Math 5190: Lecture notes.

1. Geometric Preliminaries

1.1. **Differential equations and their solutions.** We begin by reviewing the concepts of an ordinary differential equation and initial value problem. Vector fields, integral curves, and flows furnish us with a geometric formulation of these concepts.

Let $U \subset \mathbb{R}^{n+1}$ be an open domain and $\mathbf{V} : U \to \mathbb{R}^n$ a C^1 (continuously differentiable) vector valued function.

Definition 1.1. *An ordinary differential equation is a constraint of the form*

$$\phi'(t) = \mathbf{V}(t, \phi(t)), \tag{1}$$

where $\phi : I \to U$ *is a continuously differentiable function called* the solution *of the equation. An initial value problem (IVP) is an ODE together with a constraint of the form*

$$\phi(0) = \boldsymbol{x}_0, \tag{2}$$

where $\boldsymbol{x}_0 \in U$ *is a given initial position.*

Definition 1.2. *We call the ODE* (1) *autonomous if* $\mathbf{V}$ *does not depend on* t*. The autonomization of* (1) *is the autonomous ODE in* $n+1$ *dependent variables having the form*

$$\xi'(t) = 1, \tag{3}$$

$$\phi'(t) = \mathbf{V}(\xi(t), \phi(t)) \tag{4}$$

Using autonomization no generality is lost if we develop the general aspects of ODE theory strictly in terms of autonomous equations.

Intuitively, an ODE is an assignment of a velocity to every point in space and time. A solution curve represents the motion of a particle whose velocity is determined by the time and the particle's position. An autonomous ODE simply describes a velocity field that depends only on position and is independent of time. Therefore, it is useful to consider an autonomous ODE as a vector (velocity) field.

In our discussion, it is important to make a distinction between points and vectors. The former should be considered as elements of some domain $U \subset \mathbb{R}^n$, whereas the latter are elements of the vector space $\mathbb{R}^n$. Geometrically, it makes sense to take linear combinations of vectors, but there is little profit in trying to "add" points. Likewise, in analysis of ODEs, the zero vector plays a very special role. However, there is nothing special about the point $(0, \dots, 0)$; it is just an arbitrarily chosen origin, and may not even be included in U.

Thus, a vector field is a function $\mathbf{V} : U \to \mathbb{R}^n$ that assigns to every point of U a vector in $\mathbb{R}^n$. We will also write a vector field as

$$\mathbf{V} = \sum_{i=1}^{n} V^i(x^1, \dots, x^n)\mathbf{e}_i, \quad \boldsymbol{x} = (x^1, \dots, x^n) \in U;$$

here $\mathbf{e}_1, \dots, \mathbf{e}_n$ is the standard basis of $\mathbb{R}^n$ and $V^1, \dots, V^n$ are the scalar components of a vector in $\mathbb{R}^n$. If we assume that the components V^i are continuously differentiable functions, we will speak of a C^1 vector field. If the components are analytic (expressible as power series) functions we will speak of an analytic vector field.

For $f : U \to \mathbb{R}$ a differentiable function, we define $\mathbf{V}[f]$, the directional derivative of f with respect to V, to be the function

$$\mathbf{V}[f] = \sum_{i=1}^{n} V^i \, \mathrm{D}_i f \tag{5}$$

For this reason, we identify vector fields with first order differential operators, and write

$$\mathbf{V} = \sum_{i=1}^{n} V^i(x^1, \dots, x^n) \frac{\partial}{\partial x^i}.$$

In this formulation the standard vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ correspond to the partial derivative operators $\partial/\partial x^1, \dots, \partial/\partial x^n$.

As we already mentioned a vector field is exactly the same as an autonomous ODE. An *integral curve* is a parameterized curve $\phi : I \to U$, where $I \subset \mathbb{R}$ is an open interval, such that

$$\phi'(t) = \mathbf{V}(\phi(t)), \quad t \in I. \tag{6}$$

In other words, an integral curve is a particular solution of the above autonomous ODE. The general solution of an autonomous, n-dimensional ODE depends on n constants of integration. If we judiciously include the constants of integration as parameters in a general solution, we obtain something called a flow.

Let $U \subset \mathbb{R}^n$ be an n-dimensional domain, $\hat{U} \subset \mathbb{R} \times U$ an $(n+1)$-dimensional domain, and $\Phi : \hat{U} \to U$ a C^1 function.

Definition 1.3. *We say that Φ is a flow if it satisfies the following conditions.*

(A) *For all $\boldsymbol{x} \in U$, the quantity $\Phi(0, \boldsymbol{x})$ is defined and equal to $\boldsymbol{x}$;*
(B) *We have*

$$\Phi(s, \Phi(t, \boldsymbol{x})) = \Phi(s + t, \boldsymbol{x}); \tag{7}$$

(C) *For a fixed* $\boldsymbol{x} \in U$ *the set*

$$U_{\boldsymbol{x}} = \{t \in \mathbb{R} : (t, \boldsymbol{x}) \in \hat{U}\}; \tag{8}$$

is an open interval in $\mathbb{R}$.

If in addition to the above 3 conditions,

(D) *the RHS of* (10) *is defined whenever the LHS is,*

we call $\Phi(t, \boldsymbol{x})$ *a maximal flow.*

A flow is a special way to specify a general solution of an ODE, one where the constants of integration represent initial position. What is the relationship between vector fields and flows?

Definition 1.4. *Let* $\mathbf{V} : U \to \mathbb{R}^n$ *be a vector field and* $\Phi(t, \boldsymbol{x})$, $\boldsymbol{x} \in U$ *a flow. We say that* $\mathbf{V}$ generates *the flow if*

$$\mathbf{V}(\boldsymbol{x}) = (\mathrm{D}_1\Phi)(0, \boldsymbol{x}). \tag{9}$$

Note: above $\mathrm{D}_1\Phi$ denotes the partial derivative with respect to the first, "time" variable.

For each $t \in \mathbb{R}$ let us set $U_t = \{\boldsymbol{x} \in \mathbb{R}^n : (t, \boldsymbol{x}) \in \hat{U}\}$ and define the transformation $\Phi_t : U_t \to U$ by

$$\Phi_t(\boldsymbol{x}) = \Phi(t, \boldsymbol{x}), \quad (t, \boldsymbol{x}) \in \hat{U}.$$

Condition (A) can be restated by saying that Φ_0 is the identity transformation on U. Condition (B) can be restated as the 1-parameter group law

$$\Phi_{s+t} = \Phi_s \circ \Phi_t. \tag{10}$$

For this reason, we often refer to a flow as a 1-parameter group generated by a vector field

Theorem 1.5 (Existence)**.** *Let* $\mathbf{V} : U \to \mathbb{R}^n$ *be a* C^1 *vector field. Then, there exist a* C^1 *flow generated by* $\mathbf{V}$.

Theorem 1.6 (Uniqueness)**.** *Let* $\mathbf{V} : U \to \mathbb{R}^n$ *be a* C^1 *vector field. Any two flows generated by* $\mathbf{V}$ *agree on some open neighbourhood of* $U \subset \mathbb{R} \times U$. *There exists a unique maximal flow generated by* $\mathbf{V}$.

We will present the proof later. Here, we present some examples.

Example 1.7. Consider the differential equation

$$\dot{x} = \frac{1}{x}, \quad x > 0.$$

The general solution is

$$x = \sqrt{K + 2t}.$$

The corresponding vector field is

$$V = \frac{1}{x}\frac{\partial}{\partial x}, \quad x > 0.$$

The domain of the vector field is $U = \{x \in \mathbb{R} : x > 0\}$. The corresponding flow is given by

$$\Phi(t, x) = \sqrt{x^2 + 2t}.$$

The domain of the flow is

$$\hat{U} = \{(t, x) : x > 0 \text{ and } -\frac{1}{2}x^2 < t < +\infty\}.$$

Note that the limit of $\Phi(t, x)$ as $t \to +\infty$ does not exist, while the limit of $\Phi(t, x)$ as $t \to -x^2/2$ is equal to 0, a value outside the domain U.

Example 1.8. Let's consider the non-autonomous ODE

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

The general solution is obtained by rewriting the ODE as

$$e^x \frac{de^{-x}y}{dx} = x,$$

separating variables and integrating. We obtain

$$y = -x - 1 + Ce^x.$$

where C is a constant of integration.

Let us rewrite the general solution as a flow. First, we autonomize the equation by rewriting it as

$$\begin{aligned}\dot{x} &= 1,\\ \dot{y} &= x + y,\end{aligned}$$

and writing the solution as

$$\begin{aligned}x &= x_0 + t,\\ y &= -(x_0 + t) - 1 + Ce^{x_0+t}.\end{aligned}$$

We set $t = 0$ in the second equation and obtain

$$y_0 = -x_0 - 1 + Ce^{x_0},$$

or equivalently,

$$C = e^{-x_0}(y_0 + x_0 + 1).$$

Therefore, the general solution, written as a flow, is

$$\begin{aligned}x &= x_0 + t\\ y &= (y_0 + x_0 + 1)e^t - (x_0 + t) - 1.\end{aligned}$$

The domain of the flow is all real (t, x, y). A straightforward calculation shows that condition (10) is satisfied.

Finally, let us discuss the relationship between the general solution to an ODE and a flow. We would like to say that a flow is the same thing as a general solution where the initial values serve as the constants of integration. In one direction, we have the following

Proposition 1.9. *Let $\Phi(t, \boldsymbol{x})$ be a flow generated by a C^1 vector field $\mathbf{V} : U \to \mathbb{R}^n$. Then,*

$$\frac{\partial \Phi(t, \boldsymbol{x})}{\partial t} = \mathbf{V}(\Phi(t, \boldsymbol{x})).$$

Proof. Homework. □

For the other direction, we will make use of the following fundamental Theorem regarding solutions of initial value problems.

Theorem 1.10. *Let $\mathbf{V} : U \to \mathbb{R}^n$ be a C^1 vector field and $\boldsymbol{x}_0 \in U$ a fixed position. Then, there exists a C^2 curve $\phi(t)$ such that*

$$\begin{aligned} \phi'(t) &= \mathbf{V}(\phi(t)), \\ \phi(0) &= \boldsymbol{x}_0. \end{aligned}$$

The curve in question is locally unique. This means that all such curves agree in some neighborhood of $t = 0$.

Proposition 1.11. *Let $\mathbf{V}(\boldsymbol{x})$ be a vector field, and let $\Phi(t, \boldsymbol{x})$ be a general solution of the ODE $\dot{\boldsymbol{x}} = \mathbf{V}(\boldsymbol{x})$; i.e.;*

$$\frac{\partial \Phi(t, \boldsymbol{x})}{\partial t} = \mathbf{V}(\Phi(t, \boldsymbol{x})).$$

If in addition, $\Phi(0, \boldsymbol{x}) = \boldsymbol{x}$, then $\Phi(t, \boldsymbol{x})$ is a flow.

Proof. Let us establish that Φ satisfies the 1-parameter group law (B). Fix a position $\boldsymbol{x}_0$ and a time t_1 set

$$\boldsymbol{x}_1 = \Phi(t, \boldsymbol{x}_0),$$

and define curves

$$\begin{aligned} \phi(t) &= \Phi(t + t_1, \boldsymbol{x}_0), \\ \psi(t) &= \Phi(t, \boldsymbol{x}_1). \end{aligned}$$

Observe that

$$\begin{aligned}
&\dot{\phi}(t) = \dot{\Phi}(t + t_1, \boldsymbol{x}_0) = \mathbf{V}(\Phi(t + t_1, \boldsymbol{x}_0)) = \mathbf{V}(\phi(t)),\\
&\psi(0) = \boldsymbol{x}_1,\\
&\dot{\psi}(t) = \dot{\Phi}(t, \boldsymbol{x}_1) = \mathbf{V}(\Phi(t, \boldsymbol{x}_1)) = \mathbf{V}(\psi(t)),\\
&\psi(0) = \boldsymbol{x}_1.
\end{aligned}$$

Therefore, by the uniqueness of solutions to initial value problems, we must have

$$\Phi(t + t_1, \boldsymbol{x}_0) = \phi(t) = \psi(t) = \Phi(t, \Phi(t_1, \boldsymbol{x}_0)),$$

as was to be shown. $\square$

2. The reparameterization theorem.

A vector can be informally defined as a quantity having the properties of magnitude and direction. When the motion of a hypothetical particle is governed by a field of velocity vectors, two distinct aspects of the motion are determined. First, is the trajectory of the particle, that is to say the geometric shape traced out by the particle's motion. Second, is the speed with which the particle traverses that trajectory. If we modify a field of velocities by changing the magnitude, but not the direction of the vectors, it is reasonable to expect that the corresponding trajectories retain their shape, but that the hypothetical test particles move along these trajectories with a different speed.

In this regard, it is important to distinguish between a curve and its parameterization. For example, consider the unit semicircle in $\mathbb{R}^2$ given by

$$x^2 + y^2 = 1, \quad y > 0.$$

Here are two parameterizations of this curve:

$$x = s, \quad y = \sqrt{1 - s^2}, \quad -1 < s < 1,$$

and

$$x = \cos t, \quad y = \sin t, \quad 0 < t < \pi.$$

The two parameterizations are related by an invertible reparameterization function, namely

$$s = \cos t, \quad 0 < t < \pi.$$

In general, let $\phi(t) \in \mathbb{R}^n$ be a parameterization of a curve in n-dimensional space. We say that $\psi(s) \in \mathbb{R}^n$ is a reparameterization of the same curve if there exists a continuously differentiable function $t = \tau(s)$ such that $\tau'(s) \neq 0$ for all s and such that $\psi(s) = \phi(\tau(s))$.

Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field. Let $g : U \to \mathbb{R}$ be a C^1, nowhere vanishing function: $g(\boldsymbol{a}) \neq 0$, $\boldsymbol{a} \in U$. Scaling $\boldsymbol{A}$ by g gives us a new vector field, namely $\boldsymbol{B} = g\boldsymbol{A}$.

Theorem 2.1. *The integral curves of $\boldsymbol{A}$ and $\boldsymbol{B}$ are related by a reparameterization.*

Proof. Suppose that $\boldsymbol{B} = g\boldsymbol{A}$. Consider the integral curves $\phi(t), \psi(t)$ of $\boldsymbol{A}, \boldsymbol{B}$, respectively, both originating at some fixed initial point $\boldsymbol{x} \in U$;

$$\phi(0) = \psi(0) = \boldsymbol{x}.$$

We need to construct a reparameterization function $\tau(s)$ such that

$$\psi(s) = \phi(\tau(s)).$$

We will do this in a way that incorporates dependence on the initial point $\boldsymbol{x}$ explicitly.

Let $\Phi(t, \boldsymbol{x})$ be the flow corresponding to $\boldsymbol{A}(\boldsymbol{x})$. For each $t, \boldsymbol{x}$ in the domain of $\Phi(t, \boldsymbol{x})$, set

$$\sigma(t, \boldsymbol{x}) = \int_0^t \frac{du}{g(\Phi(u, \boldsymbol{x}))}.$$

For $\boldsymbol{x} \in U$ and $t_1, t_2 \geq 0$ set

$$\boldsymbol{y} = \Phi(t_1, \boldsymbol{x}), \quad s_1 = \sigma(t_1, \boldsymbol{x}), \quad s_2 = \sigma(t_2, \boldsymbol{y}).$$

We have

$$\begin{aligned}
\sigma(t_1 + t_2, \boldsymbol{x}) &= \int_0^{t_1+t_2} \frac{du}{g(\Phi(u, \boldsymbol{x}))} && (11)\\
&= \int_0^{t_1} \frac{du}{g(\Phi(u, \boldsymbol{x}))} + \int_{t_1}^{t_1+t_2} \frac{du}{g(\Phi(u, \boldsymbol{x}))} && (12)\\
&= \sigma(t_1, \boldsymbol{x}) + \int_0^{t_2} \frac{dv}{g(\Phi(v, \boldsymbol{y}))} && (13)\\
&= s_1 + s_2, && (14)
\end{aligned}$$

The step from (12) to (13) is justified by making a change of variables $u = v + t_1$ and by noting that

$$\Phi(u, \boldsymbol{x}) = \Phi(v + t_1, \boldsymbol{x}) = \Phi(v, \boldsymbol{y}), \quad 0 \leq v \leq t_2.$$

By assumption,

$$\dot{\sigma}(t, \boldsymbol{x}) = \frac{1}{g(\Phi(t, \boldsymbol{x}))} \neq 0.$$

Hence, by the inverse function theorem there exists an inverse function $\tau(s, \boldsymbol{x})$ such that

$$\tau(\sigma(t, \boldsymbol{x}), \boldsymbol{x}) = t.$$

We now define

$$\Psi(s, \boldsymbol{x}) = \Phi(\tau(s, \boldsymbol{x}), \boldsymbol{x}), \tag{15}$$

and claim that $\Psi(s, \boldsymbol{x})$ is the flow determined by $\boldsymbol{B}$. Since $\sigma(0, \boldsymbol{x}) = 0$, we have $\tau(0, \boldsymbol{x}) = 0$, whence

$$\Psi(0, \boldsymbol{x}) = \Phi(0, \boldsymbol{x}) = \boldsymbol{x}.$$

By (14) we have

$$s_1 + s_2 = \sigma(t_1 + t_2, \boldsymbol{x})$$

Hence, by (11), we have

$$\tau(s_1 + s_2, \boldsymbol{x}) = t_1 + t_2 = \tau(s_1, \boldsymbol{x}) + \tau(s_2, \boldsymbol{x}),$$

and

$$\boldsymbol{y} = \Phi(\tau(s_1, \boldsymbol{x}), \boldsymbol{x}) = \Psi(s_1, \boldsymbol{x}).$$

Therefore,

$$\begin{aligned}
\Psi(s_1 + s_2, \boldsymbol{x}) &= \Phi(\tau(s_1 + s_2, \boldsymbol{x}), \boldsymbol{x}) \\
&= \Phi(\tau(s_2, \boldsymbol{y}) + \tau(s_1, \boldsymbol{x}), \boldsymbol{x}) \\
&= \Phi(\tau(s_2, \boldsymbol{y}), \Phi(\tau(s_1, \boldsymbol{x}), \boldsymbol{x})) \\
&= \Phi(\tau(s_2, \boldsymbol{y}), \boldsymbol{y}) \\
&= \Psi(s_2, \boldsymbol{y}) \\
&= \Psi(s_2, \Psi(s_1, \boldsymbol{x})).
\end{aligned}$$

Finally, observe that

$$\begin{aligned}
\dot{\Psi}(0, \boldsymbol{x}) &= \dot{\tau}(0, \boldsymbol{x})\dot{\Phi}(0, \boldsymbol{x}) \\
&= \frac{1}{\dot{\sigma}(0, \boldsymbol{x})}\boldsymbol{A}(\boldsymbol{x}) \\
&= g(\boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}) \\
&= \boldsymbol{B}(\boldsymbol{x}).
\end{aligned}$$

Therefore, Ψ is the flow generated by $\boldsymbol{B}$. However, every integral curve of $\boldsymbol{A}, \boldsymbol{B}$, respectively, has the form

$$\phi(t) = \Phi(t, \boldsymbol{x}), \quad \psi(t) = \Psi(t, \boldsymbol{x}), \quad \boldsymbol{x} \in U.$$

By construction, all such curves are related by the reparameterization shown in (15). □

Example 2.2. Consider the non-autonomous scalar differential equation

$$\frac{dy}{dx} = -\frac{x}{y}, \quad y > 0.$$

The general solution is

$$y = \sqrt{K - x^2}.$$

Autonomizing the above gives the planar ODE

$$\dot{y} = -\frac{x}{y}, \quad \dot{x} = 1,$$

which corresponds to the vector field

$$\boldsymbol{A} = \frac{\partial}{\partial x} - \frac{x}{y}\frac{\partial}{\partial y}.$$

This vector field determines the following flow:

$$\Phi(t, x, y) = \left(x + t, \sqrt{x^2 + y^2 - (t + x)^2}\right), \quad |t + x| < \sqrt{x^2 + y^2}, \ y > 0$$

Multiplying $\boldsymbol{A}$ by y yields the vector field

$$\boldsymbol{B} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}.$$

The corresponding ODE is linear, with a linear flow, namely

$$\Psi(s,x,y) = \begin{pmatrix} \cos(s) & \sin(s) \\ -\sin(s) & \cos(s) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Let us construct the reparameterization function, as per the proof of the above theorem. We have

$$g(x,y) = y, \quad g(\Phi(t,x,y)) = \sqrt{x^2+y^2-(x+t)^2}.$$

Hence,

$$\begin{aligned} \sigma(t) &= \int_0^t \frac{du}{\sqrt{x^2+y^2-(x+u)^2}} \\ &= \int_x^{t+x} \frac{dv}{\sqrt{x^2+y^2-v^2}} \\ &= \sin^{-1}\left(\frac{t+x}{r}\right) - s_0, \end{aligned}$$

where

$$r = \sqrt{x^2+y^2}, \quad s_0 = \sin^{-1}\left(\frac{x}{r}\right).$$

This is the position dependent reparameterization function from Ψ to Φ. The reparameterization function from Φ to Ψ is given by the inverse, namely

$$\begin{aligned} \tau(s,x,y) &= r\sin(s+s_0) - x, \\ &= r\sin(s)\cos(s_0) + r\cos(s)\sin(s_0) - x, \\ &= y\sin(s) + x\cos(s) - x. \end{aligned}$$

Composing with the reparameterization function with the flow gives

$$\begin{aligned} \Phi(\tau(s,x,y),x,y) &= \left(x+\tau(s,x,y), \sqrt{x^2+y^2-(x+\tau(s,x,y))^2}\right) \\ &= (x\cos(s)+y\sin(s), \sqrt{x^2+y^2-(x\cos(s)+y\sin(s))^2}) \\ &= (x\cos(s)+y\sin(s), \sqrt{(y\cos(s)-x\sin(s))^2}) \\ &= (x\cos(s)+y\sin(s), y\cos(s)-x\sin(s)), \\ &= \Psi(s,x,y). \end{aligned}$$

Next, let us check the domains of the reparameterizations. The time domains for the first flow Φ are given by $-r-x < t < r-x$. The domain of t cannot be extended further because $\Phi(t,x,y)$ tends to $(r,0)$ as $t \to r-x$ and tends to $(-r,0)$ as $t \to -r-x$. However these limits are outside the domain of our vector field.

