + \frac{\partial^2 F}{\partial t \partial q_k} + \frac{\partial^2 F}{\partial q_j \partial q_k}\dot{q}_j.$$

On the other hand

$$\frac{\partial L_2}{\partial q_k} = \frac{\partial L_1}{\partial q_k} + \frac{\partial^2 F}{\partial q_k \partial t} + \frac{\partial^2 F}{\partial q_k \partial q_j}\dot{q}_j.$$

Subtracting shows that indeed

$$\frac{d}{dt}\left(\frac{\partial L_2}{\partial \dot{q}_k}\right) - \frac{\partial L_2}{\partial q_k} = \frac{d}{dt}\left(\frac{\partial L_1}{\partial \dot{q}_k}\right) - \frac{\partial L_1}{\partial q_k}.$$

The two Lagrangians are equivalent, since they differ by a total time derivative, because they lead to the same DEs.

Of course, another way to see the equivalence of two Lagrangians that differ only by a total time derivative is straight from Hamilton's principle:

$$A_{L_2}[q] = A_{L_1}[q] + F(q(t_B),\, t_B) - F(q(t_A),\, t_A).$$

The two additional terms do not depend on the trajectory at all, rather only on the end points. So they cannot alter the equations that determine the path.

Quite generally, a symmetry should be a transformation or a set of transformations that leave a figure or an equation somehow the same. A symmetry of a mechanical system should leave the system dynamically the same, meaning that it should leave the DEs of motion the same, and hence it should permute the solutions. It should take one possible mechanical response into another (perhaps the same) possible motion. The DEs must be the same, and so the Lagrangian function must transform into another Lagrangian that is equivalent to it, or in other words, that is related to it by at most a gauge transformation.

So consider a somewhat general group of transformations, not the most general,

$$\begin{aligned} q_k &\to \tilde{q}_k(q,\, t,\, \alpha) = q_k + \alpha\xi_k(q,\, t) + \cdots \\ t &\to \tilde{t}(q,\, t,\, \alpha) = t + \alpha\tau(q,\, t) + \cdots, \end{aligned}$$

which transforms the generalized coordinates and the time. One would need also the first prolongation, so

$$\begin{aligned} \dot{q}_k \to \frac{d\tilde{q}_k}{d\tilde{t}} \equiv \tilde{q}_k' &= \frac{dq_k + \alpha d\xi_k + \cdots}{dt + \alpha d\tau + \cdots}, \\ \tilde{q}_k' &= \dot{q}_k + \alpha\left(\dot{\xi}_k - \dot{\tau}\dot{q}_k\right) + \cdots \\ \tilde{q}_k' &= \dot{q}_k + \alpha\pi_k + \cdots \qquad \text{where} \qquad \pi_k = \dot{\xi}_k - \dot{\tau}\dot{q}_k. \end{aligned}$$

Now I need to do two things. The first is to decide exactly how to determine if this group is a symmetry of the mechanical problem, and the other is to derive the conclusion of the theorem. Roughly, Nöther's theorem is:

To each continuous symmetry there corresponds a conserved quantity.

But in fact, once the symmetry is identified, the theorem tells how to compute the conserved quantity.

So, applying the group to the action functional,

$$A_L[\tilde{q}] = \int_{\tilde{t}_A}^{\tilde{t}_B} L(\tilde{q},\, \tilde{q}',\, \tilde{t})d\tilde{t}.$$

In order to see how to make the action the same before and after, I want to make the integration range the same, namely from t_A to t_B. I assume the time and the transformed time are related monotonically. So I can change the integration variable back from $\tilde{t}$ to t,

$$A_L[\tilde{q}] = \int_{t_A}^{t_B} L(\tilde{q},\, \tilde{q}',\, \tilde{t})\frac{d\tilde{t}}{dt}dt.$$

