Introduction to invariant and equivariant problems

The curve completion problem

Consider the 'curve completion problem', which is a subproblem of the much more complex 'inpainting problem'. Suppose we are given a partially obscured curve in the plane, as in Figure 0.1, and we wish to fill in the parts of the curve that are missing. If the missing bit is small, then a straight line edge can be a cost effective solution, but this does not always give an aesthetically convincing look. Considering possible solutions to the curve completion problem (Figure 0.2), we arrive at three requirements on the resulting curve:

- it should be sufficiently smooth to fool the human eye,
- if we rotate and translate the obscured curve and then fill it in, the result should be the same as filling it in and then rotating and translating,
- it should be the 'simplest possible' in some sense.

The first requirement means that we have boundary conditions to satisfy as well as a function space in which we are working. The second means the formulation of the problem needs to be 'equivariant' with respect to the standard action of the Euclidean group in the plane, as in Figure 0.3. This condition arises naturally: for example, if the image being repaired is a dirty photocopy, the result should not depend on the angle at which the original is fed into the photocopier.

All three conditions can be satisfied if we require the resulting curve to be such as to minimise an integral which is invariant under the group action,

$$\int L(s, \kappa, \kappa_s, \ldots)\, \mathrm{d}s, \tag{0.1}$$

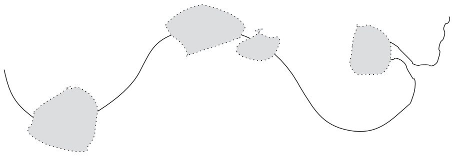

Figure 0.1 A curve in the plane with occlusions.

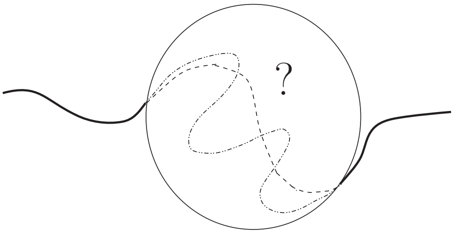

Figure 0.2 Which infilling is best?

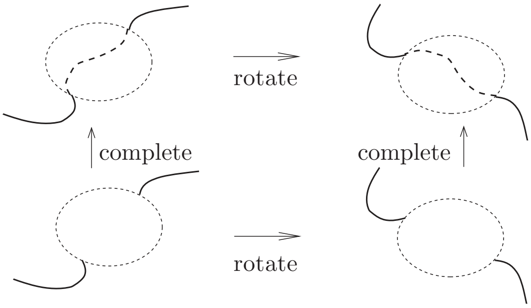

Figure 0.3 The solution is equivariant.

where s is arc length and κ the Euclidean curvature,

$$\kappa = \frac{u_{xx}}{(1+u_x^2)^{3/2}}, \qquad \frac{\mathrm{d}}{\mathrm{d}s} = \frac{1}{\sqrt{1+u_x^2}}\frac{\mathrm{d}}{\mathrm{d}x} \tag{0.2}$$

and $\mathrm{d}s = \sqrt{1+u_x^2}\,\mathrm{d}x$.

The theory of the Calculus of Variations is about finding curves that minimise integrals such as equation (0.1), and the most famous Lagrangian in this family is

$$\mathcal{L}[u] = \int \kappa^2\,\mathrm{d}s. \tag{0.3}$$

The main theorem in the Calculus of Variations is that the minimising curves satisfy a differential equation called the Euler–Lagrange equation. There are quite a few papers and even textbooks that either 'prove' or assume the wrong Euler–Lagrange equation for (0.3), namely that the minimising curve is a circle, that is, satisfying $\kappa = c$. The correct result, calculated by Euler himself, is that the curvature of the minimising curve satisfies

$$\kappa_{ss} + \frac{1}{2}\kappa^3 = 0, \tag{0.4}$$

which is solved by an elliptic function. Solutions are called 'Euler's elastica' and have many applications. See Chan *et al.* (2002) for a discussion relevant to the inpainting problem.