What about the time domains of the second flow Ψ? Here we must ask the question: what is the domain of $\boldsymbol{B}$? If the domain is all of $\mathbb{R}^2$, then $\Psi(s,x,y)$ is defined for all s,x,y. However, in order for us to compare $\boldsymbol{A}$ and $\boldsymbol{B}$, we must restrict the domain of $\boldsymbol{B}$ to the domain of $\boldsymbol{A}$, namely the upper half plane $\{(x,y) \in \mathbb{R}^2 : y > 0\}$. Hence, the constraint on s becomes

$$-x\sin(s) + y\cos(s) > 0, \quad y > 0,$$

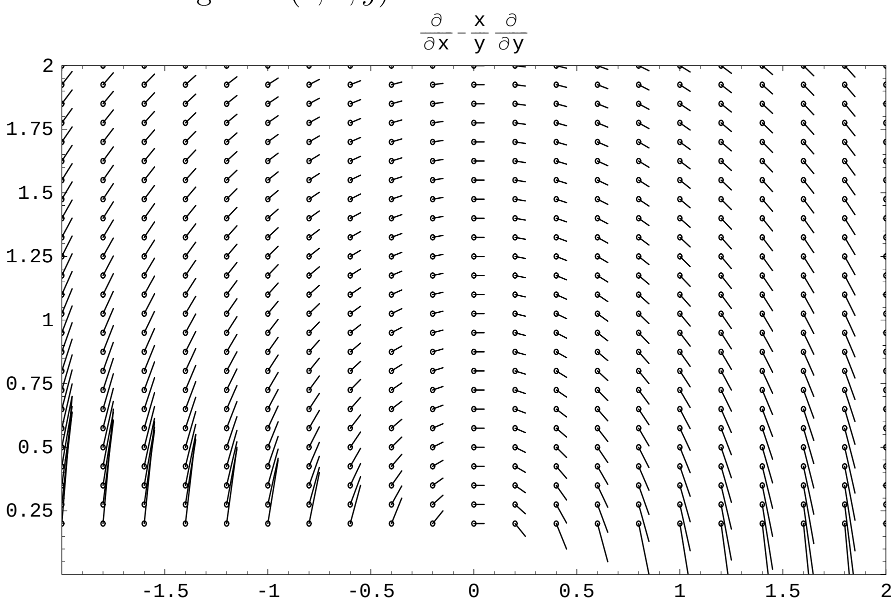
or equivalently,

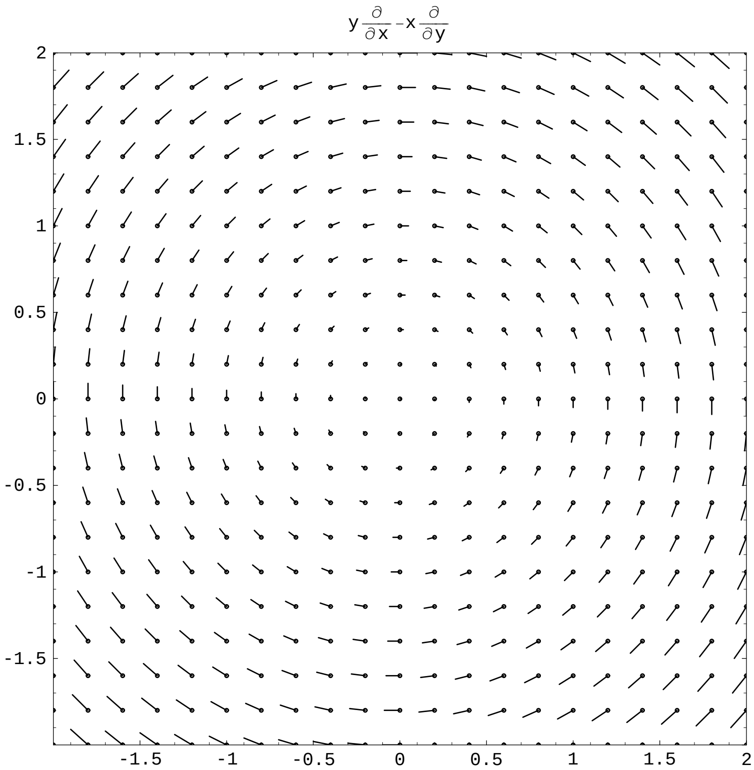
$$-\frac{\pi}{2} - s_0 < s < \frac{\pi}{2} - s_0, \quad s_0 = \tan^{-1}(x/y).$$

The domain of

$$\sigma(t,x,y) = \sin^{-1}\left(\frac{t+x}{r}\right) - s_0$$

is the interval $-r-x < t < r-x$. The image of $s = \sigma(t,x,y)$ is the interval $-\frac{\pi}{2} - s_0 < s < \frac{\pi}{2} - s_0$. Hence the latter interval becomes the domain of the inverse function $\tau(s,x,y)$, while the former interval becomes the image of $\tau(s,x,y)$.




$y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$
2
1.5
1
0.5
0
-0.5
-1
-1.5
-1.5
-1
-0.5
0
0.5
1
1.5
2

3. Diffeomorphisms and coordinate systems.

Definition 3.1. *Let $U, V \subset \mathbb{R}^n$ be domains. A diffeomorphism $\boldsymbol{F} : U \to V$ is an invertible, continuously differentiable transformation such that the inverse transformation is also continuously differentiable.*

Equivalently, we can consider n functions of n variables, $F^i(x^1, \dots, x^n)$, $i = 1, \dots, n$ so that equations

$$y^i = F^i(x^1, \dots, x^n), \quad i = 1, \dots, n \tag{16}$$

have the property that if we restrict $\boldsymbol{y} \in V$, then a solution $\boldsymbol{x} \in U$ exists and is unique. The transformation of the givens $y^1, \dots, y^n$ to the solutions $x^1, \dots, x^n$ defines inverse functions

$$x^i = G^i(y^1, \dots, y^n), \quad i = 1, \dots, n. \tag{17}$$

To have a diffeomorphism we require that all partial derivatives

$$\mathrm{D}_i F^j(x^1, \dots, x^n) = \frac{\partial y^j}{\partial x^i}, \tag{18}$$

$$\mathrm{D}_j G^i(y^1, \dots, y^n) = \frac{\partial x^i}{\partial y^j}, \quad i, j = 1, \dots, n \tag{19}$$

exist and are continuous.

Definition 3.2. *Let $\boldsymbol{F} : U \to \mathbb{R}^n$ be a continuously differentiable transformation. The Jacobian transformation $\mathcal{J}\boldsymbol{F} : U \to \mathrm{Mat}_n \mathbb{R}$ is defined to be the matrix of the partial derivatives of $\boldsymbol{F}$; to wit,*

$$\mathcal{J}\boldsymbol{F} = \begin{pmatrix} \mathrm{D}_1 F^1 & \dots & \mathrm{D}_n F^1 \\ \vdots & \ddots & \vdots \\ \mathrm{D}_1 F^n & \dots & \mathrm{D}_n F^n \end{pmatrix}. \tag{20}$$

Note: here $\mathrm{Mat}_n \mathbb{R}$ denotes the vector space of $n \times n$ real matrices.

Proposition 3.3 (Linear approximation)**.** *Let $\boldsymbol{F} : U \to \mathbb{R}^n$ be a C^1 transformation. Then,*

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}_0) + \mathcal{J}\boldsymbol{F}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|), \quad \boldsymbol{x}, \boldsymbol{x}_0 \in U. \tag{21}$$

Here, the little o notation means that the difference of the LHS and the RHS is a remainder function $\boldsymbol{R}(\boldsymbol{x}, \boldsymbol{x}_0)$ such that

$$\frac{\boldsymbol{R}(\boldsymbol{x}, \boldsymbol{x}_0)}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \to 0 \quad \text{as } \boldsymbol{x} \to \boldsymbol{x}_0.$$

We can also restate the above result by saying that the linear approximation to $\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x})$ at $\boldsymbol{x}_0$ is the affine function

$$\hat{\boldsymbol{y}} = \boldsymbol{y}_0 + M(\boldsymbol{x} - \boldsymbol{x}_0),$$

where $M = \mathcal{J}\boldsymbol{F}(\boldsymbol{x}_0)$ is the indicated $n \times n$ constant matrix and where $\boldsymbol{y}_0 = \boldsymbol{F}(\boldsymbol{x}_0)$. The assertion that $\boldsymbol{F}$ is a C^1 function is equivalent to the condition that

$$\boldsymbol{y} - \hat{\boldsymbol{y}} = o(\|\boldsymbol{x} - \boldsymbol{x}_0\|).$$

Proposition 3.4. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism. Let $\boldsymbol{x}_0 \in U$ be given and set $\boldsymbol{y}_0 = \boldsymbol{F}(\boldsymbol{x}_0)$. Then $\mathcal{J}\boldsymbol{F}(\boldsymbol{x}_0)$ and $\mathcal{J}\boldsymbol{G}(\boldsymbol{y}_0)$ are inverse matrices. Indeed, the linear approximation to the inverse function $\boldsymbol{x} = \boldsymbol{G}(\boldsymbol{y})$ at $\boldsymbol{y}_0 = \boldsymbol{F}(\boldsymbol{x}_0)$ is given by*

$$\hat{\boldsymbol{x}} = \boldsymbol{x}_0 + M^{-1}(\boldsymbol{y} - \boldsymbol{y}_0),$$

where $M = \mathcal{J}\boldsymbol{F}(\boldsymbol{x}_0)$, as before.

We can summarize the above result by saying that linear approximations of inverse functions are inverses of each other. In other words, the affine function $\boldsymbol{x} \mapsto \hat{\boldsymbol{y}}$ is the inverse of the affine function $\boldsymbol{y} \mapsto \hat{\boldsymbol{x}}$.

Theorem 3.5 (Inverse Function Theorem)**.** *If $\boldsymbol{F} : U \to V$ is a diffeomorphism, then*

$$\det \mathcal{J}\boldsymbol{F}(\boldsymbol{a}) \neq 0 \tag{22}$$

for all $\boldsymbol{a} \in U$. Conversely, if $\boldsymbol{F}$ is continuously differentiable and if (22) *holds for a particular $\boldsymbol{a} \in U$, then $\boldsymbol{F}$ is a diffeomorphism provided we restrict the domain to a sufficiently small neighborhood of $\boldsymbol{a}$.*

Example 3.6. Consider the function $F(x) = x^3$. This function is invertible, with inverse $G(x) = x^{1/3}$. It is continuously differentiable: $F'(x) = 3x^2$. However, the inverse isn't: $G'(x) = (1/3)x^{-2/3}$ is undefined at $x = 0$. Therefore $F(x)$ is not a diffeomorphism of the real line. However, if we restrict the domain to, for example to $U = \{a : a > 0\}$, we obtain a diffeomorphism from the half-line U to U.

Example 3.7. The following example illustrates that there exist diffeomorphisms belonging strictly to class C^1. Consider the function

$$y = F(x) = \operatorname{sgn}(x)x^2 + 2x.$$

We claim that this is a diffeomorphism of the real line $\mathbb{R}$. We have $F'(x) = 2|x| + 2$; the derivative exists and is continuous at all $x \in \mathbb{R}$. The inverse function

$$x = G(y) = \begin{cases} \sqrt{y+1} - 1 & \text{if } y \geq 0; \\ 1 - \sqrt{1-y} & \text{if } y < 0 \end{cases} \tag{23}$$

is also continuously differentiable for all $y \in \mathbb{R}$. Indeed,

$$G'(y) = \begin{cases} \dfrac{1}{2\sqrt{y+1}} & \text{if } y \geq 0; \\ \dfrac{1}{2\sqrt{1-y}} & \text{if } y < 0. \end{cases} \tag{24}$$

However, quite clearly $F(x)$ is does not belong to class C^2; the second-order derivative does not exist at $x = 0$.

Example 3.8. Consider the continuously differentiable function $F(x) = x^2$. Observe that $F'(x) = 2x$ vanishes at $x = 0$. Hence, $F(x)$ does not define a diffeomorphism of the real line. However, $F'(x) \neq 0$ for $x > 0$. Therefore, we can define a continuously differentiable inverse, namely $G(x) = \sqrt{x},\ x > 0$ by restricting the domain. The resulting function is a diffeomorphism of the half-line $\{a \in \mathbb{R} : a > 0\}$.

Example 3.9. Consider the transformation of $\mathbb{R}^2 \cong \mathbb{C}$ given by

$$\boldsymbol{F}(x, y) = \exp(x + \mathrm{i}\, y), \quad x, y \in \mathbb{R},$$

or equivalently by

$$u = e^x \cos y, \quad v = e^x \sin y.$$

We calculate

$$\begin{aligned} \mathfrak{J}\,\boldsymbol{F}(x, y) &= \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}, \\ \det \mathfrak{J}\,\boldsymbol{F}(x, y) &= e^{2x} \neq 0. \end{aligned}$$

Hence, by the inverse function theorem, an inverse exists, but only locally. For example we could take

$$x = \frac{1}{2}\log(u^2 + v^2), \quad y = \tan^{-1}(v/u), \quad u > 0.$$

In this way we obtain a diffeomorphism $\boldsymbol{F} : U \to V$ of the planar regions

$$\begin{aligned} U &= \{\boldsymbol{a} \in \mathbb{R}^2 : -\pi/2 < a_2 < \pi/2\}, \\ V &= \boldsymbol{F}(U) = \{\boldsymbol{a} \in \mathbb{R}^2 : a_1 > 0\}. \end{aligned}$$

Speaking informally, a system of n-dimensional coordinates is a way of assigning n numbers to points in n-dimensional space in such a way that (1) every point has a unique "address" and so that (2) we can take partial derivatives of functions with respect to the coordinates in question.

Thus, we require n functions $y^i = F^i(x^1, \ldots, x^n),\ i = 1, \ldots, n$ where the domain U is restricted in such a way that the mapping

$$\boldsymbol{a} \mapsto \boldsymbol{F}(\boldsymbol{a}), \quad \boldsymbol{a} \in U$$

is one-to-one. Condition (2) requires that the $n \times n$ partial derivatives $\left(\partial x^i / \partial y^j\right)_{ij=1}^n$ be defined and continuous. Since the $n \times n$ matrix in question is the inverse of the matrix $\left(\partial x^i / \partial y^j\right)_{ij}$, our requirement is that the Jacobian matrix $\mathfrak{J}\boldsymbol{F}$ be invertible for all points in U.

In other words, a coordinate system requires the same information as a diffeomorphism, but this information is interpreted differently. We regard a diffeomorphism as a transformation of position. However, a change of coordinates does not alter position, but rather transforms the *address* of the position.

In this regard, $\mathbb{R}^n$ is equipped with a special coordinate system that corresponds to the identity diffeomorphism. These are the *standard coordinates*, which we will denote by $x^1, \ldots, x^n$. Thus x^1 is a function that maps an element of $\mathbb{R}^n$ to its first component, etc. However, in deference to tradition we will use $x = x^1, y = x^2$ to denote the standard coordinates of $\mathbb{R}^2$ and use $x = x^1, y = x^2, z = x^3$ to denote the standard coordinates of $\mathbb{R}^3$.

Example 3.10. Consider the usual polar coordinates r, θ, where

$$x = r\cos\theta, \quad y = r\sin\theta. \tag{25}$$

If we write $(x, y) = \boldsymbol{G}(r, \theta)$, we define a transformation $\boldsymbol{G} : \mathbb{R}^2 \to \mathbb{R}^2$, but that transformation fails to be a diffeomorphism; it is not one-to-one. In order to assign a unique polar coordinate address to every point we must solve for r, θ. This requires us to restrict the domain, somehow. One way to do this is to set

$$r = \sqrt{x^2 + y^2}, \quad \theta = \operatorname{Arg}(x + iy). \tag{26}$$

Here Arg is the principal argument function whose domain is the complex plane minus the non-positive x-axis; i.e. the set

$$U = \{(a_1, a_2) \in \mathbb{R}^2 : a_1 > 0 \text{ or } a_2 \neq 0\}. \tag{27}$$

If we now write $(r, \theta) = \boldsymbol{F}(x, y)$, we define a diffeomorphism $\boldsymbol{F} : U \to V$ where

$$V = \{(a_1, a_2) \in \mathbb{R}^2 : a_1 > 0,\ -\pi < a_2 < \pi\}. \tag{28}$$

Other choices of domain are possible, of course. Each of these choices corresponds to a distinct diffeomorphism.