So now it appears that if the transformation is a symmetry of the mechanical system, the integrand must be a new Lagrangian equivalent at each value of the group parameter α to the old one. So

$$L(\tilde{q},\, \tilde{q}',\, \tilde{t})\frac{d\tilde{t}}{dt} = L(q,\, \dot{q},\, t) + \frac{dF(q,\, t,\, \alpha)}{dt}$$

for all α. It is convenient to write $\tilde{L} = \tilde{L}(q,\, \dot{q},\, t,\, \alpha)$ for $L(\tilde{q},\tilde{q}',\, \tilde{t})$. The group parameter α and the time t are considered to be independent. The expression can be simplified, in the spirit of any typical calculation with Lie groups, by expanding in α, since indeed one assumes only first order knowledge of the transformation. This should be sufficient, if the transformations do indeed form a group, as the group can be reconstructed from knowing only this much. So

$$L + \alpha\frac{d}{d\alpha}\left(\tilde{L}\frac{d\tilde{t}}{dt}\right)\Bigg|_{\alpha=0} + \cdots = L + \alpha\frac{d}{d\alpha}\left(\frac{dF}{dt}\right)\Bigg|_{\alpha=0} + \cdots.$$

I have assumed F goes to zero as α does. The derivatives in the last term are each partial derivatives with respect to α or t, where in each case the other variable is held fixed. Thus

$$\frac{d}{d\alpha}\left(\tilde{L}\frac{d\tilde{t}}{dt}\right)\Bigg|_{\alpha=0} = \frac{d}{d\alpha}\left(\frac{dF}{dt}\right)\Bigg|_{\alpha=0} = \frac{d}{dt}\left(\frac{dF}{d\alpha}\right)\Bigg|_{\alpha=0}.$$

The symmetry condition is that

$$\frac{d}{d\alpha}\left(\tilde{L}\frac{d\tilde{t}}{dt}\right)\Bigg|_{\alpha=0} = \frac{df}{dt}$$

for some arbitrary function $f(q,\, t)$. So, to test whether a particular group of transformations represents a continuous symmetry of the mechanical system, compute the left-hand side and see if the result becomes the time derivative of something. Often this something will turn out to be zero.

So that is the hypothesis of Nöther's theorem. To find that the system admits a continuous symmetry is to find a Lie group such that this condition is satisfied. Now for the conclusion of the theorem one wants to find out what exactly is conserved.

Let me compute the left-land side. It seems this must turn out to be a time derivative also in order to conclude that the time derivative of some quantity should vanish. In other words, in order to find some conserved quantity,

$$\frac{d}{d\alpha}\left(\tilde{L}\frac{d\tilde{t}}{dt}\right) = \left(\frac{\partial\tilde{L}}{\partial\tilde{q}_k}\frac{\partial\tilde{q}_k}{\partial\alpha} + \frac{\partial\tilde{L}}{\partial\tilde{q}_k'}\frac{\partial\tilde{q}_k'}{\partial\alpha} + \frac{\partial\tilde{L}}{\partial\tilde{t}}\frac{\partial\tilde{t}}{\partial\alpha}\right)\frac{d\tilde{t}}{dt} + \tilde{L}\frac{d}{d\alpha}\left(\frac{d\tilde{t}}{dt}\right),$$

and

$$\frac{d\tilde{t}}{dt} = 1 + \alpha\dot{\tau} + \cdots.$$

Then when I assume the hypothesis of the theorem is satisfied for some $f(q, t)$ it means that when $\alpha \to 0$,

$$\left(\frac{\partial L}{\partial q_k}\xi_k + \frac{\partial L}{\partial \dot{q}_k}\pi_k + \frac{\partial L}{\partial t}\tau\right) + \dot{\tau}L = \frac{df}{dt}.$$

From the definition of the canonical momentum $p_k = \partial L/\partial \dot{q}_k$ and then from the Lagrange equations $\dot{p}_k = \partial L/\partial q_k$,

$$\dot{p}_k\xi_k + p_k\left(\dot{\xi}_k - \dot{\tau}\dot{q}_k\right) + \frac{\partial L}{\partial t}\tau + \dot{\tau}L = \frac{df}{dt}.$$

The left-hand side should be the time derivative of something. The last two terms on the left suggest a time derivative of the product τL. I will add and subtract the two missing terms:

$$\dot{p}_k\xi_k + p_k\left(\dot{\xi}_k - \dot{\tau}\dot{q}_k\right) - \dot{p}_k\dot{q}_k\tau - p_k\ddot{q}_k\tau + \left(\frac{\partial L}{\partial t} + \frac{\partial L}{\partial q_k}\dot{q}_k + \frac{\partial L}{\partial \dot{q}_k}\ddot{q}_k\right)\tau + \dot{\tau}L = \frac{df}{dt}.$$

The conclusion is finally

$$\frac{d}{dt}\left(\xi_k p_k - \tau\dot{q}_k p_k + \tau L - f\right) = 0.$$

The quantity

$$c = \xi_k p_k - \tau\dot{q}_k p_k + \tau L - f$$

is conserved, if the group is such that the hypothesis is satisfied.