While Euler–Lagrange equations can be found routinely by symbolic computation packages, and then rewritten in terms of historically known invariants, this process reveals little to nothing of why the Euler–Lagrange equation has the terms and features it does. The motivating force behind Chapter 7 was to bring out and understand the structure of Euler–Lagrange equations for variational problems where the integrand, called a Lagrangian, is invariant under a group; the groups relevant here are not finite groups, but *Lie groups*, those that can be parametrised by real or complex numbers, such as translations and rotations.

One of the most profound theorems of the Calculus of Variations is Noether's Theorem, giving formulae for first integrals of Euler–Lagrange equations for Lie group invariant Lagrangians. Most Lagrangians arising in physics have such an invariance; the laws of nature typically remain the same under translations and rotations, also pseudorotations in relativistic calculations, and so on, and thus Noether's Theorem is well known and much used.

If one calculates Noether's first integrals for the variational problem (0.3), the result can be written in the form,

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{1+u_x^2}} & -\dfrac{u_x}{\sqrt{1+u_x^2}} & 0 \\ \dfrac{u_x}{\sqrt{1+u_x^2}} & \dfrac{1}{\sqrt{1+u_x^2}} & 0 \\ \dfrac{xu_x - u}{\sqrt{1+u_x^2}} & \dfrac{uu_x + x}{\sqrt{1+u_x^2}} & 1 \end{pmatrix} \begin{pmatrix} -\kappa^2 \\ -2\kappa_s \\ 2\kappa \end{pmatrix} \tag{0.5}$$

where the c_i are the constants of integration. The first component comes from translation in x, the second from translation in u and the third from rotation in the (x, u) plane about the origin. The 3×3 matrix appearing in (0.5), which I denote here by $B(x, u, u_x)$, has a remarkable property. If one calculates the induced action of the group of rotations and translations in the plane, that is, the special Euclidean group $SE(2)$, on B, componentwise, then one has

$$B(g \cdot x, g \cdot u, g \cdot u_x) = R(g)B(x, u, u_x), \qquad \text{for all } g \in SE(2)$$

where $R(g)$ is a particular matrix representation of $SE(2)$ called the Adjoint representation. In other words, $B(x, u, u_x)$ is *equivariant* with respect to the group action, and is thus an equivariant map from the space with coordinates $(x, u, u_x, u_{xx}, \ldots)$ to $SE(2)$. The equivariance can be used to understand how the group action takes solutions of the Euler–Lagrange equations to solutions.

Equivariant maps are, in fact, the secret to success for the invariant calculus. They are denoted as a 'moving frame' and are the central theme of Chapter 4. In Chapter 7 we prove results that give the structure of both the Euler–Lagrange equations and the set of first integrals for invariant Lagrangians, using the symbolic invariant calculus developed in Chapters 4 and 5. The fact that the formula for Noether's Theorem yields the very map required to establish the symbolic invariant calculus, used in turn to understand the structure of the results, continues to amaze me.

Curvature flows and the Korteweg–de Vries equation

Consider the group of 2×2 real matrices with determinant 1, called $SL(2)$, which we write as

$$SL(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid ad - bc = 1 \right\}.$$

We are interested in actions of this group on, say, curves in the (x, u) plane, that evolve in time, so our curves are parametrised as $(x, t, u(x, t))$. Suppose for $g \in SL(2)$ we impose that the group acts on curves via the map

$$g \cdot x = x, \qquad g \cdot t = t, \qquad g \cdot u = \frac{au + b}{cu + d}.$$

Using the chain rule, we can induce an action on u_x and higher derivatives, as

$$g \cdot u_x = \frac{\partial(g \cdot u)}{\partial(g \cdot x)} = \frac{u_x}{(cu + d)^2},$$

and

$$g \cdot u_{xx} = \frac{\partial(g \cdot u_x)}{\partial(g \cdot x)}$$

and so on. It is then a well-established historical fact that the lowest order invariants are

$$W = \frac{u_t}{u_x}, \qquad V = \frac{u_{xxx}}{u_x} - \frac{3}{2}\frac{u_{xx}^2}{u_x^2} := \{u; x\}.$$