4. Transformation of vector fields.

Let $\boldsymbol{F} : U \to V$ be a diffeomorphism of domains $U, V \subset \mathbb{R}^n$. Let $y^j = F^j(x^1, \ldots, x^n)$ be the corresponding change of coordinates. Let's consider the meaning of the symbols $\partial y^j/\partial x^i$ and $\partial x^i/\partial y^j$. For the former, we write

$$\frac{\partial y^j}{\partial x^i} = \frac{\partial F^j(x^1, \ldots, x^n)}{\partial x^i} = (\mathrm{D}_i F^j)(x^1, \ldots, x^n).$$

Thus, $\partial y^j/\partial x^i$ is the same thing as the function $\mathrm{D}_i F^j : U \to \mathbb{R}$. Let $\boldsymbol{G} : V \to U$ be the inverse transformation. Since $x^i = G^i(y^1, \ldots, y^n)$, we would like to be able to write

$$\begin{aligned} \frac{\partial x^i}{\partial y^j} &= \frac{\partial G^i(y^1, \ldots, y^n)}{\partial y^j} \\ &= (\mathrm{D}_j G^i)(y^1, \ldots, y^n) \\ &= (\mathrm{D}_j G^i \circ \boldsymbol{F})(x^1, \ldots, x^n). \end{aligned} \tag{29}$$

Definition 4.1. *To that end, we define $\partial/\partial y^j$, $j = 1, \ldots, n$ to be the vector field on the domain U whose component in the $\mathbf{e}_i$ direction is the function $\mathrm{D}_j G^i \circ \boldsymbol{F} : U \to \mathbb{R}$; in other words,*

$$\frac{\partial}{\partial y^j} = \sum_{i=1}^{n} \frac{\partial x^i}{\partial y^j} \frac{\partial}{\partial x^i}, \quad j = 1, \ldots, n, \tag{30}$$

where $\partial x^i/\partial y^j$ is the function shown in (29). *We call these the* fundamental vector fields *corresponding to the coordinates $y^1, \ldots, y^n$. Given a function $u = f(x^1, \ldots, x^n)$, we define $\partial u/\partial y^j$ to be the directional derivative of f with respect to the j^{th} fundamental vector field.*

Using the above definitions we can express the multi-variable chain rule in the familiar fashion, namely

$$\frac{\partial u}{\partial y^j} = \sum_{i=1}^{n} \frac{\partial u}{\partial x^i} \frac{\partial x^i}{\partial y^j}, \quad j = 1, \ldots, n. \tag{31}$$

Example 4.2. Let $(r, \theta) = \boldsymbol{F}(x, y)$ be the transformation (26) from Cartesian to polar coordinates. Let us calculate the fundamental vector fields $\partial/\partial r$ and $\partial/\partial\theta$. We calculate

$$\begin{aligned}
\frac{\partial x}{\partial r} &= \cos\theta = \frac{x}{r} = \frac{x}{\sqrt{x^2+y^2}},\\
\frac{\partial y}{\partial r} &= \sin\theta = \frac{y}{r} = \frac{y}{\sqrt{x^2+y^2}},\\
\frac{\partial}{\partial r} &= \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y}\\
&= \frac{x}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{\partial x}{\partial \theta} &= -r\sin\theta = -y,\\
\frac{\partial y}{\partial \theta} &= r\cos\theta = x,\\
\frac{\partial}{\partial \theta} &= \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y}\\
&= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.
\end{aligned}$$

Proposition 4.3. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and $y^j = F^j(x^1, \ldots, x^n)$ the corresponding change of coordinates. Then,*

$$\frac{\partial}{\partial x^i} = \sum_{j=1}^{n} \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j}, \quad i = 1, \ldots, n. \tag{32}$$

Proof. Let $\boldsymbol{G} : V \to U$ be the inverse transformation. By definition, $\boldsymbol{G} \circ \boldsymbol{F}$ is the identity transformation. The Jacobian matrix of the identity transformation is the identity matrix. Hence, by the chain rule

$$\mathfrak{J}\boldsymbol{F} = \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \cdots & \frac{\partial y^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y^n}{\partial x^1} & \cdots & \frac{\partial y^n}{\partial x^n} \end{pmatrix} \quad \text{and} \quad \mathfrak{J}\boldsymbol{G} \circ \boldsymbol{F} = \begin{pmatrix} \frac{\partial x^1}{\partial y^1} & \cdots & \frac{\partial x^1}{\partial y^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial y^1} & \cdots & \frac{\partial x^n}{\partial y^n} \end{pmatrix}$$

are inverse matrices. We obtain (32) by inverting the relations (30). $\square$

Proposition 4.4. *The fundamental vector field $\partial/\partial y^j$, $j = 1, \ldots, n$ is the unique vector field $\boldsymbol{A} : U \to \mathbb{R}^n$ that satisfies*

$$\boldsymbol{A}[y^k] = \delta^k_j = \begin{cases} 1 & k = j, \\ 0 & k \neq j \end{cases} \tag{33}$$

Proof. By the definition (30), we have

$$\frac{\partial y^k}{\partial y^j} = \sum_{i=1}^{n} \frac{\partial x^i}{\partial y^j} \frac{\partial y^k}{\partial x^i}.$$

Above, we established that the matrix of partials $\partial x^i/\partial y^j$ is inverse to the matrix $\partial y^k/\partial x^i$. This establishes that the matrix formed by the partials $\partial y^k/\partial y^j$ is the $n \times n$ identity matrix.

To prove the converse consider a vector field

$$\boldsymbol{A} = \sum_{i=1}^{n} A^i(x^1, \ldots, x^n) \frac{\partial}{\partial x^i},$$

such that (33) holds. Hence, the functions A^i, $i = 1, \ldots, n$ satisfies the system of equations

$$\begin{pmatrix} \dfrac{\partial y^1}{\partial x^1} & \cdots & \dfrac{\partial y^1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y^n}{\partial x^1} & \cdots & \dfrac{\partial y^n}{\partial x^n} \end{pmatrix} \begin{pmatrix} A^1 \\ \vdots \\ A^n \end{pmatrix} = \mathbf{e}_j.$$

Using the inverse matrix of partials to solve the above gives

$$A^i(x^1, \ldots, x^n) = \frac{\partial x^i}{\partial y^j}.$$

□

Note: δ^i_j is known as the Kronecker symbol.

Example 4.5. Returning to the example of polar coordinates, we have

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} \qquad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$$
$$\frac{\partial \theta}{\partial x} = \frac{-y}{x^2 + y^2}, \qquad \frac{\partial \theta}{\partial y} = \frac{x}{x^2 + y^2}.$$

Note: here we used the fact that

$$\theta = \operatorname{Arg}(x + iy) = \tan^{-1}(y/x) + \text{const.}$$

Observe that the following matrices are inverses:

$$
\begin{aligned}
\mathcal{J}\boldsymbol{F}(x,y) &= \begin{pmatrix} \dfrac{\partial r}{\partial x} & \dfrac{\partial r}{\partial y} \\ \dfrac{\partial \theta}{\partial x} & \dfrac{\partial \theta}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ \dfrac{-y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}, \\
(\mathcal{J}\boldsymbol{G}\circ\boldsymbol{F})(x,y) &= \begin{pmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & -y \\ \dfrac{y}{\sqrt{x^2+y^2}} & x \end{pmatrix},
\end{aligned}
$$

Previously, we showed that

$$
\begin{pmatrix} \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} \end{pmatrix} \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & -y \\ \dfrac{y}{\sqrt{x^2+y^2}} & x \end{pmatrix}
$$

Right-multiplying the above relation by the inverse matrix gives

$$
\begin{pmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} \end{pmatrix} \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ \dfrac{-y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}.
$$

Consequently, we have

$$
\begin{aligned}
\frac{\partial r}{\partial r} &= \frac{\partial}{\partial r}\left(\sqrt{x^2+y^2}\right) \\
&= \left(\frac{x}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y}\right)\left(\sqrt{x^2+y^2}\right) \\
&= 1.
\end{aligned}
$$

Similar calculations show that

$$\begin{aligned}
\frac{\partial\theta}{\partial r} &= \left(\frac{x}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y}\right)\left(\tan^{-1}(y/x)\right) \\
&= 0, \\
\frac{\partial r}{\partial\theta} &= \frac{\partial}{\partial\theta}\left(\sqrt{x^2+y^2}\right) \\
&= \left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right)\left(\sqrt{x^2+y^2}\right) \\
&= 0, \\
\frac{\partial\theta}{\partial\theta} &= \left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right)\left(\tan^{-1}(y/x)\right) \\
&= 1.
\end{aligned}$$

Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field on a domain $U \subset \mathbb{R}^n$. The standard vector fields $\partial/\partial x^i = \mathbf{e}_i,\ i = 1,\ldots,n$ are a basis, and hence $\boldsymbol{A}$ has a unique expression of the form

$$\boldsymbol{A} = \sum_{i=1}^{n} A^i(x^1,\ldots,x^n)\frac{\partial}{\partial x^i}.$$

Now let $y^j = F^j(x^1,\ldots,x^n)$ be another system of coordinates on the same domain. Since the fundamental vector fields $\partial/\partial y^j,\ j = 1,\ldots,n$ are also a basis, we can also express the given vector field as

$$\boldsymbol{A} = \sum_{j=1}^{n} B^j(y^1,\ldots,y^n)\frac{\partial}{\partial y^j}.$$

Definition 4.6. *We call the functions $B^j : V \to \mathbb{R}$ the components of $\boldsymbol{A}$ relative to coordinates $y^1,\ldots,y^n$.*

In order to understand the geometry underlying the transformation from components $A^i : U \to \mathbb{R}$ to the components $B^j : V \to \mathbb{R}$, we need to introduce the push-forward transformation.

Definition 4.7. *Let $\boldsymbol{F} : U \to V$ be a C^2 diffeomorphism, and let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field. We call the C^1 vector field $\boldsymbol{F}_*\boldsymbol{A} : V \to \mathbb{R}^n$, defined by*

$$\boldsymbol{F}_*\boldsymbol{A} = (\mathfrak{J}\,\boldsymbol{F}\cdot\boldsymbol{A})\circ\boldsymbol{F}^{-1}. \tag{34}$$

the push-forward of $\boldsymbol{A}$.

Note: the dot in $\mathfrak{J}\,\boldsymbol{F}\cdot\boldsymbol{A}$ refers to matrix vector multiplication.

Proposition 4.8. *Let $\boldsymbol{F}, \boldsymbol{A}$ be as above, and let $\boldsymbol{B} = \boldsymbol{F}_*\boldsymbol{A}$. Let $A^i : U \to \mathbb{R}$ and $B^j : V \to R$ be the component functions of the vector fields $\boldsymbol{A}$ and $\boldsymbol{B}$. Letting $y^j = F^j(x^1, \dots, x^n)$, we have*

$$\boldsymbol{A} = \sum_{i=1}^{n} A^i(x^1, \dots, x^n)\frac{\partial}{\partial x^i} = \sum_{j=1}^{n} B^j(y^1, \dots, y^n)\frac{\partial}{\partial y^j}. \tag{35}$$

In particular, if $\boldsymbol{G} : V \to U$ denotes the inverse diffeomorphism and $\partial/\partial\hat{x}^j$ denotes to the constant vector field in direction $\mathbf{e}_j$ on the domain V, then

$$\frac{\partial}{\partial y^j} = \boldsymbol{G}_*\frac{\partial}{\partial \hat{x}^j}, \quad j = 1, \dots, n,$$

Example 4.9. Let $(r, \theta) = \boldsymbol{F}(x, y)$, where $\boldsymbol{F} : U \to V$ as before. Above, we showed that

$$\begin{aligned}
\frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} \\
&= \frac{x}{\sqrt{x^2 + y^2}}\frac{\partial}{\partial r} - \frac{y}{x^2 + y^2}\frac{\partial}{\partial \theta} \\
&= \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}.
\end{aligned}$$

Letting $\hat{x}, \hat{y}$ denote the standard coordinates on the domain V, we have

$$\boldsymbol{F}_*\frac{\partial}{\partial x} = \cos\hat{y}\frac{\partial}{\partial \hat{x}} - \frac{\sin\hat{y}}{\hat{x}}\frac{\partial}{\partial \hat{y}}.$$

Let us calculate this result directly. Using our previous calculations,

$$(\partial\boldsymbol{F} \cdot \mathbf{e}_1)(x, y) = \begin{pmatrix} \dfrac{x}{\sqrt{x^2 + y^2}} \\ \dfrac{-y}{x^2 + y^2} \end{pmatrix}.$$

Hence,

$$\begin{aligned}
((\partial\boldsymbol{F} \cdot \mathbf{e}_1) \circ \boldsymbol{G})(\hat{x}, \hat{y}) &= (\partial\boldsymbol{F} \cdot \mathbf{e}_1)(\hat{x}\cos(\hat{y}), \hat{x}\sin(\hat{y})) \\
&= \begin{pmatrix} \cos(\hat{y}) \\ -\dfrac{\sin(\hat{y})}{\hat{x}} \end{pmatrix} \\
&= \cos(\hat{y})\,\mathbf{e}_1 - \frac{\sin(\hat{y})}{\hat{x}}\,\mathbf{e}_2,
\end{aligned}$$

as was to be shown.

Thus, the push-forward of a vector field, or what is the same a change of coordinates, requires the knowledge of both the forward and the inverse diffeomorphism. Below we prove that only the knowledge of the inverse diffeomorphism is required, provided we are willing to invert a matrix.

Proposition 4.10. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and $\boldsymbol{G} = \boldsymbol{F}^{-1} : V \to U$ its inverse. If $\boldsymbol{A} : U \to \mathbb{R}^n$ is a vector field, then*

$$\boldsymbol{F}_* \boldsymbol{A} = (\mathcal{J}\, \boldsymbol{G})^{-1} \cdot (\boldsymbol{A} \circ \boldsymbol{G}). \tag{36}$$

Proof. The above formula follows from the definition of $\boldsymbol{F}_* \boldsymbol{A}$ and from fact that $\mathcal{J}\, \boldsymbol{G}$ and $\mathcal{J}\, \boldsymbol{F} \circ \boldsymbol{G}$ are inverse matrices. □

Example 4.11. Let us introduce two-dimensional parabolic coordinates (u, v) by setting

$$x = u^2 - v^2, \quad y = 2uv. \tag{37}$$

Let us call the forward transformation $(u, v) = \boldsymbol{F}(x, y)$ and the inverse

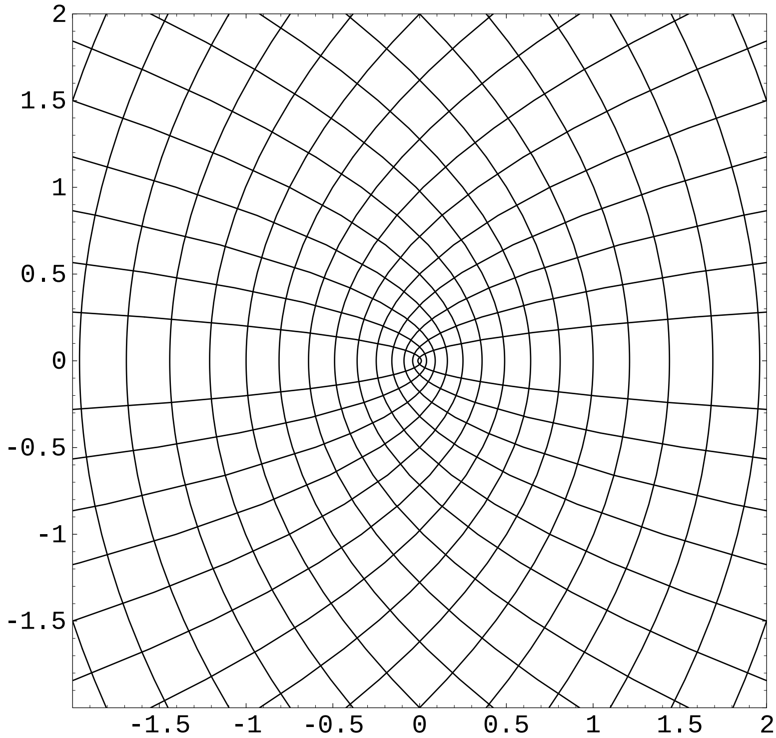


FIGURE 1. Parabolic coordinates

$(x, y) = \boldsymbol{G}(u, v)$. We have explicit formulas for $\boldsymbol{G}$, but explicit formulas for $\boldsymbol{F}$ would be "messy" and involve radicals. In any case, $\boldsymbol{G}$ is more convenient for generating the coordinate grid for the u, v coordinates. Half of the grid lines in the above figure were generated by setting $u = u_0$ treating v as a parameter, and then plotting the corresponding curve $(x, y) = \boldsymbol{G}(u_0, v)$. The other grid lines were generated by plotting $(x, y) = \boldsymbol{G}(u, v_0)$ for various constant values $v = v_0$.

In order to express a vector field using parabolic coordinates it would seem that we find explicit expressions for $\boldsymbol{F}$. However, we can avoid this by using (36). For example, let us express $x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$ using parabolic coordinates. We have

$$\eth\, \boldsymbol{G}(u, v) = \begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix} = 2\begin{pmatrix} u & -v \\ v & u \end{pmatrix}.$$

Taking the inverse of the above matrix, we have

$$\eth\, \boldsymbol{F}(x, y) = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \dfrac{u}{u^2+v^2} & \dfrac{v}{u^2+v^2} \\ \dfrac{-v}{u^2+v^2} & \dfrac{u}{u^2+v^2} \end{pmatrix} \tag{38}$$

Hence,

$$\begin{aligned} \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) &= \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} \\ &= \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \begin{pmatrix} \dfrac{u}{2(u^2+v^2)} & \dfrac{v}{2(u^2+v^2)} \\ \dfrac{-v}{2(u^2+v^2)} & \dfrac{u}{2(u^2+v^2)} \end{pmatrix} \end{aligned}$$

We can now switch to parabolic coordinates.

$$\begin{aligned} x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial x} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix} \begin{pmatrix} u^2 - v^2 \\ 2uv \end{pmatrix} \\ &= \frac{1}{2}\left(u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}\right). \end{aligned}$$

Note that we didn't write $\eth\, \boldsymbol{F}$ in terms of x, y, because we don't have formulas for u and v in terms of x and y. We therefore express all our calculations in terms of u and v.

5. Covariance

The principle of covariance states that any operation defined in terms of intrinsic geometry can be carried out in any system of coordinates with suitable transformation laws relating the operation and the answers. Let's consider how the principle of covariance manifests in the case of directional derivatives.

First, let us consider how functions transform under a change of coordinates.

Definition 5.1. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and let $g : V \to \mathbb{R}$ be a function of n variables. The function $f : U \to \mathbb{R}$ defined by*

$$f = g \circ \boldsymbol{F}. \tag{39}$$

is called the pull-back of g by $\boldsymbol{F}$ and is denoted by $f = \boldsymbol{F}^* g$.

Setting $y^j = F^j(x^1, \dots, x^n)$ allows us to express (39) succinctly as

$$f(x^1, \dots, x^n) = g(y^1, \dots, y^n).$$

Therefore, the pull-back transformation is the geometric description of the process of transforming a function by making a change of variables.

Recall that the directional derivative operation combines a C^1 vector field and a C^2 function to produce a C^1 function. For example, given vector field $\boldsymbol{A} : U \to \mathbb{R}^n$ and function $f : U \to \mathbb{R}$ the directional derivative gives us the function $\boldsymbol{A}[f] : U \to \mathbb{R}$. The push-forward and the pull-back are adjoint operations relative to the directional derivative bracket; pushing forward a vector field is, in a sense, equivalent to pulling back a function.

Proposition 5.2. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field, $\boldsymbol{F} : U \to V$ a C^2 diffeomorphism, and $g : V \to R$ a C^2 function. Letting $\boldsymbol{B} = \boldsymbol{F}_* \boldsymbol{A}$ and $f = \boldsymbol{F}^* g$, we have*

$$\boldsymbol{F}^*(\boldsymbol{B}[g]) = \boldsymbol{A}[f]. \tag{40}$$

Proof. Setting $\boldsymbol{y} = \boldsymbol{F}(\boldsymbol{x})$, the chain rule gives us

$$\begin{aligned} \frac{\partial f(x^1, \dots, x^n)}{\partial x^i} &= \frac{\partial g(y^1, \dots, y^n)}{\partial x^i} \\ &= \sum_{j=1}^n \frac{\partial g(y^1, \dots, y^n)}{\partial y^j} \frac{\partial y^j}{\partial x^i}, \quad i = 1, \dots, n. \end{aligned}$$

Equivalently,

$$\mathrm{D}_i f = \sum_{j=1}^n (\mathrm{D}_j g \circ \boldsymbol{F}) \, \mathrm{D}_i F^j, \quad i = 1, \dots, n$$

Hence,

$$\begin{aligned}
\boldsymbol{A}[f] &= \sum_{i=1}^{n} A^i \,\mathrm{D}_i f \\
&= \sum_{i,j=1}^{n} A^i \,(\mathrm{D}_j g \circ \boldsymbol{F})\; \mathrm{D}_i F^j \\
&= \sum_{j=1}^{n} \boldsymbol{A}[F^j](\mathrm{D}_j g \circ \boldsymbol{F}) \\
&= \sum_{j=1}^{n} (B^j \circ \boldsymbol{F})(\mathrm{D}_j g \circ \boldsymbol{F}) \\
&= \boldsymbol{B}[g] \circ F,
\end{aligned}$$

as was to be shown. □

The upshot of the above result is that a directional derivative calculated in a different coordinate system agrees with the directional derivative calculated relative to standard coordinates. Indeed, let $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{F}, f, g$ be as in the Proposition. By definition,

$$\boldsymbol{A} = \sum_{i=1}^{n} A^i(x^1, \ldots, x^n) \frac{\partial}{\partial x^i} = \sum_{j=1}^{n} B^j(y^1, \ldots, y^n) \frac{\partial}{\partial y^j},$$

where $A^i : U \to \mathbb{R}$ are the components of $\boldsymbol{A}$ and where $B^j : V \to \mathbb{R}$ are the components of $\boldsymbol{B} = \boldsymbol{F}_* \boldsymbol{A}$. Since $f = F^* g$,

$$f(x^1, \ldots, x^n) = g(y^1, \ldots, y^n)$$

is the same scalar function expressed in two different ways. By definition of directional derivative,

$$\begin{aligned}
\boldsymbol{A}[f] &= \sum_{i=1}^{n} A^i \,\mathrm{D}_i f, \\
\boldsymbol{B}[g] &= \sum_{j=1}^{n} B^j \,\mathrm{D}_j g.
\end{aligned}$$

The Proposition asserts that

$$\boldsymbol{A}[f](x^1, \ldots, x^n) = \boldsymbol{B}[g](y^1, \ldots, y^n),$$

or what is equivalent

$$\sum_{i=1}^{n} A^i(x^1, \ldots, x^n) \frac{\partial f(x^1, \ldots, x^n)}{\partial x^i} = \sum_{j=1}^{n} B^i(y^1, \ldots, y^n) \frac{\partial g(y^1, \ldots, y^n)}{\partial y^j}.$$

Example 5.3. Next, let us illustrate the principle of covariance. Let $f(x,y)$ be a function and set

$$g(u,v) = f(u^2 - v^2, 2uv) = (f \circ \boldsymbol{G})(u,v).$$

In other words, g gives the form of the expression $f(x,y)$ in parabolic coordinates. Geometrically, $f : U \to \mathbb{R}$ and $g = \boldsymbol{G}^* f : V \to \mathbb{R}$ is the pullback function (We will choose suitable domains U, V below). Let us calculate the directional derivative

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \right) [f(x,y)]$$

in both coordinate systems and compare the answers. We obtain

$$\begin{aligned}
&\frac{1}{2}\left(u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v} \right) [g(u,v)] \\
&= \frac{1}{2}\left(u\frac{\partial f(u^2 - v^2, 2uv)}{\partial u} + v\frac{\partial f(u^2 - v^2, 2uv)}{\partial v} \right) \\
&= \big(u^2(\mathrm{D}_1 f)(x,y) + uv(\mathrm{D}_2 f)(x,y) - v^2(\mathrm{D}_1 f)(x,y) + uv(\mathrm{D}_2 f)(x,y)\big) \\
&= \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \right) [f(x,y)],
\end{aligned}$$

as was to be shown.