To say that c is a conserved quantity might obscure the fact that it will sometimes be an explicit function of time. Then the coordinates need to change in such a way that the implicit and explicit time dependences cancel. So another way to describe c is to say it is a first integral of the dynamical equations.

Exercises. To conclude this series of lectures and to illustrate some of the points mentioned above, the following exercises are suggested.

1. Consider a mechanical system comprising a single particle moving in a potential that is rotationally symmetric around the z axis or the q_3 axis. That means the Lagrangian function is completely unchanged when the system is rotated. So the Lagrangian function $L(q_1, q_2, q_3, \dot{q}_1, \dot{q}_2, \dot{q}_3, t)$ is left completely unchanged by the rotations

   $$q_1 \to \tilde{q}_1 = q_1\cos\alpha + q_2\sin\alpha, \qquad q_2 \to \tilde{q}_2 = q_2\cos\alpha - q_1\sin\alpha,$$

   with the time and the other coordinate left invariant. Find the group generator, and then from Nöther's theorem, find the conserved quantity. You should recognize it.

2. Consider another mechanical system in which the Lagrangian does not depend explicitly on time. Thus the transformation

$$t \to \tilde{t} = t + \alpha$$

leaves the Lagrangian completely unchanged. Find the Nötherian conserved quantity. Again you should recognize it.

3. Imagine a charged particle, perhaps an electron, moving in the presence of an electrified, helical wire. Perhaps the wire would be a filament. The surface of the filament is an equipotential surface, so the electric potential in space near the filament has the same helicoidal symmetry. That is to say, the potential energy of the charged particle, and hence also the Lagrangian function, is invariant with respect to the transformations

$$q_1 \to q_1 \cos\alpha + q_2 \sin\alpha,\ q_2 \to q_2 \cos\alpha - q_1 \sin\alpha,\ q_3 \to q_3 + k\alpha.$$

Find the conserved quantity implied by Nöther's theorem. How can you describe this?

4. A particle of mass m is attracted to the center of coordinates by an inverse cube law force. It will remain in a plane as it moves due to conservation of angular momentum. The Lagrangian function in polar coordinates $q_1 = r$ and $q_2 = \theta$ is

$$L = \frac{1}{2}m\left(\dot{q}_1^2 + q_1^2\dot{q}_2^2\right) + \frac{k}{q_1^2}.$$

This problem has a scaling symmetry. Find the relation between α and β such that the scaling transformation

$$q_1 \to \tilde{q}_1 = e^{\alpha}q_1, \qquad q_2 \to q_2, \qquad t \to \tilde{t} = e^{\beta}t$$

leaves the Lagrangian unchanged. Find the conserved quantity c. In this case, although it must be constant in time it is explicitly time dependent.

References

[1] Stephani H 1989 *Differential Equations: Their Solution Using Symmetry* (New York: Cambridge University Press)

[2] Cohen A 1911 *An Introduction to the Lie Theory of One-parameter Groups with Applications to the Solution of Differential Equations* (New York: Heath)

[3] Dresner L 1999 *Applications of Lie's Theory of Ordinary and Partial Differential Equations* (Bristol: IOP Publishing)

[4] Page J M 1897 *Ordinary Differential Equations, with an Introduction to Lie's Theory of the Group of One Parameter* (London: MacMillan)

[5] Wolfram Research 2016 *Mathematica* Version 10.0, Champaign, IL

[6] Goldstein H 1980 *Classical Mechanics* (Reading, MA: Addison-Wesley)

[7] Desloge E A and Karch R I 1977 Noether's theorem in classical mechanics *Am. J. Phys.* **45** 336–40