The invariant V is called the Schwarzian derivative of u and is often denoted as $\{u; x\}$. This derivative featured strongly in the differential geometry of a bygone era; it is used today in the study of integrable systems. The reason is as follows. The invariants V and W are functionally independent, but there is a differential identity or syzygy,

$$\frac{\partial}{\partial t} V = \underbrace{\left(\frac{\partial^3}{\partial x^3} + 2V\frac{\partial}{\partial x} + V_x \right)}_{\mathcal{H}} W.$$

The operator $\mathcal{H}$ appearing in this equation is one of the two Hamiltonian operators for the Korteweg–de Vries equation, see Olver (1993), Example 7.6, with $V = u/3$. Thus, if $W = V$, that is if $u_t = u_x\{u; x\}$, then $V(x, t)$ satisfies the Korteweg–de Vries equation.

In fact there are many examples like this, where syzygies between invariants give rise to pairs of partial differential equations that are integrable, with one of the pair being in terms of the invariants of a given smooth group action. Another example of such a pair is the vortex filament equation and the non-linear Schrödinger equation. In that case, the group action is the standard action of the group of rotations and translations in $\mathbb{R}^3$. We refer to Mansfield and van der Kamp (2006) and to Marí Beffa (2004, 2007, 2008a, 2008b) for more information.

The essential simplicity of the main idea

For many applications, what seems to be wanted is the following:

> given the smooth group action, derive the invariants and their syzygies *algorithmically*, that is, without prior knowledge of 100 years of differential geometry, and with *minimal effort*.

To show the essential simplicity of the main idea, we consider a simple set of transformations of curves $(x, u(x))$ in the plane given by

$$x \mapsto \widetilde{x} = \lambda x + k, \qquad u \mapsto \widetilde{u} = \lambda u, \qquad \lambda \neq 0. \tag{0.6}$$

The induced action on tangent lines to the curves is given by the chain rule:

$$u_x \mapsto \frac{\mathrm{d}\widetilde{u}}{\mathrm{d}\widetilde{x}} = \frac{\mathrm{d}\widetilde{u}}{\mathrm{d}x}\left(\frac{\mathrm{d}\widetilde{x}}{\mathrm{d}x}\right)^{-1} = \frac{\lambda u_x}{\lambda} = u_x$$

and so u_x is an invariant. Continuing, we obtain

$$u_{xx} \mapsto \frac{u_{xx}}{\lambda}, \qquad u_{xxx} \mapsto \frac{u_{xxx}}{\lambda^2}$$

and so on. Of course, in this simple example, we can see what the invariants have to be. But let us pretend we do not for some reason, and derive a set of invariants.

The basic idea is to solve two equations for the two parameters λ and k. If we take $\widetilde{x} = 0$ and $\widetilde{u} = 1$, we obtain

$$\lambda = \frac{1}{u}, \qquad k = -\frac{x}{u}. \tag{0.7}$$

We give these particular values of the parameters the grand title ‘the frame’. If we now evaluate the images of $u_{xx}, u_{xxx}, \ldots$ under the mapping, with λ and k given by the frame parameters in equation (0.7), we obtain

$$u_{xx} \mapsto \frac{u_{xx}}{\lambda} \mapsto uu_{xx}, \qquad u_{xxx} \mapsto \frac{u_{xxx}}{\lambda^2} \mapsto u^2u_{xxx}, \qquad \ldots.$$

We now observe that the final images of our maps are all invariants. Indeed,

$$u^2u_{xxx} \mapsto (\lambda u)^2\left(\frac{u_{xxx}}{\lambda^2}\right) = u^2u_{xxx}$$

and so on. The method of ‘solve for the frame, then back-substitute’ has produced an invariant of every order, specifically

$$I_n = u^{n-1}u_{\underbrace{xx\ldots x}_{n \text{ terms}}}.$$

It is easy to show that any invariant can be expressed in terms of the I_n. Indeed, if $F(x, u, u_x, \ldots)$ is an invariant, then

$$F(x, u, u_x, u_{xx}, \ldots) = F\left(\lambda x + k, \lambda u, u_x, \frac{u_{xx}}{\lambda}, \ldots\right) \tag{0.8}$$

for all λ and k. If I use the ‘frame’ values of the parameters in equation (0.8), I obtain

$$F(x, u, u_x, u_{xx}, \ldots) = F(0, 1, u_x, uu_{xx}, \ldots) = F(0, 1, I_1, I_2, \ldots).$$

Since any invariant at all can be written in terms of the I_n, we have what is called a *generating set* of invariants.