What happens if there are multiple systems of coordinates, say standard coordinates S, coordinates A and coordinates B? Does a directional derivative calculated in coordinates A agree with the same calculation using coordinates B? The answer must be "yes", because both calculation must agree with the answer obtained using standard coordinates. In this type of situation there are 6 different transformations we must consider: standard coordinates to A, standard coordinates to B, A to B, and all the inverse transformations. Of course a transformation from S to A, followed a transformation from A to B, should be the same thing as the transformation from S to B. This kind of coherence property is our next subject.

Proposition 5.4. *Let $\boldsymbol{F} : U \to V$ and $\boldsymbol{G} : V \to W$ be diffeomorphisms. Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field. Then,*

$$\boldsymbol{G}_*(\boldsymbol{F}_*\boldsymbol{A}) = (\boldsymbol{G} \circ \boldsymbol{F})_*\boldsymbol{A}. \tag{41}$$

Similarly, let $h : W \to \mathbb{R}$ be a function. Then,

$$\boldsymbol{F}^*(\boldsymbol{G}^* h) = (\boldsymbol{G} \circ \boldsymbol{F})^* h. \tag{42}$$

Proof. Homework. □

Example 5.5. Let's consider three coordinate system in the plane: the standard (x, y) coordinates, the polar (r, θ) coordinates and the parabolic (u, v) coordinates. These are related by the following diffeomorphisms:

$$\begin{aligned}(x, y) &= \boldsymbol{G}_1(r, \theta) = (r\cos(\theta), r\sin(\theta)),\\ (u, v) &= \boldsymbol{G}_2(r, \theta) = (\sqrt{r}\cos(\theta/2), \sqrt{r}\sin(\theta/2)),\\ (x, y) &= \boldsymbol{G}_3(u, v) = (u^2 - v^2, 2uv).\end{aligned}$$

Here $\boldsymbol{G}_1 : V \to U$, $\boldsymbol{G}_2 : W \to V$, and $\boldsymbol{G}_2 : W \to U$ are diffeomorphisms of the following domains:

$$\begin{aligned}U &= \{(a_1, a_2) : a_1 > 0 \text{ or } a_2 \neq 0\},\\ V &= \{(a_1, a_2) : a_1 > 0 \text{ and } -\pi < a_2 < \pi\},\\ W &= \{(a_1, a_2) : a_1 > 0\}.\end{aligned}$$

In order for the above definitions to be sound we must have

$$\boldsymbol{G}_1 = \boldsymbol{G}_3 \circ \boldsymbol{G}_2.$$

The following calculation verifies this. Throughout, it is convenient to identify vectors in $\mathbb{R}^2$ with complex numbers. This allows us to write

$$\begin{aligned}\boldsymbol{G}_2(r, \theta) &= \sqrt{r}\exp(\mathrm{i}\,\theta/2),\\ \boldsymbol{G}_3(u, v) &= (u + \mathrm{i}\,v)^2.\end{aligned}$$

Hence,

$$\begin{aligned}(\boldsymbol{G}_3 \circ \boldsymbol{G}_2)(r, \theta) &= \boldsymbol{G}_3(\sqrt{r}\exp(\mathrm{i}\,\theta/2))\\ &= r\exp(\mathrm{i}\,\theta)\\ &= \boldsymbol{G}_1(r, \theta).\end{aligned}$$

Next, consider a function $f(x, y)$. The same function can be expressed using polar coordinates as

$$f(x, y) = g_1(r, \theta)$$

where $g_1 = \boldsymbol{G}_1^* f$, because

$$f(x, y) = f(r\cos\theta, r\sin\theta) = (f \circ \boldsymbol{G}_1)(r, \theta) = g_1(r, \theta).$$

Similarly, to express $f(x, y)$ in parabolic coordinates we use $g_2 = \boldsymbol{G}_3^* f$ because

$$g_2(u, v) = (f \circ \boldsymbol{G}_2)(u, v) = f(u^2 - v^2, 2uv) = f(x, y).$$

However, we can also express $g_2(u, v)$ as a function of r, θ by setting

$$g_2(u, v) = g_3(r, \theta),$$

where $g_3 = \boldsymbol{G}_2^* g_2$ because

$$g_3(r,\theta) = (g_2 \circ \boldsymbol{G}_3)(r,\theta) = g_2(\sqrt{r}\exp(\mathrm{i}\,\theta/2)) = g_2(u,v).$$

However all of the above are equal to $f(x,y)$ and so $g_3 = g_1$, or to put it another way

$$\boldsymbol{G}_3^*(\boldsymbol{G}_2^* f) = \boldsymbol{G}_1^* f.$$

That's exactly what (42) asserts.

Similarly, if we start with a given vector field

$$\boldsymbol{A} = A^1(x,y)\frac{\partial}{\partial x} + A^2(x,y)\frac{\partial}{\partial y},$$

we can write it using polar coordinates as

$$\boldsymbol{A} = B^1(r,\theta)\frac{\partial}{\partial r} + B^2(r,\theta)\frac{\partial}{\partial \theta},$$

or using parabolic coordinates:

$$\boldsymbol{A} = C^1(u,v)\frac{\partial}{\partial u} + C^2(u,v)\frac{\partial}{\partial v}.$$

Here $\boldsymbol{B} = (\boldsymbol{F}_1)_*\boldsymbol{A}$ and $\boldsymbol{C} = (\boldsymbol{F}_2)_*\boldsymbol{A}$, where $\boldsymbol{F}_1 = (\boldsymbol{G}_1)^{-1}$ and $\boldsymbol{F}_2 = (\boldsymbol{G}_2)^{-1}$ are the inverse diffeomorphisms. However, a vector field expressed in parabolic coordinates can be rewritten using polar coordinates by making use of the $\boldsymbol{F}_3 = (\boldsymbol{G}_3)^{-1}$ transformation; in other words,we also have $\boldsymbol{B} = (\boldsymbol{F}_3)_*\boldsymbol{C}$. Thus, we have two different ways of calculating $B^1(r,\theta)$ and $B^2(r,\theta)$. The above Proposition asserts that both calculations yield the same answer.

Let's do a specific calculation of this type. Take $\boldsymbol{A} = \partial/\partial x$. Earlier, we established that

$$\begin{aligned}
\frac{\partial}{\partial x} &= \frac{\partial u}{\partial x}\frac{\partial}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial}{\partial v} \\
&= \frac{u}{2(u^2+v^2)}\frac{\partial}{\partial u} - \frac{v}{2(u^2+v^2)}\frac{\partial}{\partial v}, \\
&= \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right)\begin{pmatrix} \dfrac{u}{2(u^2+v^2)} \\ -\dfrac{v}{2(u^2+v^2)} \end{pmatrix}.
\end{aligned}$$

We also have

$$\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right) = \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \begin{pmatrix} \dfrac{\partial u}{\partial r} & \dfrac{\partial u}{\partial \theta} \\ \dfrac{\partial v}{\partial r} & \dfrac{\partial v}{\partial \theta} \end{pmatrix}$$

$$= \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \begin{pmatrix} \frac{\cos(\theta/2)}{2\sqrt{r}} & -\frac{\sqrt{r}}{2}\sin(\theta/2) \\ \frac{\sin(\theta/2)}{2\sqrt{r}} & \frac{\sqrt{r}}{2}\cos(\theta/2) \end{pmatrix}$$

Taking inverses gives

$$\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) = \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right) \begin{pmatrix} 2\sqrt{r}\cos(\theta/2) & 2\sqrt{r}\sin(\theta/2) \\ -2\sin(\theta/2)/\sqrt{r} & 2\cos(\theta/2)/\sqrt{r} \end{pmatrix}$$

Combining the above relations and switching to the r, θ variables gives

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{1}{2r}\left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}\right) \begin{pmatrix} \sqrt{r}\cos(\theta/2) \\ -\sqrt{r}\sin(\theta/2) \end{pmatrix} \\ &= \frac{1}{2r}\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right) \begin{pmatrix} 2\sqrt{r}\cos(\theta/2) & 2\sqrt{r}\sin(\theta/2) \\ -2\sin(\theta/2)/\sqrt{r} & 2\cos(\theta/2)/\sqrt{r} \end{pmatrix} \begin{pmatrix} \sqrt{r}\cos(\theta/2) \\ -\sqrt{r}\sin(\theta/2) \end{pmatrix} \\ &= \left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right) \begin{pmatrix} \cos(\theta) \\ -\sin(\theta)/r \end{pmatrix}. \end{aligned}$$

However, we could have calculated the above directly by using the transformation from polar to Cartesian coordinates:

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} \\ &= \cos(\theta)\frac{\partial}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial}{\partial \theta} \end{aligned}$$

The two answers agree.

6. Coordinate transformations and ODEs

Consider a system of autonomous ODEs

$$\dot{x}^i = A^i(x^1, \ldots, x^n), \quad \boldsymbol{x} \in U. \tag{43}$$

defined on some domain $U \subset \mathbb{R}^n$. Let $\boldsymbol{F} : U \to V$ be a diffeomorphism, and let $y^j = F^j(x^1, \ldots, x^n)$ be the corresponding system of coordinates on U. We wish to write a system

$$\dot{y}^j = B^j(y^1, \ldots, y^n), \quad \boldsymbol{y} \in V \tag{44}$$

that is in some sense equivalent to the given system (43). Since

$$\begin{aligned} \frac{dy^j}{dt} &= \sum_{i=1}^n \frac{\partial y^j}{\partial x^i} \frac{dx^i}{dt} \\ &= \sum_{i=1}^n \frac{\partial y^j}{\partial x^i} A^i(x^1, \ldots, x^n) \\ &= B^j(y^1, \ldots, y^n) \end{aligned}$$

we recognize that ODEs are governed by the same transformation laws as vector fields. Expressing the system (43) relative to the y^j coordinates is equivalent to expressing the vector field

$$\boldsymbol{A} = \sum_{i=1}^n A^i(x^1, \ldots, x^n) \frac{\partial}{\partial x^i} = \sum_{j=1}^n B^j(y^1, \ldots, y^n) \frac{\partial}{\partial y^j}$$

relative to the y^j coordinate system; the calculation the push-forward $\boldsymbol{B} = \boldsymbol{F}_* \boldsymbol{A}$ is an equivalent process.

However we interpret the transformation from (43) to (44), it can only be meaningful if the solutions of the two systems are in some sense equivalent. We need a transformation law for solutions.

Let $\Phi(t, \boldsymbol{x})$ denote the flow generated by $\boldsymbol{A}$ and let us write the solution to (43) as

$$\hat{x}^i = \Phi^i(t, x^1, \ldots, x^n), \quad i = 1, \ldots, n.$$

Thus, $\hat{x}^1, \ldots, \hat{x}^n$ are functions on the flow domain $\hat{U} \subset \mathbb{R} \times U$ that satisfy

$$\frac{\partial \hat{x}^i}{\partial t} = A^i(\hat{x}^1, \ldots, \hat{x}^n), \quad i = 1, \ldots, n. \tag{45}$$

Similarly, let $\Psi(t, \boldsymbol{y})$ denote the flow generated by $\boldsymbol{B}$ and let

$$\hat{y}^j = \Psi^j(t, y^1, \ldots, y^n), \quad j = 1, \ldots, n \tag{46}$$

express the general solution to (44). Since $y^j = F^j(x^1, \dots, x^n)$ it makes sense that compatibility of the two solution be the condition

$$\hat{y}^j = F^j(\hat{x}^1, \dots, \hat{x}^n). \tag{47}$$

The above condition can be expressed in a component-free language as

$$\boldsymbol{F} \circ \Phi_t = \Psi_t \circ \boldsymbol{F}, \tag{48}$$

where, as before, $\Phi_t(\boldsymbol{x}) = \Phi(t, \boldsymbol{x})$ and $\Psi_t(\boldsymbol{y}) = \Psi(t, \boldsymbol{y})$.

Definition 6.1. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and let $\Phi(t, \boldsymbol{x})$ be a flow defined on U. We define the push-forward $\Psi = \boldsymbol{F}_*\Phi$ to be the mapping defined by*

$$\Psi_t = \boldsymbol{F} \circ \Phi_t \circ \boldsymbol{F}^{-1}. \tag{49}$$

The domain of Ψ is the set

$$\hat{V} = \{(a_0, a_1, \dots, a_n) \in \mathbb{R}^{n+1} : (a_0, \boldsymbol{F}^{-1}(a_1, \dots, a_n)) \in \hat{U}\}, \tag{50}$$

where $\hat{U}$ is the domain of Φ.

Note that with the above definition, the push-forward flow satisfies condition (48).

Proposition 6.2. *Let $\boldsymbol{F}, \Phi, \Psi$ be as above. With the above definition, Ψ is a flow.*

Proof. Let $\boldsymbol{G} = \boldsymbol{F}^{-1}$ denote the inverse transformation. We have

$$\Psi_0(\boldsymbol{y}) = \boldsymbol{F}(\Phi_0(\boldsymbol{G}(\boldsymbol{y}))) = \boldsymbol{F}(\boldsymbol{G}(\boldsymbol{y})) = \boldsymbol{y}, \quad \boldsymbol{y} \in V.$$

As well,

$$\begin{aligned}
\Psi_{t_1+t_2} &= \boldsymbol{F} \circ \Phi_{t_1+t_2} \circ \boldsymbol{G} \\
&= \boldsymbol{F} \circ \Phi_{t_1} \circ \Phi_{t_2} \circ \boldsymbol{G} \\
&= \boldsymbol{F} \circ \Phi_{t_1} \circ \boldsymbol{G} \circ \boldsymbol{F} \circ \Phi_{t_2} \circ \boldsymbol{G} \\
&= \Psi_{t_1} \circ \Psi_{t_2}.
\end{aligned}$$

The various domain requirements of a flow are left as exercises for the interested reader. □

Proposition 6.3. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field and let Φ denote the flow generated by $\boldsymbol{A}$. Let $\boldsymbol{F} : U \to V$ be a diffeomorphism, and set $\boldsymbol{B} = \boldsymbol{F}_*\boldsymbol{A}$ and $\Psi = \boldsymbol{F}_*\Phi$. Then, Ψ is the flow generated by $\boldsymbol{B}$.*

Proof. By the chain rule and by Proposition 1.11,

$$\begin{aligned}\frac{\partial \Psi(t,\boldsymbol{y})}{\partial t} &= \frac{\partial \boldsymbol{F}(\Phi(t,\boldsymbol{x}))}{\partial t}\\ &= \mathcal{J}\,\boldsymbol{F}(\Phi(t,\boldsymbol{x}))\cdot\frac{\partial \Phi(t,\boldsymbol{x})}{\partial t}\\ &= \mathcal{J}\,\boldsymbol{F}(\Phi(t,\boldsymbol{x}))\cdot \boldsymbol{A}(\Phi(t,\boldsymbol{x}))\end{aligned}$$

By definition of the push-forward of a vector field,

$$\boldsymbol{B}(\boldsymbol{y}) = \mathcal{J}\,\boldsymbol{F}(\boldsymbol{x})\cdot\boldsymbol{A}(\boldsymbol{x}).$$

Hence,

$$(\mathrm{D}_1\Psi)(0,\boldsymbol{y}) = \mathcal{J}\,\boldsymbol{F}(\boldsymbol{x})\cdot\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{y}),$$

as was to be shown. □

Example 6.4. Let r,θ be the usual polar coordinates. Earlier, we showed that

$$\frac{\partial}{\partial x} = \cos(\theta)\frac{\partial}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial}{\partial \theta}$$

In other words, the system

$$\begin{aligned}\dot r &= \cos(\theta)\\ \dot\theta &= -\sin(\theta)/r\end{aligned}\tag{51}$$

is equivalent to the system

$$\begin{aligned}\dot x &= 1,\\ \dot y &= 0.\end{aligned}$$

The general solution, in the form of a flow, of the latter system is

$$\Phi(t,x,y) = (x+t,y).$$

To obtain the solution of the system in polar coordinates, call it $\Psi(t,r,\theta)$ we have to calculate the push-forward of Φ relative to the transformation $(r,\theta) = \boldsymbol{F}(x,y)$. We do so by writing

$$\begin{aligned}\Psi(t,r,\theta) &= \boldsymbol{F}(\Phi(t,r\cos\theta,r\sin\theta))\\ &= \boldsymbol{F}(r\cos\theta+t,r\sin\theta)\\ &= \left(\sqrt{r^2+2tr\cos\theta+t^2},\tan^{-1}\left(\frac{r\sin(\theta)}{r\cos(\theta)+t}\right)\right)\end{aligned}$$

A straight-forward calculation verifies that the above is, indeed, the general solution to (51). Figure 2 shows the push forward $\boldsymbol{F}_*\boldsymbol{A}$, while Figure 3 displays the push-forward of the Cartesian coordinate grid via $\boldsymbol{F}$; i.e. the grid consisting of the curves $x\cos(y) = C_1$ and the curves $x\sin(y) = C_2$.

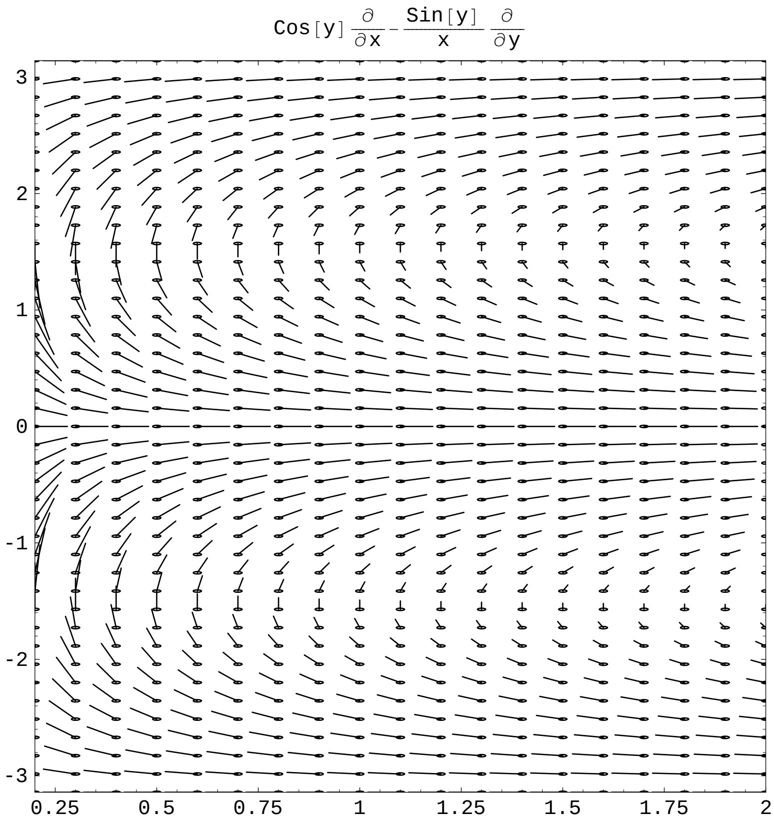


FIGURE 2. Push-forward of $\partial/\partial x$.

7. FIRST INTEGRALS

We have already mentioned the close connection between autonomous ODEs and vector fields. On the other hand, we identify vector fields with first-order differential operators. In this section we discuss first integrals, a geometric idea that ties together solutions of an ODE and the directional derivative operation.

Let us introduce a system of autonomous ODEs

$$\dot{x}^i = A^i(x^1, \ldots, x^n), \quad \boldsymbol{x} \in U.$$

Consider the variation of a scalar function $u = f(x^1, \ldots, x^n)$ along a particular integral curve $\boldsymbol{\phi}(t) \in U$. Setting $\hat{x}^i = \phi^i(t)$, we have

$$\begin{aligned}
\dot{u} &= (f \circ \boldsymbol{\phi})'(t) \\
&= \sum_{i=1}^{n} \mathrm{D}_i f(\hat{x}^1, \ldots, \hat{x}^n) A^i(\hat{x}^1, \ldots, \hat{x}^n) \\
&= (\boldsymbol{A}[u] \circ \boldsymbol{\phi})(t).
\end{aligned}$$

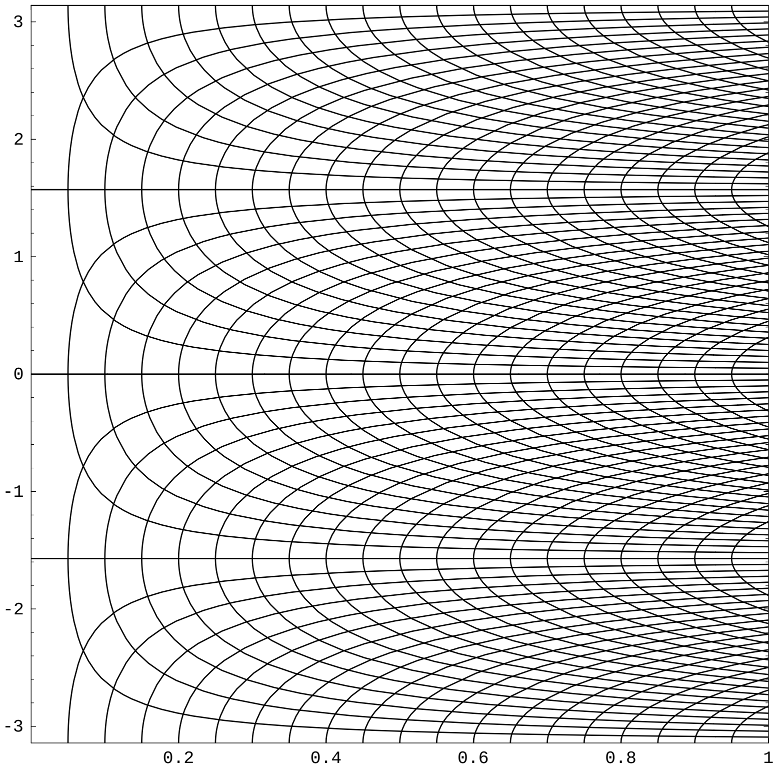


FIGURE 3. Push-forward of Cartesian coordinate grid.

This then is the geometric meaning of the directional derivative: it names the rate of change of a given scalar function along an integral curve of a given vector field.

Definition 7.1. *Let* $\boldsymbol{A} : U \to \mathbb{R}^n$ *be a* C^1 *vector field and* $f : U \to R$ *a* C^2 *function. We call* f *a first integral whenever* $\boldsymbol{A}[f] = 0$.

Proposition 7.2. *A function* $f : U \to R$ *is a first integral of a vector field* $\boldsymbol{A} : U \to \mathbb{R}^n$ *if and only if it is constant along all integral curves of* $\boldsymbol{A}$*. More precisely, if* $\Phi(t, \boldsymbol{x})$ *is the flow generated by* $\boldsymbol{A}$*, then* $f(\boldsymbol{x})$ *is a first integral if and only if*

$$(f \circ \Phi)(t, \boldsymbol{x}) = f(\boldsymbol{x}).$$

Example 7.3. Consider the vector field $\boldsymbol{A} = -y\partial/\partial x + x\partial/\partial y$. The function $f(x, y) = x^2 + y^2$ is a first integral because $\boldsymbol{A}[x^2 + y^2] = 0$ (a simple calculation). Equivalently, the flow generated by $\boldsymbol{A}$ is the function

$$\Phi(t, x, y) = (x \cos t - y \sin t, x \sin t + y \cos t).$$

Observe that

$$\begin{aligned}(f \circ \Phi)(t,x,y) &= (x\cos t - y\sin t)^2 + (x\sin t + y\cos t)^2 \\ &= x^2 + y^2 = f(x,y),\end{aligned}$$

as was to be shown.

Analytic geometry in the plane is based on the notion that a curve is defined by an equation in 2 variables. Analogously, a curve in n dimensions should require $n-1$ independent equation. The following definition and propositions make this more precise.

Definition 7.4. *Let $U \subset \mathbb{R}^n$ be a domain. We say that continuously differentiable functions $f^1,\dots,f^j : U \to \mathbb{R}$ are functionally independent if the $j \times n$ matrix of partial derivatives*

$$\begin{pmatrix} \mathrm{D}_1 f^1 & \dots & \mathrm{D}_n f^1 \\ \vdots & \ddots & \vdots \\ D_1 f^j & \dots & D_n f^j \end{pmatrix} \tag{52}$$

has rank j at all points of U.

Proposition 7.5. *Let $f^1,\dots,f^{j-1} : U \to R$ be C^1 functions and set $f^j = F(f^1,\dots,f^{j-1})$ where F is a C^1 function of $j-1$ variables. Then $f^1,\dots,f^{j-1},f^j$ are functionally dependent. Conversely, if $f^1,\dots,f^j$ are C^1 and functionally dependent, then after rearranging the order of $f^1,\dots,f^j$ and restricting the domain, we have $f^j = F(f^1,\dots,f^{j-1})$ for some C^1 function F of $j-1$ variables.*

Example 7.6. By way of example, let us consider the functions $u = (x+y)/x,\ x > 0$ and $v = \tan^{-1}(y/x),\ x > 0$. The matrix of partial derivatives

$$\begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} -y/x^2 & 1/x \\ -y/(x^2+y^2) & x/(x^2+y^2) \end{pmatrix} \tag{53}$$

is singular; the rank is < 2 everywhere. We can therefore express one function in terms of the other. Indeed,

$$u = \tan(v) + 1.$$

Proposition 7.7. *Let $f^1,\dots,f^{n-1} : U \to \mathbb{R}$ be functionally independent, continuously differentiable functions. Fix a point $(a^1,\dots,a^n) \in U$ and set $C^i = f^i(a^1,\dots,a^n),\ i = 1,\dots,n-1$. Then, there exists a C^1 curve $\hat{x}^i = \phi^i(t)$ that passes through $(a^1,\dots,a^n)$ and satisfies the equations*

$$f^i(\hat{x}^1,\dots,\hat{x}^n) = C^i, \quad i = 1,\dots,n-1.$$

Every other such curve is related to $\boldsymbol{\phi}(t)$ by a reparameterization.

The proof of the above Propositions relies on the implicit function theorem; we will not present it here. However, using the inverse function theorem, we can prove the following.

Proposition 7.8. *Let $U \subset \mathbb{R}^n$ be an n-dimensional domain and $F^1, \ldots, F^n : U \to \mathbb{R}$ functionally independent, C^1 functions. Then, after suitably restricting the domain to $U_1 \subset U$ the mapping $\boldsymbol{F} : U_1 \to V$, where $V = \boldsymbol{F}(U_1) \subset \mathbb{R}^n$, is a diffeomorphism.*

The upshot is that n functionally independent functions define a system of coordinates (provided the domain is suitably restricted).

In light of Proposition 7.7, the knowledge of $n-1$ first integrals of an n-dimensional vector field $\boldsymbol{A} : U \to \mathbb{R}^n$ determines the trajectories of the integral curves of the corresponding ODE $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})$. To obtain the flow/general solution one needs only to determine the correct parameterization of these integral curves. This will be explored carefully in the next section. Here we limit our discussion to some examples.

Example 7.9. We have already established that $I_1 = x^2 + y^2$ is a first integral of the vector field $-y\partial/\partial x + x\partial/\partial y$. Since this is 2-dimensional vector field, a single first integral suffices to determine the integral curves. In this case, for values $I_1 > 0$, the integral curves are circles centered at the origin. The actual solutions of the ODE, namely

$$x = x_0 \cos t - y_0 \sin t, \quad y = x_0 \sin t + y_0 \cos t$$

describe a parameterization of these circles.

Example 7.10. Consider the following 3-dimensional ODE:

$$\begin{aligned} \dot{x} &= 0 \\ \dot{y} &= x \\ \dot{z} &= y. \end{aligned}$$

The following functions are first integrals:

$$I_1 = x \tag{54}$$

$$I_2 = 2xz - y^2. \tag{55}$$

It's quite evident that $I_1 = x$ is a first integral. As for I_2, observe that

$$\dot{I}_2 = 2z\dot{x} + 2x\dot{z} - 2y\dot{y} = 2xy - 2xy = 0.$$

Therefore, the integral curves are parabolas described by equations (54) where we restrict I_1, I_2 to specific values.

A solution of the ODE is given by a certain parameterization of these parabolas. We can solve the system in question in a step-by-step fashion, taking anti-derivatives at each step. We obtain the following general solution

$$\begin{aligned} x &= x_0, \\ y &= x_0\, t + y_0, \\ z &= x_0\, t^2/2 + y_0\, t + z_0. \end{aligned}$$

This is a parameterization of the curve (54). Having the general solution, an explicit calculation verifies that

$$\begin{aligned} x &= x_0 = I_1, \\ 2xz - y^2 &= 2x_0 z_0 - y_0^2 = I_2. \end{aligned}$$

8. Rectification

Definition 8.1. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field and $\boldsymbol{F} : U \to V$ a C^1 diffeomorphism. We say that $\boldsymbol{F}$ rectifies $\boldsymbol{A}$ if*

$$\boldsymbol{F}_*\boldsymbol{A} = \frac{\partial}{\partial \hat{x}^1},$$

where $\hat{x}^1, \ldots, \hat{x}^n$ are the standard coordinates on V. Equivalently, setting $y^j = F^j(x^1, \ldots, x^n)$ we can say that $\boldsymbol{F}$ rectifies $\boldsymbol{A}$ if

$$\boldsymbol{A} = \frac{\partial}{\partial y^1}.$$

We can easily describe the flow generated by a unit vector. Therefore, if we can rectify a vector field, we can solve the corresponding ODE.

Example 8.2. Earlier, we saw that

$$\boldsymbol{A} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} = \frac{\partial}{\partial \theta},$$

where r, θ are the usual polar coordinates. Therefore, the transformation

$$\theta = \tan^{-1}(y/x), \quad r = \sqrt{x^2 + y^2}$$

rectifies the given vector field $\boldsymbol{A}$. Hence, the flow generated by $\boldsymbol{A}$ can be given by

$$\theta = \theta_0 + t, \quad r = r_0.$$

Rewriting the above using x, y coordinates yields the usual expression for the flow, namely

$$\begin{aligned}
x &= r\cos\theta = r_0\cos(\theta_0 + t) = r_0\cos(\theta_0)\cos(t) - r_0\sin(\theta_0)\sin(t) \\
&= x_0\cos(t) - y_0\sin(t), \\
y &= r\sin\theta = r_0\sin(\theta_0 + t) \\
&= x_0\sin(t) + y_0\cos(t).
\end{aligned}$$

Also, note that r is a first integral.

Proposition 8.3 (Rectification gives flows)**.** *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field and $\boldsymbol{F} : U \to V$ a diffeomorphism that rectifies $\boldsymbol{A}$. Then, the flow generated by $\boldsymbol{A}$ is given by*

$$\Phi(t, x^1, \ldots, x^n) = \boldsymbol{G}(y^1 + t, y^2, \ldots, y^n),$$

where $y^j = F^j(x^1, \ldots, x^n)$, and where $\boldsymbol{G} = \boldsymbol{F}^{-1}$ is the inverse diffeomorphism.

Proof. Let
$$\Psi(t, \hat{x}^1, \ldots, \hat{x}^n) = (\hat{x}^1 + t, \hat{x}^2, \ldots, \hat{x}^n)$$
denote the flow generated by $\partial/\partial\hat{x}^1$. Our definition of Ψ is equivalent to $\Phi = \boldsymbol{G}_*\Psi$. Since
$$\boldsymbol{A} = \boldsymbol{G}_* \frac{\partial}{\partial \hat{x}^1},$$
the conclusion follows by Proposition 6.3. □

Proposition 8.4 (Rectification gives first integrals.)**.** *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field and $\boldsymbol{F} : U \to V$ a diffeomorphism that rectifies $\boldsymbol{A}$. Then*
$$y^j = F^j(x^1, \ldots, x^n), \quad j = 2, \ldots, n$$
are first integrals of $\boldsymbol{A}$.

Proof. By definition,
$$\boldsymbol{A} = \frac{\partial}{\partial y^1}.$$
Hence, by the covariance of the directional derivative, we have
$$\begin{aligned} \boldsymbol{A}[y^1] &= 1, \\ \boldsymbol{A}[y^j] &= 0, \quad j = 2, \ldots, n. \end{aligned}$$
□

Next, suppose that $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field such that $A^1(0, x^2, \ldots, x^n) \neq 0$ for all $\boldsymbol{x} \in U_0$ where
$$U_0 = \{(0, x^2, \ldots, x^n) \in U\}$$
is the intersection of U and the $x^1 = 0$ hyperplane. Let
$$\boldsymbol{\Phi}(t, \boldsymbol{x}) = (\Phi^1(t, x^1, \ldots, x^n), \ldots, \Phi^n(t, x^1, \ldots, x^n))$$
be the flow generated by $\boldsymbol{A}$. Consider the equation
$$\Phi^1(\hat{t}, x^1, \ldots, x^n) = 0, \tag{56}$$
where $\hat{t}$ is the unknown. Observe that $\Phi^1(0, 0, x^2, \ldots, x^n) = 0$ and that
$$\dot{\Phi}^1(0, 0, x^2, \ldots, x^n) = A^1(0, x^2, \ldots, x^n) \neq 0.$$
Hence, by the implicit function theorem, there exists a C^1 function $\hat{t} = \tau(x^1, \ldots, x^n)$ defined in some neighborhood $\hat{U} \subset U$ of the hyperplane slice U_0. With $\hat{t}$ as above, set
$$I^j = \Phi^j(\hat{t}, x^1, \ldots, x^n).$$

Proposition 8.5 (Flows give first integrals)**.** *With the above definitions, the functions $I^2, \ldots, I^n$ are first integrals of $\boldsymbol{A}$.*

Proof. By the 1-parameter group property, we have

$$\Phi^1(\tau(x) - t, \boldsymbol{\Phi}(t, \boldsymbol{x})) = \Phi^1(\tau(\boldsymbol{x}), \boldsymbol{x}).$$

Since locally, $\hat{t} = \tau(\boldsymbol{x})$ is the unique solution of the equation system we must have

$$\tau(\boldsymbol{\Phi}(t, \boldsymbol{x})) = \tau(\boldsymbol{x}) - t.$$

Let us define functions $Q^j(x^1, \ldots, x^n),\ j = 2, \ldots, n$ by writing

$$I^j = Q^j(x^1, \ldots, x^n) = \Phi^j(\tau(\boldsymbol{x}), \boldsymbol{x}).$$

We have

$$\begin{aligned} Q^j(\Phi(t, \boldsymbol{x})) &= \Phi^j(\tau(\Phi(t, \boldsymbol{x})), \Phi(t, \boldsymbol{x})) \\ &= \Phi^j(\tau(\boldsymbol{x}) - t, \Phi(t, \boldsymbol{x})) \\ &= Q^j(\boldsymbol{x}). \end{aligned}$$

This establishes that $I^2, \ldots, I^n$ are first integrals. □

Next, let us discuss how to solve an autonomous ODE using first integrals. As we already indicated, knowledge of $n-1$ functionally independent first integrals serves as a description of the unparameterized integral curves. We can obtained the desired parameterization by employing a special coordinate system built using the first integrals, and by means of a reparameterization akin to the one shown in the proof to Theorem 2.1.

The next two Proposition indicate how to obtain a rectification by means of first integrals.

Proposition 8.6. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field such that $A^1(x^1, \ldots, x^n) \neq 0$ and let*

$$\xi^j = F^j(x^1, \ldots, x^n), \quad j = 2, \ldots, n \tag{57}$$

be $n-1$ functionally independent, C^1, first integrals. Then, setting

$$\xi^1 = F^1(x^1, \ldots, x^n) := x^1, \tag{58}$$

we obtain a diffeomorphism $\boldsymbol{F} : U_1 \to V$, where $U_1 \subset U$ is a suitably restricted subdomain $U_1 \subset U$. Furthermore,

$$\boldsymbol{A} = B^1(\xi^1, \ldots, \xi^n) \frac{\partial}{\partial \xi^1}, \tag{59}$$

where the function B^1 satisfies and is defined by the relation

$$B^1(\xi^1, \ldots, \xi^n) = A^1(x^1, \ldots, x^n). \tag{60}$$

Proof. Our first claim is that $\xi^1, \xi^2, \dots, \xi^n$ are functionally independent. Suppose not. By assumption, $\xi^2, \dots, \xi^n$ are functionally independent, and so $x^1 = \xi^1 = F(\xi^2, \dots, \xi^n)$ for some C^1 function F of $n-1$ variables. However, since $\xi^2, \dots, \boldsymbol{x}^n$ are first integrals, this would imply that x^1 is also a first integral. This would necessitate

$$\boldsymbol{A}[\xi^1] = A^1(x^1, \dots, x^n) = 0,$$

a contradiction.

Therefore $\xi^1, \dots, x^n$ are functionally independent. Hence, by taking

$$F^1(x^1, \dots, x^n) = x^1$$

and F^j as in (57), and by restricting the domain to some $U_1 \subset U$ we obtain a diffeomorphism $\boldsymbol{F} : U_1 \to V$.

Finally, let us prove (59). Let $\boldsymbol{B} = \boldsymbol{F}_* A$. By definition,

$$A = \sum_{j=1}^{n} B^j(\xi^1, \dots, \xi^n) \frac{\partial}{\partial \xi^j}.$$

By assumption,

$$\begin{aligned} \boldsymbol{A}[\xi^1] &= A^1(x^1, \dots, x^n), \\ \boldsymbol{A}[\xi^j] &= 0, \quad j = 2, \dots, n. \end{aligned}$$

Therefore, $B^j(\xi^1, \dots, \xi^n) = 0$ for $j = 2, \dots, n$ and

$$B^1(\xi^1, \dots, \xi^n) = A^1(x^1, \dots, x^n),$$

as was to be shown. □

Proposition 8.7. *Let $A^1 : U \to \mathbb{R}$ be a non-zero, C^1 function and let $\boldsymbol{A} = A^1(x^1, \dots, x^n) \frac{\partial}{\partial x^1}$ be the indicated vector field. Define*

$$\xi^1 = F^1(x^1, x^2, \dots, x^n) := \int_0^{x^1} \frac{du}{A^1(u, x^2, \dots, x^n)}, \tag{61}$$

$$\xi^j = F^j(x^1, \dots, x^n) := x^j, \quad j = 2, \dots, n. \tag{62}$$

Then,

$$\boldsymbol{A} = \frac{\partial}{\partial \xi^1}; \tag{63}$$

i.e., $\boldsymbol{F}$ rectifies $\boldsymbol{A}$.