But that is not all. If I use the same approach on the derivative operator

$$\frac{\mathrm{d}}{\mathrm{d}x} \mapsto \frac{\mathrm{d}}{\mathrm{d}\widetilde{x}} = \left(\frac{\mathrm{d}\widetilde{x}}{\mathrm{d}x}\right)^{-1} \frac{\mathrm{d}}{\mathrm{d}x} = \lambda \frac{\mathrm{d}}{\mathrm{d}x} \mapsto u \frac{\mathrm{d}}{\mathrm{d}x}$$

then the final result,

$$\mathcal{D} = u \frac{\mathrm{d}}{\mathrm{d}x}$$

is invariant, that is,

$$\mathcal{D} \mapsto \widetilde{u} \frac{\mathrm{d}}{\mathrm{d}\widetilde{x}} = u \frac{\mathrm{d}}{\mathrm{d}x} = \mathcal{D}.$$

Differentiating an invariant with respect to an invariant differential operator must yield an invariant, and indeed we obtain

$$\mathcal{D}I_1 = I_2, \qquad \mathcal{D}I_2 = I_3 + I_1 I_2 \tag{0.9}$$

and so on.

Equations of the form (0.9) are called *symbolic differentiation formulae*. The major advance made by Fels and Olver (Fels and Olver, 1998, 1999) was to find a way to obtain equations (0.9) without knowing the frame, but only the equations used to define the frame, which in this case were $\widetilde{x} = 0$, $\widetilde{u} = 1$.

If we now look at a matrix form of our mapping,

$$\begin{pmatrix} \widetilde{x} \\ \widetilde{u} \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & k \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ u \\ 1 \end{pmatrix}$$

and evaluate the matrix of parameters on the frame, we obtain 'the matrix form of the frame',

$$\varrho(x, u) = \begin{pmatrix} \dfrac{1}{u} & 0 & -\dfrac{x}{u} \\ 0 & \dfrac{1}{u} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Going one step further, if we act on this matrix $\varrho(x, u)$, we obtain

$$
\begin{aligned}
\varrho(\widetilde{x}, \widetilde{u}) &= \begin{pmatrix} \dfrac{1}{\widetilde{u}} & 0 & -\dfrac{\widetilde{x}}{\widetilde{u}} \\ 0 & \dfrac{1}{\widetilde{u}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \dfrac{1}{u} & 0 & -\dfrac{x}{u} \\ 0 & \dfrac{1}{u} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{1}{\lambda} & 0 & -\dfrac{k}{\lambda} \\ 0 & \dfrac{1}{\lambda} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= \varrho(x, u) \begin{pmatrix} \lambda & 0 & k \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}.
\end{aligned}
$$

What this result means is that ‘the frame’ is *equivariant* with respect to the mapping (0.6).

The miracle is that the entire symbolic calculus can be built from the equivariance of the frame and ordinary multivariable calculus, even if you do not know the frame explicitly, that is, even if you cannot solve the equations giving the frame for the parameters.

The one caveat is that not any old mapping involving parameters can be studied this way; the mapping (0.6) is in fact a *Lie group action*, where the Lie group is the set of 3×3 matrices

$$
\left\{ \begin{pmatrix} \lambda & 0 & k \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \mid \lambda, k \in \mathbb{R}, \lambda > 0 \right\}
$$

(amongst other representations) which is closed under multiplication and inversion.

Not all group actions are linear like (0.6), and since we do not need to assume linearity for any of the theory to be valid, we do not assume it. However, often the version of a theorem assuming a linear action is easier to state and prove, and so we tend to do both.

Overview of this book

My primary aim in writing this book was to bring the theory and applications of moving frames to an audience not wishing to learn Differential Geometry first, to show how the calculations can be done using primarily undergraduate calculus, and to provide a discussion of a range of applications in a fully detailed way so that readers can do their own calculations without undue headscratching.