Proof. Observe that

$$\begin{aligned} \boldsymbol{A}[\xi^1] &= (\mathrm{D}_1 F^1)(x^1, \dots, x^n) A^1(x^1, \dots, x^n) \\ &= \frac{A^1(x^1, \dots, x^n)}{A^1(x^1, \dots, x^n)} = 1, \\ \boldsymbol{A}[\xi^j] &= 0, \quad j = 2, \dots, n. \end{aligned}$$

□

Example 8.8. Consider the generator of the 2-dimensional rotation group:

$$\boldsymbol{A} = -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}.$$

We have already established that $r = \sqrt{x^2 + y^2}$ is a first integral. Let us therefore take

$$\xi = x, \quad \eta = \sqrt{x^2 + y^2}$$

as a new system of coordinates. As our domain, we can take $U_1 = \{(a_1, a_2) \in \mathbb{R}^2 : a_2 > 0\}$. Letting $(\xi, \eta) = \boldsymbol{F}(x, y)$ denote the corresponding transformation, we have

$$V = \boldsymbol{F}(U_1) = \{(a_1, a_2) \in \mathbb{R}^2 : a_2 > |a_1|\}.$$

The inverse diffeomorphism is given by

$$x = \xi, \quad y = \sqrt{\eta^2 - \xi^2}, \quad \eta > |\xi|.$$

Therefore,

$$\begin{aligned}
\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) &= \left(\frac{\partial}{\partial \xi}, \frac{\partial}{\partial \eta}\right) \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{pmatrix} \\
&= \left(\frac{\partial}{\partial \xi}, \frac{\partial}{\partial \eta}\right) \begin{pmatrix} 1 & 0 \\ \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \end{pmatrix}, \\
-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} &= -\sqrt{\eta^2 - \xi^2}\frac{\partial}{\partial \xi}
\end{aligned}$$

The next step is to reparameterize the simple translational flow for $\partial/\partial\xi$ so as to obtain the flow for $-\sqrt{\eta^2 - \xi^2}\partial/\partial\xi$. According to Proposition 8.7 we can do this by making the change of coordinates

$$\begin{aligned}
\theta &= \int_0^\xi \frac{du}{-\sqrt{\eta^2 - \xi^2}} = \tan^{-1}\frac{\sqrt{\eta^2 - (\xi)^2}}{\xi} - \pi/2 = \tan^{-1}(y/x) - \pi/2, \\
r &= \eta = \sqrt{x^2 + y^2}.
\end{aligned}$$

Example 8.9. Let's return to the differential equation

$$\dot{x} = 0, \quad \dot{y} = x, \quad \dot{z} = y.$$

Earlier we demonstrated that

$$\xi^2 = x, \quad \xi^3 = 2xz - y^2$$

are first integrals. Let us use this information to rectify the corresponding vector field. We begin by introducing coordinates $\xi^1 = y$ with ξ^2, ξ^3 as above, and restricting the domain to $y > 0$. We calculate

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \left(\frac{\partial}{\partial \xi^2}, \frac{\partial}{\partial \xi^1}, \frac{\partial}{\partial \xi^3}\right) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2y & 2z & 2x \end{pmatrix},$$

$$x\frac{\partial}{\partial y} + y\frac{\partial}{\partial z} = \xi^2 \frac{\partial}{\partial \xi^1}.$$

Next, we follow Proposition 8.7 and introduce coordinates

$$\begin{aligned} \eta^1 &= \int_0^{\xi^1} \frac{du}{\xi^2} = \frac{\xi^1}{\xi^2}, \\ &= y/x, \\ \eta^2 &= \xi^2 = x, \\ \eta^3 &= \xi^3 = 2xz - y^2. \end{aligned}$$

Relative to these coordinates, we have

$$x\frac{\partial}{\partial y} + y\frac{\partial}{\partial z} = \xi^2 \frac{\partial}{\partial \xi^1} = \frac{\partial}{\partial \eta^1}.$$

We can use this rectification to integrate the given ODE. The inverse transformation is given by

$$x = \eta^2, \quad y = \eta^1 \eta^2, \quad z = \frac{1}{2}\left(\eta^3/\eta^2 + (\eta^1)^2 \eta^2\right).$$

Therefore, the flow is given by

$$\begin{aligned} &\eta^1 = \eta^1_0 + t, \quad \eta^2 = \eta^2_0, \quad \eta^3 = \eta^3_0. \\ x &= \eta^2 = x_0, \\ y &= \eta^1\eta^2 = (y_0/x_0 + t)x_0 = x_0 t + y_0, \\ z &= \frac{1}{2}\left((2x_0 z_0 - y_0^2)/x_0 + (y_0/x_0 + t)^2 x_0\right) \\ &= \frac{1}{2} x_0 t^2 + 2t y_0 + z_0. \end{aligned}$$

9. Symmetries

Definition 9.1. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a C^1 vector field, $U_1, V_1 \subset U$ subdomains, and $\boldsymbol{F} : U_1 \to V_1$ a diffeomorphism. We say that $\boldsymbol{F}$ is a symmetry of $\boldsymbol{A}$ if $\boldsymbol{F}_*\boldsymbol{A} = \boldsymbol{A}$. Equivalently, setting $y^j = F^j(x^1, \ldots, x^n)$ we say that $\boldsymbol{F}$ is a symmetry if*

$$\boldsymbol{A} = \sum_{i=1}^{n} A^i(x^1, \ldots, x^n)\frac{\partial}{\partial x^i} = \sum_{i=1}^{n} A^i(y^1, \ldots, y^n)\frac{\partial}{\partial y^i}. \tag{64}$$

We say that $\boldsymbol{F}$ is a conformal *symmetry if $\boldsymbol{F}_*\boldsymbol{A} = f\boldsymbol{A}$ where $f : V_1 \to \mathbb{R}$ is a non-vanishing function.*

Example 9.2. Consider the rotation generator $-y\partial/\partial x + x\partial/\partial y$ and the following transformations.

(1) The transformation $\hat{x} = -y, \hat{y} = x$ is a symmetry because

$$-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} = -\hat{y}\frac{\partial}{\partial \hat{x}} + \hat{x}\frac{\partial}{\partial \hat{y}}. \tag{65}$$

(2) The transformation

$$x = \sqrt{u^2 + v^2}\cos u, \quad y = \sqrt{u^2 + v^2}\sin u$$

is a conformal symmetry because

$$-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y} = \frac{\partial}{\partial u} - \frac{u}{v}\frac{\partial}{\partial v} = \left(-\frac{1}{v}\right)\left(-v\frac{\partial}{\partial u} + u\frac{\partial}{\partial v}\right).$$

(3) Let k be a real number. The homothety transformations

$$\hat{x} = e^k x, \quad \hat{y} = e^k y$$

form a 1-parameter group of symmetries, because for every k relation (65) holds.

Definition 9.3. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field and $\Phi(t, \boldsymbol{x})$ the flow generated by $\boldsymbol{A}$. We say that a diffeomorphism $\boldsymbol{F} : U_1 \to V_1$ preserves the solutions of the corresponding ODE $\dot{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{x})$ if Φ if $\boldsymbol{F}_*\Phi = \Phi$. Similarly, we say $\boldsymbol{F}$ preserves the integral curves of $\boldsymbol{A}$ if there exists a reparameterization function $\tau(s, \boldsymbol{x})$ such that*

$$(\boldsymbol{F}_*\Phi)(s, \boldsymbol{x}) = \Phi(\tau(s, \boldsymbol{x}), \boldsymbol{x}). \tag{66}$$

Note: recall that a flow reparameterization function $\tau(s, \boldsymbol{x})$ satisfies

$$\tau(0, \boldsymbol{x}) = 0$$
$$\dot{\tau}(0, \boldsymbol{x}) \neq 0.$$

Proposition 9.4. *A diffeomorphism $\boldsymbol{F}$ is a symmetry of a vector field $\boldsymbol{A}$ if and only if $\boldsymbol{F}$ preserves the solutions of $\boldsymbol{A}$. Similarly, $\boldsymbol{F}$ is a conformal symmetry if and only it preserves the integral curves.*

Proof. Earlier we demonstrated the covariance of flows and vector fields. Thus, $\boldsymbol{F}_*\Phi$ is the flow for $\boldsymbol{F}_*\boldsymbol{A}$. Therefore $\boldsymbol{F}_*\boldsymbol{A} = \boldsymbol{A}$ if and only if $\boldsymbol{F}_*\Phi = \Phi$.

Next, let us suppose that $\boldsymbol{F}$ is a conformal symmetry of $\boldsymbol{A}$, with $\boldsymbol{F}_*\boldsymbol{A} = f\boldsymbol{A}$. Let $\Phi(t, \boldsymbol{x})$ be the flow generated by $\boldsymbol{A}$ and $\Psi(t, \boldsymbol{x})$ the flow generated by $f\boldsymbol{A}$. By the reparameterization theorem, there exists a reparameterization function $\tau(s, \boldsymbol{x})$ such that

$$\Psi(s, \boldsymbol{x}) = \Phi(\tau(s, \boldsymbol{x}), \boldsymbol{x}). \tag{67}$$

Conversely, suppose that (66) holds. It follows that $\Psi(s, \boldsymbol{x})$ as defined by (67) is the flow generated by $\boldsymbol{F}_*\boldsymbol{A}$. Observe that

$$\begin{aligned}\dot{\Psi}(0, \boldsymbol{x}) &= \dot{\Phi}(\tau(0, \boldsymbol{x}), \boldsymbol{x})\dot{\tau}(0, \boldsymbol{x}),\\ &= f(\boldsymbol{x})\dot{\Phi}(0, \boldsymbol{x}),\end{aligned}$$

where

$$f(\boldsymbol{x}) = \dot{\tau}(0, \boldsymbol{x}). \tag{68}$$

□

Example 9.5. Let us return to the symmetries of the rotation generator. The rotation flow is given by

$$x = x_0 \cos t - y_0 \sin t, \quad y = x_0 \sin t + y_0 \cos t.$$

We now consider the effects of the 3 transformations given above on this flow.

(1) We have

$$\begin{aligned}\hat{x} &= -y = -x_0 \sin t - y_0 \cos t\\ &= \hat{x}_0 \cos t - \hat{y}_0 \sin t,\\ \hat{y} &= x = \hat{x}_0 \sin t + \hat{y}_0 \cos t.\end{aligned}$$

The form of the flow is preserved by the transformation.

(2) Recall that

$$x^2 + y^2 = x_0^2 + y_0^2.$$

is a first integral, and hence the integral curves have the form $x^2 + y^2 = \text{const.}$. Next, consider the same function, but in the new coordinates:

$$u^2 + v^2 = x^2 + y^2$$

Therefore, integral curves are preserved by the transformation.

As for the reparameterization function, the flow in the u, v coordinates is given by

$$u = u_0 + t, \quad v = \sqrt{u_0^2 + v_0^2 - u^2} = \sqrt{v_0^2 - 2u_0 t - t^2}.$$

If we employ the reparameterization function

$$t = u_0 \cos s - v_0 \sin s - u_0,$$

we obtain

$$\begin{aligned} u &= u_0 \cos s - v_0 \sin s, \\ v &= \sqrt{u_0^2 + v_0^2 - u_0^2 \cos^2 s - v_0^2 \sin^2 s + 2u_0 v_0 \sin s \cos s} \\ &= \sqrt{u_0^2 \sin^2 s + v_0^2 \cos^2 s + 2u_0 v_0 \sin s \cos s} \\ &= u_0 \sin s + v_0 \cos s \end{aligned}$$

provided we restrict our flows and vector fields to the domain $v > 0$.

10. The Lie bracket

Definition 10.1. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be C^2 vector fields. We define the Lie bracket of $\boldsymbol{A}$ and $\boldsymbol{B}$ to be the vector field*

$$[\boldsymbol{A}, \boldsymbol{B}] = \mathcal{J}\,\boldsymbol{B} \cdot \boldsymbol{A} - \mathcal{J}\,\boldsymbol{A} \cdot \boldsymbol{B} \tag{69}$$

$$= \sum_{i=1}^{n} \big(\boldsymbol{A}[B^i] - \boldsymbol{B}[A^i]\big)\, \partial_{x^i} \tag{70}$$

$$= \sum_{i,j=1}^{n} \big(A^j\, \mathrm{D}_j B^i - B^j\, \mathrm{D}_i A^i\big)\, \partial_{x^i}, \tag{71}$$

where for convenience we are using

$$\partial_{x^i} = \frac{\partial}{\partial x^i}, \quad \partial_r = \frac{\partial}{\partial r}, \quad \textit{etc.}$$

Proposition 10.2. *Let $f : U \to R^n$ be a C^2 function. Then*

$$[\boldsymbol{A}, \boldsymbol{B}][f] = \boldsymbol{A}[\boldsymbol{B}[f]] - \boldsymbol{B}[\boldsymbol{A}[f]] \tag{72}$$

Indeed, $[\boldsymbol{A}, \boldsymbol{B}]$ is the unique vector field that satisfies the above property.

Proof. The first assertion follows from a straight-forward calculations. Note that all the 2nd order derivative parts cancel. As for the second assertion, observe that $\boldsymbol{A}[x^i] = A^i$, and similarly for $\boldsymbol{B}$. Hence, by (72) we have

$$[\boldsymbol{A}, \boldsymbol{B}][x^i] = \boldsymbol{A}[B^i] - \boldsymbol{B}[A^i],$$

in full agreement with (69). $\square$

Proposition 10.3. *The Lie bracket has the following properties. Throughout, $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are vector fields and $f : U \to \mathbb{R}$ is a function.*

$$[\boldsymbol{A}, \boldsymbol{B}] = -[\boldsymbol{B}, \boldsymbol{A}], \tag{73}$$

$$[\boldsymbol{A} + \boldsymbol{B}, \boldsymbol{C}] = [\boldsymbol{A}, \boldsymbol{C}] + [\boldsymbol{B}, \boldsymbol{C}], \tag{74}$$

$$[\boldsymbol{A}, [\boldsymbol{B}, \boldsymbol{C}]] + [\boldsymbol{B}, [\boldsymbol{C}, \boldsymbol{A}]] + [\boldsymbol{C}, [\boldsymbol{A}, \boldsymbol{B}]] = 0, \tag{75}$$

$$[\boldsymbol{A}, f\boldsymbol{B}] = \boldsymbol{A}[f]\boldsymbol{B} + f[\boldsymbol{A}, \boldsymbol{B}] \tag{76}$$

As a consequence of the above properties, we have

$$[\boldsymbol{A}, \boldsymbol{A}] = 0,$$
$$[f\boldsymbol{A}, \boldsymbol{B}] = -\boldsymbol{B}[f]\boldsymbol{A} + f[\boldsymbol{A}, \boldsymbol{B}].$$

Also, if c is a constant, then

$$[c\boldsymbol{A}, \boldsymbol{B}] = [\boldsymbol{A}, c\boldsymbol{B}] = c[\boldsymbol{A}, \boldsymbol{B}].$$

Proposition 10.4. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and $y^j = F^j(x^1, \dots, x^n)$ the corresponding change of coordinates. For $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$, vector fields, we have*

$$[\boldsymbol{F}_*\boldsymbol{A}, \boldsymbol{F}_*\boldsymbol{B}] = \boldsymbol{F}_*[\boldsymbol{A}, \boldsymbol{B}]. \tag{77}$$

Proof. This follow directly from Proposition 10.2 and the covariance of the directional derivative. □

Example 10.5. Observe that, by definition of the Lie bracket,

$$[\partial_{x^i}, \partial_{x^j}] = 0.$$

We can use this fact together with properties (74) (75) to calculate the Lie bracket. For example,

$$\begin{aligned}
[x\partial_x + y\partial_y, -y\partial_x + x\partial_y] &= [x\partial_x, -y\partial_x] + [x\partial_x, x\partial_y] \\
&\quad + [y\partial_y, -y\partial_x] + [y\partial_y, x\partial_y] \\
&= 0.
\end{aligned}$$

Switching to polar coordinates, we have

$$\begin{aligned}
x\partial_x + y\partial_y &= r\partial_r, \\
-y\partial_x + x\partial_y &= \partial_\theta.
\end{aligned}$$

Another straightforward calculation confirms the principle of covariance:

$$[r\partial_r, \partial_\theta] = 0.$$

11. Infinitesimal symmetries

Definition 11.1. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be vector fields. We say that $\boldsymbol{A}$ is an infinitesimal symmetry of $\boldsymbol{B}$ if $[\boldsymbol{A}, \boldsymbol{B}] = 0$. We say that $\boldsymbol{A}$ is an infinitesimal conformal symmetry of $\boldsymbol{B}$ if $[\boldsymbol{A}, \boldsymbol{B}] = f\boldsymbol{B}$, where $f : U \to \mathbb{R}^n$ is a function.*

Note: if $\boldsymbol{A}$ is a symmetry of $\boldsymbol{B}$, then necessarily $\boldsymbol{B}$ is a symmetry of $\boldsymbol{A}$. However, the analogous statement does not hold for conformal symmetries.

Proposition 11.2. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be vector fields. If $[\boldsymbol{A}, \boldsymbol{B}] = 0$, then for all functions $g : U \to \mathbb{R}^n$, the vector field $\boldsymbol{A}$ is a conformal symmetry of $g\boldsymbol{B}$. Conversely, if $\boldsymbol{A}$ is an infinitesimal conformal symmetry of $\boldsymbol{B}$, then there exists a subdomain $U_1 \subset U$ and a non-zero function $g : U_1 \to \mathbb{R}$ such that $[\boldsymbol{A}, g\boldsymbol{B}] = 0$.*

Proof. The forward implication follows from property (76) of the Lie bracket. Let us prove the converse. Rectify $\boldsymbol{A}$ so that $\boldsymbol{A} = \partial_{y^1}$ relative to some coordinates $y^j = F^j(x^1, \ldots, x^n)$, where $\boldsymbol{F} : U \supset U_1 \to V$ is a diffeomorphism. By assumption $[\boldsymbol{A}, \boldsymbol{B}] = f\boldsymbol{B}$. Let $h = \boldsymbol{G}^* f$, where $\boldsymbol{G} = \boldsymbol{F}^{-1}$. Equivalently,

$$h(y^1, \ldots, y^n) = f(x^1, \ldots, x^n).$$

Set

$$g(x^1, \ldots, x^n) = \exp\left(-\int^{y_1} h(u, y^2, \ldots, y^n)du\right).$$

Hence, by construction,

$$\begin{aligned} \boldsymbol{A}[g] &= \frac{\partial g}{\partial y^1} = -fg, \\ [\boldsymbol{A}, g\boldsymbol{B}] &= g[\boldsymbol{A}, \boldsymbol{B}] + \boldsymbol{A}[g]\boldsymbol{B} \\ &= fg\boldsymbol{B} - fg\boldsymbol{B} = 0 \end{aligned}$$

□

Proposition 11.3. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be C^2 vector fields and Φ the C^2 flow generated by $\boldsymbol{A}$. Let*

$$\boldsymbol{C}(t, \boldsymbol{x}) = \boldsymbol{C}_t(\boldsymbol{x}) = ((\Phi_t)_*\boldsymbol{B})(\boldsymbol{x})$$

be the indicated time-dependent vector field. Then,

$$\dot{\boldsymbol{C}}_t = -[\boldsymbol{A}, \boldsymbol{C}_t]. \tag{78}$$

In particular,

$$\dot{\boldsymbol{C}}_0 = -[\boldsymbol{A}, \boldsymbol{B}].$$

Proof. By definition of the push-forward, we have

$$(\boldsymbol{C}_t \circ \Phi_t)(\boldsymbol{x}) = \boldsymbol{C}(t, \Phi(t, \boldsymbol{x})) = \mathcal{J}\,\Phi_t(\boldsymbol{x}) \cdot \boldsymbol{B}(\boldsymbol{x}).$$

Taking the derivative with respect to t and using the chain rule gives

$$\dot{\boldsymbol{C}}_t \circ \Phi_t + (\mathcal{J}\,\boldsymbol{C}_t \circ \Phi_t)\dot{\Phi}_t = (\mathcal{J}\,\Phi_t)^{\cdot} \cdot \boldsymbol{B}.$$

By definition, $\dot{\Phi}_t = \boldsymbol{A} \circ \Phi_t$. Hence, By Clairaut's theorem (interchange of partial derivatives with respect to time and position),

$$\begin{aligned}(\mathcal{J}\,\Phi_t)^{\cdot} &= \mathcal{J}\,\dot{\Phi}_t \\ &= \mathcal{J}\,(\boldsymbol{A} \circ \Phi_t) \\ &= (\mathcal{J}\,\boldsymbol{A} \circ \Phi_t) \cdot \mathcal{J}\,\Phi_t\end{aligned}$$

Hence,

$$\begin{aligned}(\mathcal{J}\,\boldsymbol{A} \circ \Phi_t) \cdot \mathcal{J}\,\Phi_t \cdot \boldsymbol{B} &= \dot{\boldsymbol{C}}_t \circ \Phi_t + (\mathcal{J}\,\boldsymbol{C}_t \circ \Phi_t) \cdot (\boldsymbol{A} \circ \Phi_t) \\ &= (\dot{\boldsymbol{C}}_t + \mathcal{J}\,\boldsymbol{C}_t \cdot \boldsymbol{A}) \circ \Phi_t.\end{aligned}$$

However, since

$$\mathcal{J}\,\Phi_t \cdot \boldsymbol{B} = \boldsymbol{C}_t \circ \Phi_t,$$

we have

$$\mathcal{J}\,\boldsymbol{A} \cdot \boldsymbol{C}_t = \dot{\boldsymbol{C}}_t + \mathcal{J}\,\boldsymbol{C}_t \cdot \boldsymbol{A},$$

thereby establishing (78). □

Proposition 11.4. *Let $\boldsymbol{A} : U \to \mathbb{R}^n$ be a vector field and Φ_t the flow generated by A. Then Φ_t is a symmetry of $\boldsymbol{A}$ for all t, while $\boldsymbol{A}$ is its own infinitesimal symmetry.*

More generally, we have

Proposition 11.5. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be vector fields and Φ_t the flow generated by $\boldsymbol{A}$. Then $\boldsymbol{A}$ is an infinitesimal symmetry of $\boldsymbol{B}$ if and only if for all t, Φ_t is a symmetry of $\boldsymbol{B}$.*

Proof. Set $\boldsymbol{C}_t = (\Phi_t)_*\boldsymbol{B}$ and suppose that $\boldsymbol{C}_t = \boldsymbol{B}$ for all t. Then, since $\boldsymbol{C}_t$ is, by assumption, constant with respect to t, we have, by Proposition 11.3, $[\boldsymbol{A}, \boldsymbol{B}] = 0$.

Let us now prove the converse. Suppose that $[\boldsymbol{A}, \boldsymbol{B}] = 0$. Hence, by Proposition 10.4, and using the fact $(\Phi_t)_*\boldsymbol{A} = \boldsymbol{A}$ we have

$$(\Phi_t)_*[\boldsymbol{A}, \boldsymbol{B}] = [\boldsymbol{A}, \boldsymbol{C}_t] = 0. \tag{79}$$

By Proposition (11.3), we have

$$\dot{\boldsymbol{C}}_t = -[\boldsymbol{C}_t, \boldsymbol{A}] = 0,$$

Since $\boldsymbol{C}_0 = \boldsymbol{B}$ we conclude that $\boldsymbol{C}_t = \boldsymbol{B}$ for all t. □

Proposition 11.6. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be vector fields and Φ_s, Ψ_t the respective flows. Then $[\boldsymbol{A}, \boldsymbol{B}] = 0$ if and only if*

$$\Phi_s \circ \Psi_t = \Psi_t \circ \Phi_s$$

for all s, t for which the above transformations are defined.

Proof. Above, we showed that $[\boldsymbol{A}, \boldsymbol{B}] = 0$ if and only if $\Phi_{t*}\boldsymbol{B} = \boldsymbol{B}$ for all t. By Proposition 9.4, the latter is true if and only if

$$\Phi_{t*}\Psi_s = \Phi_t \circ \Psi_s \circ \Phi_{-t} == \Psi_s$$

for all s, as was to be shown. □

Example 11.7. Let $\boldsymbol{A} = x\partial_x + y\partial_y$ and $\boldsymbol{B} = -y\partial_x + x\partial_y$. Let

$$\begin{aligned} \Phi_t(x, y) &= (e^t x, e^t y) \\ \Psi_t(x, y) &= (\cos(t)x - \sin(t)y, \sin(t)x + \cos(t)y) \end{aligned}$$

be the corresponding flows. Consider the transformation

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

We have

$$\begin{aligned} (\partial_x, \partial_y) &= (\partial_u, \partial_v) \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix}, \\ x\partial_x + y\partial_y &= (\partial_u, \partial_v) \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} e^{-t}u \\ e^{-t}v \end{pmatrix} \\ &= u\partial_u + v\partial_v, \\ -y\partial_x + x\partial_y &= (\partial_u, \partial_v) \begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix} \begin{pmatrix} -e^{-t}v \\ e^{-t}u \end{pmatrix} \\ &= -v\partial_u + u\partial_v \end{aligned}$$

Similarly, consider the transformation

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

Equivalently,

$$\begin{pmatrix} u \\ v \end{pmatrix} = R_t \begin{pmatrix} x \\ y \end{pmatrix},$$

where

$$R_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

is the indicated rotation matrix. We have

$$\begin{aligned}
(\partial_x, \partial_y) &= (\partial_u, \partial_v) R_t, \\
x\partial_x + y\partial_y &= (\partial_u, \partial_v) R_t \begin{pmatrix} x \\ y \end{pmatrix}, \\
&= (\partial_u, \partial_v) R_t R_t^{-1} \begin{pmatrix} u \\ v \end{pmatrix}, \\
&= u\partial_u + v\partial_v.
\end{aligned}$$

As well,

$$\begin{aligned}
-y\partial_x + x\partial_y &= (\partial_u, \partial_v) R_t \begin{pmatrix} -y \\ x \end{pmatrix} \\
&= (\partial_u, \partial_v) R_t R_{\frac{\pi}{2}} \begin{pmatrix} x \\ y \end{pmatrix}, \\
&= (\partial_u, \partial_v) R_t R_{\frac{\pi}{2}} R_{-t} \begin{pmatrix} u \\ v \end{pmatrix}, \\
&= (\partial_u, \partial_v) R_{\frac{\pi}{2}} \begin{pmatrix} u \\ v \end{pmatrix}, \\
&= -v\partial_u + u\partial_v
\end{aligned}$$

Proposition 11.8. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and $\boldsymbol{G} : V \to U$ the inverse transformation. For $\boldsymbol{A} : U \to \mathbb{R}^n$, a vector field, and $f : U \to \mathbb{R}$, a function, we have*

$$\boldsymbol{F}_*(g\boldsymbol{B}) = \boldsymbol{G}^*(g)\boldsymbol{F}_*(\boldsymbol{B}). \tag{80}$$

Proof. Homework. □

Proposition 11.9. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be vector fields and Φ_t the flow generated by $\boldsymbol{A}$. Then $\boldsymbol{A}$ is an infinitesimal conformal symmetry of $\boldsymbol{B}$ if and only if for all t, Φ_t is a conformal symmetry of $\boldsymbol{B}$.*

Proof. Suppose that $\boldsymbol{A}$ is a conformal symmetry of $\boldsymbol{B}$. Then, there exists a non-zero function $g : U \to \mathbb{R}$ such that $[\boldsymbol{A}, g\boldsymbol{B}] = 0$. Hence,

$$\Phi_{t*}(g\boldsymbol{B}) = g\boldsymbol{B}.$$

By the definition of push-forward and the above Proposition,

$$\Phi_{t*}(g\boldsymbol{B}) = \Phi_{-t}^*(g)\Phi_{t*}(\boldsymbol{B}).$$

Hence,

$$\Phi_{t*}(\boldsymbol{B}) = h\boldsymbol{B},$$

where

$$h = g/\Phi_{-t}^*(g).$$

Conversely, suppose that Φ_t is a conformal symmetry for all t. Hence, there exists a function $h_t : U \to \mathbb{R}$ such that

$$\Phi_{t*}\boldsymbol{B} = h_t\boldsymbol{B}.$$

Taking d/dt, setting $t = 0$, and using Proposition 11.5 gives

$$[\boldsymbol{A}, \boldsymbol{B}] = \dot{h}_0\boldsymbol{B},$$

as was to be shown. □

Example 11.10. The radial vector field $\boldsymbol{A} = x\partial_x + y\partial_y$ is an infinitesimal conformal symmetry of the vector field $\boldsymbol{B} = \partial_x - (x/y)\partial_y$. Let Φ_t be the flow generated by $\boldsymbol{A}$. Fix t and consider the transformation

$$(u, v) = \Phi_t(x, y) = (e^t x, e^t y).$$

We have

$$\begin{aligned} \partial_x - (x/y)\partial_y &= (\partial_u, \partial_v)\begin{pmatrix} e^t & 0 \\ 0 & e^t \end{pmatrix}\begin{pmatrix} 1 \\ -u/v \end{pmatrix} \\ &= e^t(\partial_u - (u/v)\partial_v) \end{aligned}$$

Therefore, Φ_t is a conformal symmetry of $\boldsymbol{B}$.

Proposition 11.11. *Let $\boldsymbol{A}, \boldsymbol{B} : U \to \mathbb{R}^n$ be vector fields and Φ_t the flow generated by $\boldsymbol{A}$. If $\boldsymbol{A}$ is an infinitesimal conformal symmetry of $\boldsymbol{B}$, then Φ_t preserves the integral curves of $\boldsymbol{B}$.*

Proof. See Proposition 9.4. □

Example 11.12. Let us return to the vector fields of the preceding examples. The integral curves of $\boldsymbol{B}$ are the semicircles about the origin, $x^2 + y^2 = R^2$ in the upper half-plane, $y > 0$. Changing the variables to u, v the equation of the circle becomes $u^2 + v^2 = (e^t R)^2$. Therefore the Φ_t transforms a circle of radius R to a circle of radius $e^t R$.

12. Symmetry Integration

Definition 12.1. *Let $\boldsymbol{A} : U \to \mathbb{R}^2$ where $U \subset \mathbb{R}^2$ be a C^1 vector field and*

$$dy/dx = \omega(x, y) \tag{81}$$

a non-autonomous ODE, where $\omega : U \to \mathbb{R}$ is a C^1 function. We call $\boldsymbol{A}$ an infinitesimal symmetry of such an ODE if $\boldsymbol{A}$ is an infinitesimal conformal symmetry of $\partial_x + \omega(x, y)\partial_y$.

Proposition 12.2. *A vector field $\boldsymbol{A} = \xi\,\partial_x + \eta\,\partial_y$ is an infinitesimal symmetry of* (81) *if and only if*

$$\omega^2\xi_y + \omega(\xi_x - \eta_y) - \eta_x + \xi\omega_x + \eta\omega_y = 0 \tag{82}$$

Proof. We have

$$[\xi\partial_x + \eta\partial_y, \partial_x + \omega\partial_y] = (-\omega\xi_y - \xi_x)\partial_x + (-\omega\eta_y - \eta_x + \xi\omega_x + \eta\omega_y)\partial_y$$

The RHS is proportional to $\partial_x - \omega\partial_y$ if and only if (82) holds. $\square$

Equation (82) is called the symmetry determining equation.

Proposition 12.3. *Let $u = F^1(x, y), v = F^2(x, y)$ be a change of coordinates. An ODE* (81) *in standard coordinates is equivalent to an ODE*

$$\frac{dv}{du} = \hat{\omega}(u, v),$$

where

$$\hat{\omega}(u, v) = \frac{v_x + \omega(x, y)v_y}{u_x + \omega(x, y)u_y}, \tag{83}$$

where $u_x = \partial_x(u) = \partial u/\partial x$, etc.

Proof. By the usual change of coordinate formula for vector fields, we have

$$\begin{aligned} \partial_x &= u_x\partial_u + v_x\partial_v \\ \partial_y &= u_y\partial_u + v_y\partial_v \\ \partial_x + \omega\partial_y &= (u_x + \omega u_y)\partial_u + (v_x + \omega v_y)\partial_v \\ &= (u_x + \omega u_y)(\partial_u + \hat{\omega}(u, v)\partial_v). \end{aligned}$$

$\square$

Definition 12.4. *Let $U_1, V_1 \subset U$ be subdomains. We call a diffeomorphism $\boldsymbol{F} : U_1 \to V_1$ a symmetry of the ODE* (81) *if*

$$\hat{\omega}(u, v) = \omega(u, v)$$

where $(u, v) = \boldsymbol{F}(x, y)$ and where $\hat{\omega}$ is defined as per (83).

Proposition 12.5. *A vector field $\boldsymbol{A} : U \to \mathbb{R}^2$ is an infinitesimal symmetry of the ODE* (81) *if and only if the corresponding flow Φ_t is a symmetry for all t.*

Proposition 12.6. *The unit vector field ∂_y is a symmetry of* (81) *if and only if $\omega(x, y) = f(x)$ is independent of y. In this case, the general solution of* (81) *is given by*

$$y = \int^x f(s)\, ds$$

In light of the above proposition, one method of integration is to rectify the symmetry and then find the general solution by means of a quadrature.

Proposition 12.7. *The vector field $g(y)\partial_y$ is a symmetry of an ODE* (81) *if and only if there exists a function $f(x)$ such that*

$$\omega(x, y) = f(x)g(y).$$

In this case, the general solution is given by

$$\int^y \frac{ds}{g(s)} = \int^x f(s) ds,$$

i.e, by separation of variables.

Proof. Let us introduce coordinates

$$u = x, \quad v = \int^y \frac{ds}{g(s)},$$

and thereby rectify

$$g(y)\partial_y = \partial_v.$$

By the preceding Proposition, in these new coordinates the ODE has the form

$$\frac{dv}{du} = f(u),$$

for some function $f(u) = f(x)$. Since

$$\partial_u + f(u)\partial_v = \partial_x + f(x)g(y)\partial_y$$

the form of the ODE in standard coordinates is

$$\frac{dy}{dx} = f(x)g(y),$$

a separable equation. Again, by the preceding proposition, the general solution has the form

$$v = \int^u f(s),$$

as was to be shown. $\square$

Proposition 12.8. *The vector field $\boldsymbol{A} = x\partial_x + y\partial_y$ is a symmetry of an ODE* (81) *if and only if*

$$\omega(x, y) = f(y/x),$$

for some function f of one variable. In this case, the general solution is given by

$$x = \exp \int^{y/x} \frac{ds}{f(s) - s} \tag{84}$$

Proof. We rectify $\boldsymbol{A}$ by introducing coordinates $u = y/x$, $v = \log(x)$. Now

$$\begin{aligned} x\partial_x &= \partial_v - u\partial_u, \\ y\partial_y &= u\partial_u, \\ x\partial_x + y\partial_y &= \partial_v. \end{aligned}$$

By (83) the ode $dv/du = f_1(u)$ is equivalent to (81) where

$$f_1(u) = \frac{1/x}{-y/x^2 + \omega(x, y)/x},$$

or equivalently,

$$\omega(x, y) = f(y/x), \quad \text{where} \quad f(u) = \frac{1}{f_1(u)} - u.$$

Therefore, the general solution is given by

$$v = \int^u f_1(s)\, ds,$$

or equivalently, by

$$\log(x) = \int^{y/x} \frac{ds}{f(s) - s}.$$

$\square$

13. Differential 1-forms.

Definition 13.1. *Let V be an n-dimensional vector space. A linear form is a linear mapping $V \to R$. The set of all such forms the n-dimensional vector space V^*, called the dual of V. Let $v_1, \dots, v_n$ be a basis of V. The unique basis $\epsilon^1, \dots, \epsilon^n \in V^*$ such that*

$$\epsilon^i(v_j) = \delta^i_j, \quad i, j = 1, \dots, n$$

is called the dual basis of $v_1, \dots, v_n$.

Note: it's customary to represent elements of $\mathbb{R}^n$ as column vectors and elements of $(\mathbb{R}^n)^*$ as row vectors.

Definition 13.2. *Let $U \subset \mathbb{R}^n$ be a domain. A differential 1-form, or simply a differential, is a mapping $\boldsymbol{\alpha} : U \to (\mathbb{R}^n)^*$. For a vector field $\boldsymbol{A} : U \to \mathbb{R}^n$, we define $\boldsymbol{\alpha} \cdot \boldsymbol{A} : U \to R$ to be the function*

$$(\boldsymbol{\alpha} \cdot \boldsymbol{A})(\boldsymbol{x}) = \boldsymbol{\alpha}(\boldsymbol{x})(\boldsymbol{A}(\boldsymbol{x})).$$

Expressing a differential as a row vector of functions

$$\boldsymbol{\alpha} = \big(\alpha_1, \dots, \alpha_n\big)$$

and a vector field as a column vector of functions

$$\boldsymbol{A} = \begin{pmatrix} A^1 \\ \vdots \\ A^n \end{pmatrix},$$

we can express the product of a differential 1-form and a vector field as

$$\boldsymbol{\alpha} \cdot \boldsymbol{A} = \sum_{i=1}^{n} \alpha_i A^i.$$

We say that $\boldsymbol{\alpha}$ is of class C^1, C^2, analytic, etc if its components $\alpha_1, \dots, \alpha_n$ are C^1, C^2, analytic, etc.

Definition 13.3. *Let $f : U \to \mathbb{R}$ be a differentiable function. We call the 1-form*

$$df := \left(\frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n} \right)$$

the differential of f. We call the constant differentials $dx^1, \dots, dx^n$ the basic 1-forms and write

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}\, dx^i.$$

Proposition 13.4. *Let $\boldsymbol{F} : U \to V$ be a diffeomorphism and $y^j = F^j(x^1, \ldots, x^n)$ the corresponding change of coordinates. Then,*

$$dy^j = \sum_{i=1}^{n} \frac{\partial y^j}{\partial x^i} dx^i \tag{85}$$

are the unique differential 1-forms having the property that

$$dy^j \cdot \frac{\partial}{\partial y_i} = \delta_i^j. \tag{86}$$

Furthermore, every differential 1-form $\boldsymbol{\alpha} : U \to (\mathbb{R}^n)^*$ has the unique expression

$$\boldsymbol{\alpha} = \sum_{j=1}^{n} \tilde{\alpha}^j(y^1, \ldots, y^n) dy^j, \tag{87}$$

where $\tilde{\alpha}^j : V \to \mathbb{R}$ are functions defined by

$$\begin{aligned} \tilde{\alpha}^j(y^1, \ldots, y^n) &= \left(\boldsymbol{\alpha} \cdot \frac{\partial}{\partial y^j}\right)(x^1, \ldots, x^n) \\ &= \sum_{i=1}^{n} \frac{\partial x^i}{\partial y^j} \alpha_i(x^1, \ldots, x^n) \end{aligned}$$

Proposition 13.5. *Let $f : U \to \mathbb{R}$ be a C^1 function and $y^j = F^j(x^1, \ldots, x^n)$ a coordinates system. Then,*

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial y^j}\, dy^j. \tag{88}$$

Indeed, df is the unique differential 1-form characterized by the condition that for every C^0 vector field $\boldsymbol{A} : U \to \mathbb{R}^n$ we have

$$df \cdot \boldsymbol{A} = \boldsymbol{A}[f]. \tag{89}$$

A differential can also be regarded as the integrand of a line integral.

Proposition 13.6. *Let $\boldsymbol{\alpha} : U \to \mathbb{R}$ be a continuous differential and $\boldsymbol{\gamma} : I \to U$ a C^1 curve, where $I = [t_0, t_1]$. The line integral*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{\alpha} = \int_I (\boldsymbol{\alpha} \circ \boldsymbol{\gamma})(\dot{\boldsymbol{\gamma}}) = \int_{t_0}^{t_1} \sum_{i=1}^{n} \alpha^i(\boldsymbol{\gamma}(t)) \gamma^{i\prime}(t)\, dt$$

is invariant under orientation preserving reparameterization.