The main subject matter is, first and foremost, smooth group actions on smooth spaces. Surprisingly, this includes applications to many seemingly discrete problems. The groups referred to in this book are *Lie groups*, groups that depend on real or complex parameters. In Chapter 1 we discuss the basic notions concerning Lie groups and their actions, particularly their actions as prolonged to derivative terms. Since there is a wealth of excellent texts on this topic, we cruise through the examples, calculations and basic definitions, introducing the main examples I use throughout.

The following two chapters give foundational material for Lie theory as I use and need it for this book. I could not find a good text with exactly what was needed, together with suitable examples and exercises, so I have written this myself, proving everything from scratch. While I imagine most readers will only refer to them as necessary, hopefully others will be inspired to learn more Differential Topology and Lie Theory from texts dedicated to those topics.

In Chapter 2, I discuss how multivariable calculus extends to a calculus on Lie groups; this is mostly an introduction to standard Differential Topology for the particular cases of interest, and a discussion of the central role played by one parameter subgroups. The point of view taken in Differential Topology, on 'what is a vector' and 'what is a vector field', is radically different to that taken in Differential Geometry. The first theory bases the notion of a vector on a path, the second on the algebraic notion of a derivation acting on functions. There are serious problems with a *definition* of a vector field as a derivation.† On the other hand, the notion of a vector as a path in the space, which can be differentiated at its distinguished point in coordinates, is a powerful, all purpose, take anywhere idea that has a clear and explicit link to standard multivariable calculus. Further, anyone who has witnessed a leaf being carried by water, or a speck of dust being carried by the wind, has already developed the necessary corresponding intuitive notion. Armed with the clear and useful notion of a vector as a path, everything we need can be proved from the theorem guaranteeing the existence and uniqueness of a solution to first order differential systems. So as to give

† Not the least problem is that the chain rule needed for the transformation of vectors does not follow from this definition alone, which can apply equally well to strictly algebraic objects.

those well versed in one language insight into the other, we give some links between the two sets of ideas and the relevant notations.

In the second chapter on the foundations of Lie theory, Chapter 3, we discuss the Lie bracket of vector fields and Frobenius' Theorem and from there, the Lie algebra, the Lie bracket, and the Adjoint and adjoint actions. The two quite different appearances of the formulae for the Lie bracket, for matrix groups and transformation groups, are shown to be instances of the one general construction, which in turn relies on the Lie bracket of vector fields in $\mathbb{R}^n$. While many authors simply give the two different formulae as definitions, I was not willing to do that for reasons I make clear in the introduction to that chapter.

Chapters 4 and 5 are the central chapters of the book. The key idea underlying the symbolic invariant calculus is a formulation of a moving frame as an equivariant map from space M on which the group G acts, to the group itself. When one can solve for the frame, one has explicit invariants and invariant differential operators. When one cannot solve for the frame, then one has symbolic invariants and invariant differential operators. This is the topic of Chapter 4, which introduces the distinguished set of symbolic invariants and symbolic invariant differentiation operators used throughout the rest of the book. Chapter 5 continues the main theoretical development to discuss the differential relations or syzygies satisfied by the invariants, and introduces the curvature matrices. These are well known in differential geometry, and we discuss the famous Serret–Frenet frame, but they have other applications; in particular, they can be used to solve numerically for the frame. Both chapters have sections detailing various applications and further developments; sections designated by a star, *, can be omitted on a first reading.

From this firm theoretical foundation, a host of applications can be described. The two most developed applications in this book are to solving invariant ordinary differential equations, and to the Calculus of Variations. In fact, there is a long history of using smooth group actions to solve invariant ordinary differential equations; normally one would think of this theory as a success story, with little more to say. However, we describe in Chapter 6 just how much more can be achieved with the new ideas. Similarly, the Calculus of Variations is a classical subject that one might think of as fully mature. In Chapter 7, the use of the new ideas throws substantial light on the *structure* of the known results when invariance under a smooth group action is given.

The three applications that pleased me the most were solving the Chazy equation, finding the equations for a free rigid body without any mysterious concepts, and the final theorem of the book, showing the structure of the first integrals given by Noether's Theorem. All three came out of trying to develop interesting exercises for this book.