Definition 13.7. *If $\boldsymbol{\alpha}$ is a 1-form such that $\boldsymbol{\alpha} = df$ for some function $f : U \to \mathbb{R}$, then we call $\boldsymbol{\alpha}$ an* exact differential.

Proposition 13.8. *Let $\boldsymbol{\alpha} = df$ be an exact differential on a domain U. Let $\boldsymbol{\gamma} : I \to U$ be a parameterized curve. Then,*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{\alpha} = f(\boldsymbol{x}_1) - f(\boldsymbol{x}_0),$$

where $x_i = \boldsymbol{\gamma}(t_i)$, $i = 0, 1$. Conversely, suppose that U is a connected domain, and that for all curves $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$ with endpoints $\boldsymbol{x}_0, \boldsymbol{x}_1 \in U$ we have

$$\int_{\gamma_1} \boldsymbol{\alpha} = \int_{\gamma_2} \boldsymbol{\alpha}.$$

Then, $\boldsymbol{\alpha} = df$ is an exact differential with

$$f(\boldsymbol{x}) = \int_{\boldsymbol{x}_0}^{\boldsymbol{x}} \boldsymbol{\alpha}, \quad \boldsymbol{x} \in U$$

where the above integral is taken along an arbitrary curve that connects $\boldsymbol{x}_0$ to $\boldsymbol{x}$.

Definition 13.9. *Say that $\boldsymbol{\alpha} : U \to (\mathbb{R}^n)^*$ is a closed differential if*

$$\frac{\partial \alpha^i(x^1, \ldots, x^n)}{\partial x^j} - \frac{\partial \alpha^j(x^1, \ldots, x^n)}{\partial x^i} = 0, \tag{90}$$

for all $1 \leq i < j \leq n$.

Proposition 13.10. *An exact differential is closed. Conversely, if U is a simply connected domain, then every closed differential is exact.*

14. Symmetries and integrating factors.

In this section, we resrict our attention to 2-dimensional, simply connected domains $U \subset \mathbb{R}^n$.

Proposition 14.1. *Let $\boldsymbol{\alpha} : U \to (\mathbb{R}^2)^*$ be a C^1 differential. There after a suitable domain restriction $U_1 \subset U$ there exists a C^1 function $\mu : U_1 \to \mathbb{R}$ such that $\mu\boldsymbol{\alpha}$ is exact.*

Proof. Write $\boldsymbol{\alpha} = M\,dx + N\,dy$. Consider the vector field $\boldsymbol{A} = -N\,\partial_x + M\,\partial_y$. By rectifying this vector field, we can find a non-constant first integral $\xi = f(x, y)$ such that

$$d\xi \cdot \boldsymbol{A} = \boldsymbol{A}[\xi] = 0.$$

Hence,

$$\begin{vmatrix} M & N \\ \xi_x & \xi_y \end{vmatrix} = M\xi_y - N\xi_x = 0.$$

We now define

$$\mu = \xi_x / M$$

on a suitably restricted domain on which ξ is defined and $M \neq 0$. In this way,

$$\mu\boldsymbol{\alpha} = \xi_x\,dx + \xi_x(N/M)\,dy = \xi_x\,dx + \xi_y\,dy = d\xi.$$

If M is identically zero, we simply define $\mu = 1/N$. $\square$

Definition 14.2. *Let $\boldsymbol{\alpha} = Mdx + Ndy$ be a C^1 differential defined on a domain $U \subset \mathbb{R}^2$. We call $\boldsymbol{\gamma} : I \to U$, where $I \subset \mathbb{R}$ is an interval, an integral curve of $\boldsymbol{\alpha}$ if the function*

$$f(t) = (\boldsymbol{\alpha} \circ \boldsymbol{\gamma})(t) \cdot \dot{\boldsymbol{\gamma}}(t)$$

is identically zero for all $t \in I$.

Proposition 14.3. *Let $\boldsymbol{\alpha}, \boldsymbol{\gamma}$ be as above. Then $\boldsymbol{\gamma}$ is an integral curve of $\boldsymbol{\alpha}$ if and only if*

$$\int_{\boldsymbol{\gamma}|I_1} \boldsymbol{\alpha} = 0$$

for all subintervals $I_1 \subset I$.

Proposition 14.4. *Let $\boldsymbol{\alpha}, \boldsymbol{\gamma}$ be as above. Then $\boldsymbol{\gamma}$ is an integral curve of $\boldsymbol{\alpha}$ if and only if $\boldsymbol{\gamma}$ describes a solution of the ODE $dy/dy = -M/N$.*

Definition 14.5. *We call a function $\mu : U \to \mathbb{R}$ such that $\mu\boldsymbol{\alpha}$ is closed an integrating factor of a differential $\boldsymbol{\alpha}$.*

To solve an ODE $dy/dx = -M/N$ it suffices to find an integrating factor for the differential $\boldsymbol{\alpha} = Mdx + Ndy$ and restrict the domain.

Proposition 14.6. *Let $A = \xi\partial_x + \eta\partial_y$ be a vector field and $\boldsymbol{\alpha} = M\,dx + N\,dy$ a differential. If $\boldsymbol{\alpha}\cdot\boldsymbol{A} = 0$, then $\boldsymbol{A}$ is an infinitesimal symmetry of $dy/dx = -M/N$.*

Proof. The vector fields $\boldsymbol{A}$ and $-N\partial_x + M\partial_y$ are proportional if and only if $\boldsymbol{A}\cdot\boldsymbol{\alpha} = 0$. Observe that if f is any function, then

$$[\boldsymbol{A}, f\boldsymbol{A}] = \boldsymbol{A}[f]\boldsymbol{A}.$$

Hence, $\boldsymbol{A}$ is a conformal symmetry of $f\boldsymbol{A}$ for all functions f. □

If $\boldsymbol{\alpha}\cdot\boldsymbol{A} = 0$, we call $\boldsymbol{A}$ the *trivial symmetry* of $dy/dx = -M/N$.

Proposition 14.7. *Let $U \subset \mathbb{R}^2$ be a simply connected domain, $\boldsymbol{A} : U \to \mathbb{R}^n$ a vector field and $\boldsymbol{\alpha} = M\,dx + N\,dy$ a differential such that $\boldsymbol{\alpha}\cdot\boldsymbol{A} \neq 0$. Then $\boldsymbol{A}$ is a symmetry of the ODE $dy/dx = -M/N$ if and only if*

$$\mu = \frac{1}{\boldsymbol{\alpha}\cdot\boldsymbol{A},} \tag{91}$$

is an integrating factor of $\boldsymbol{\alpha}$.

Proof. Write $\boldsymbol{A} = \xi(x,y)\partial_x + \eta(x,y)\partial_y$. Let μ be as above, and consider the differential

$$\begin{aligned}\mu\boldsymbol{\alpha} &= \frac{M}{M\xi + N\eta}dx + \frac{N}{M\xi + N\eta}dy,\\ &= 1/(\eta - \omega(x,y)\xi)\,(dy - \omega(x,y)dx)\end{aligned}$$

where

$$\omega(x,y) = -M/N.$$

By (90) and by Proposition 13.10, the above differential is closed if and only if

$$\partial_x(\eta - \xi\omega)^{-1} + \partial_y(\omega/(\eta - \xi\omega)) = 0.$$

Taking derivatives and clearing denominators, the above condition is equivalent to

$$-\partial_x(\eta - \xi\omega) + (\eta - \xi\omega)\omega_y - \omega\partial_y(\eta - \xi\omega) = 0$$

The above, in turn, is equivalent to the determining equation (82). □

Example 14.8. Let's revisit the integration method for a homogeneous differential equations

$$\frac{dy}{dx} = f(y/x).$$

Earlier, we showed that $x\partial_x + y\partial_y$ is an infinitesimal symmetry. Hence,

$$\mu = ((dy - f(y/x)dx)\cdot(x\partial_x + y\partial_y))^{-1} = (y - xf(y/x))^{-1}$$

is an integrating factor. Hence the solution of the homogeneous ODE can be obtained by integrating

$$\int \left(\frac{-f(y/x)}{y - xf(y/x)} dx + \frac{1}{y - xf(y/x)} dy \right).$$

The above differential is closed (check!), hence locally exact. To obtain solution, let us integrate the dy term (i.e., we treate x as a constant. We obtain

$$\begin{aligned} \int \frac{dy}{y - xf(y/x)} &= \int \frac{1}{y/x - f(y/x)} \frac{dy}{x} \\ &= F(y/x) = \int^{y/x} \frac{du}{u - f(u)} \end{aligned}$$

Observe that

$$\begin{aligned} \partial_x F(y/x) = F'(y/x)(-y/x^2) &= \frac{1}{y/x - f(y/x)}(-y/x^2) \\ &= \left(-\frac{1}{x}\right) \frac{y}{y - xf(y/x)} = -\frac{1}{x} - \frac{f(y/x)}{y - xf(y/x)} \end{aligned}$$

Therefore,

$$\int \left(\frac{-f(y/x)}{y - xf(y/x)} dx + \frac{1}{y - xf(y/x)} dy \right) = \\ -\log(x) + F(y/x) + C.$$

Thus, the same general solution can be obtained by integration. The necessary changes of variables arise naturally in the course of this integration.

Proposition 14.9. *The vector field $f(x)\partial_y$ is a symmetry of an ODE* (81) *if and only if only if*

$$\omega(x, y) = p(x)y + q(x),$$

where $p(x) = f'(x)/f(x)$.

Proof. From the determining equation (82), $f(x)\partial_y$ is a symmetry if and only if

$$\omega_y = f'(x)/f(x).$$

This is true if and only if ω has the form shown above. □

Proposition 14.10. *The general solution of a linear ODE*

$$\frac{dy}{dx} = p(x)y + q(x)$$

is given by

$$y = f(x) \int^x \frac{q(s)}{f(s)} ds,$$

where $f(x) = \exp(\int^x p(s)\, ds)$.

Proof. With the above definition, we have $f'(x)/f(x) = p(x)$. Hence, $f(x)\partial_y$ is a symmetry, and hence $1/f(x)$ is an integrating factor. We have

$$\frac{dy}{f(x)} - \frac{f'(x)y}{f(x)^2}dx - \frac{q(x)}{f(x)}dx = d\left(\frac{y}{f(x)}\right) - d\int^x \frac{q(s)}{f(s)}ds.$$

Therefore,

$$\int \frac{1}{f(x)} (dy - (p(x)y + q(x))dx) =$$
$$y/f(x) - \int^x (q(s)/f(s))ds,$$

where the constant of integration is incorportated in the RHS integral. $\square$

15. Reduction of order

Consider a general second-order ordinary differential equations

$$\frac{d^2y}{dx^2} = \omega\left(x, y, \frac{dy}{dx}\right). \tag{92}$$

We convert the above ODE to a 3D autonomous system by introducing the variable $y_1 = dy/dx$ and writing

$$\begin{aligned} \dot{x} &= 1 \\ \dot{y} &= y_1, \\ \dot{y_1} &= \omega(x, y, y_1). \end{aligned}$$

In other words, solving a 2nd order ODE is equivalent to rectifying the vector field

$$\boldsymbol{A} = \partial_x + y_1\partial_y + \omega(x, y, y_1)\partial_{y_1} \tag{93}$$

on a 3-dimensional domain with coordinates (x, y, y_1). We will show that a conformal infinitesimal symmetry of $\boldsymbol{A}$ allows us to reduce the order of (92). This means that we can formulate a first order ODE

$$\frac{dv}{du} = \phi(u, v) \tag{94}$$

such that the solutions of (92) are obtained from the solutions of (94) by a quadrature (the old-fashioned name for an anti-derivative).

Example 15.1. Let us consider a 1-dimensional conservative mechanical system

$$\frac{d^2x}{dt^2} = a(x), \tag{95}$$

where $a(x)$ is a position-dependent acceleration function. We introduce a velocity variable v and seek a function $v = f(x)$ that gives velocity as a function of position. Since $v = dx/dt$, the function f must satisfy

$$\frac{dv}{dt} = f'(x(t))f(x(t)) = a(x(t)).$$

In other words, $v = f(x)$ is a solution of the following 1st order ODE

$$\frac{dv}{dx} = \frac{a(x)}{v}. \tag{96}$$

If we know a solution $v = f(x)$ of (96), we can obtain a solution $x = p(t)$ of (95) by integrating the ODE

$$\frac{dx}{dt} = f(x).$$

This can be done using a quadrature, namely

$$t = \int^x \frac{du}{f(u)}. \tag{97}$$

The general solution of (95) should depend on 2 constants of integration: the initial position and the initial velocity. One of these constants arises in (97), the other arises when (and if) we solve (96).

Now for 1-dimensional mechanical systems it is always possible to solve the reduced equations by means of a quadrature (separate variables). Indeed the general solution of (96) is given by

$$\frac{1}{2}v^2 = E - U(x), \tag{98}$$

where

$$U(x) = -\int_{x_0}^x a(s)ds \tag{99}$$

is called a potential function and x_0 is some fixed position. The E is a constant of integration known as the energy of the mechanical system. Note that if we change x_0, we change E. Thus, energy is relative in exactly the same way that the potential function is relative. The relationship between time and position can now be given implictly as

$$t = \pm \int^x \frac{ds}{\sqrt{2E - 2U(s)}}. \tag{100}$$

We use the + branch on the portions of the trajectory where the velocity is positive and − on the portions where it is negative.

From the point of view of geometry, we are trying to describe the integral curves of the 3-dimensional vector field

$$\partial_t + v\partial_x + a(x)\partial_v. \tag{101}$$

on a 3-dimensional domain (called phase space) with coordinates (t, x, v). These integral curves are describe by means of 2 first-integrals:

$$E = \frac{1}{2}v^2 + U(x),$$
$$\tau = t - \operatorname{sgn}(v) \int^x \frac{ds}{\sqrt{2E - U(s)}},$$

Definition 15.2. *We define an autonomous 2nd order ODE to be an ODE of the form*

$$\frac{d^2y}{dx^2} = \omega(y, dy/dx). \tag{102}$$

The solutions of such an ODE correspond the integral curves of the vector field

$$\partial_x + y_1\partial_y + \omega(y, y_1)\partial_{y_1} = \partial_x + y_1 \left(\partial_y + \frac{\omega(y, y_1)}{y_1}\partial_{y_1} \right).$$

A 1st integral can be obtained by integrating the 1st order ODE

$$\frac{dy_1}{dy} = \frac{\omega(y, y_1)}{y_1} \tag{103}$$

Indeed, if $y_1 = f(y)$ is a solution of (103), we can obtain a general solution of (102) by using a quadrature to solve $dy/dx = f(y)$; i.e. by taking

$$x = \int^y \frac{ds}{f(s)}.$$

The case of a conservative mechanical system presented above is an example of an 2nd order autonomous ODE. Here is another example.

Example 15.3. Consider the 2nd order ODE

$$\frac{d^2y}{dx^2} = \frac{dy}{dx}.$$

The above is an autonomous equation, so we can reduce order and consider the equation

$$\frac{dy_1}{dy} = \frac{y_1}{y_1} = 1.$$

The general solution is $y_1 = y - C_1$. We now integrate

$$\frac{dy}{dx} = y - C_1$$

using a quadrature to obtain

$$x = \log(y - C_1) - \tilde{C},$$

or equivalently, after setting $C_2 = e^{\tilde{C}}$,

$$y = C_2 e^x + C_1.$$

Thus, a strategy for integrating a 2nd order ODE is to find and rectify an infinitesimal symmetry. Doing so will, in effect, transform the given 2nd order ODE into an autonomous one. In order to put this strategy into practice we have to inquire about the transformation law for 2nd order ODEs.

Definition 15.4. *Let $(u, v) = \boldsymbol{F}(x, y)$ be a diffeomorphic transformation. The first prolongation of this diffeomorphism is the transformation $(u, v, v_1) = \boldsymbol{F}^{(1)}(x, y, y_1)$ where, c.f.,* (83)

$$v_1 = \frac{v_x + y_1 v_y}{u_x + y_1 u_y} \tag{104}$$

The second prolognation is the transformation $(u, v, v_1, v_2) = \boldsymbol{F}^{(2)}(x, y, y_1, y_2)$ where

$$v_2 = \frac{u_x v_y - u_y v_x}{(u_x + y_1 u_y)^3}\, y_2 + \frac{(\partial_x + y_1 \partial_y)[v_1]}{u_x + y_1 u_y}. \tag{105}$$

Earlier, we proved the following.

Proposition 15.5. *Let $(u, v) = \boldsymbol{F}(x, y)$ be a diffeomorphic transformation. Then the solutions of a 1st order ODE*

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

are equivalent to the solutions of

$$\frac{dv}{dy} = \tilde{\omega}(u, v)$$

where where $v_1 = \tilde{\omega}(u, v)$ is related to $y_1 = \omega(x, y)$ by the 1st prolongation formula (104).

Generalizing to 2nd order, we have the following.

Proposition 15.6. *Let $(u, v) = \boldsymbol{F}(x, y)$ be a diffeomorphic transformation. Then the solutions of a 2nd order ODE* (92) *are equivalent to the solutions of*

$$\frac{d^2 v}{du^2} = \tilde{\omega}\left(u, v, \frac{dv}{du}\right),$$

where $v_2 = \tilde{\omega}(u, v, v_1)$ is related to $y_2 = \omega(x, y, y_1)$ by the 2nd prolongation formula (105).

Proof. Observe that

$$\partial_x + y_1 \partial_y = (u_x + y_1 u_y)(\partial_u + v_1 \partial_v),$$

where v_1, y_1 are related as above. Similarly,

$$\partial_x + y_1 \partial_y + y_2 \partial_{y_1} = (u_x + y_1 u_y)(\partial_u + v_1 \partial_v + v_2 \partial_{y_2})$$

where y_2, v_2 are related by the above formula. $\square$

Definition 15.7. *Let $\boldsymbol{A} = \xi(x,y)\partial_x + \eta(x,y)\partial_y$ be a vector field defined on a 2-dimensional domain. Setting*

$$\upsilon(x,y,y_1) = (\partial_x + y_1\partial_y)[\eta(x,y) - y_1\xi(x,y)] \tag{106}$$

$$= \eta_x + y_1(\eta_y - \xi_x) - y_1^2\xi_y, \tag{107}$$

we call the 3-dimensional vector field

$$\boldsymbol{A}^{(1)} = \xi(x,y)\partial_x + \eta(x,y)\partial_y + \upsilon(x,y,y_1)\partial_{y_1} \tag{108}$$

the first prolongation *of $\boldsymbol{A}$.*

Proposition 15.8. *Let Φ_t be the flow generated by a 2-dimensional vector field $\boldsymbol{A} = \xi(x,y)\partial_x + \eta(x,y)\partial_y$. Then $\Phi_t^{(1)}$ is the flow generated by the 3-dimensional vector field $\boldsymbol{A}^{(1)}$.*

Proof. Homework. □

Definition 15.9. *Let us say that a vector field $\boldsymbol{A} = \xi\partial_x + \eta\partial_y$ is an infintesimal symmetry of a 2nd order ODE* (92) *if the prolongation $\boldsymbol{A}^{(1)}$ is an infinitesimal conformal symmetry of the vector field $\partial_x + y_1\partial_y + \omega(x,y,y_1)\partial_{y_1}$.*

Proposition 15.10. *A 2nd order ODE* (92) *admits the translation generator ∂_x as an infinitesimal symmetry if and only if it is an autonomous ODE of the form* (102)*.*

Proof. Note that the prolongation of ∂_x is equal to $\partial_x + 0\partial_y + 0\partial_{y_1}$. Calculating the Lie bracket, we obtain

$$[\partial_x, \partial_x + y_1\partial_y + \omega(x,y,y_1)\partial_{y_1}] = \partial_x(\omega(x,y,y_1))\partial_{y_1}$$

Thus in order for ∂_x to be a conformal symmetry, ω_x must vanish. □

Next, let us consider 2nd order equations that admit $x\partial_x + y\partial_y$ as a symmetry. Since we know how to rectify this vector field, we will obtain a class of equations that admit a reduction of order. We will call these types of equations, *homogeneous 2nd order ODEs.*

Proposition 15.11. *A 2nd order ODE* (81) *admits $x\partial_x + y\partial_y$ as an infinitesimal symmetry if and only if it has the form*

$$\frac{d^2y}{dx^2} = \omega(y/x, dy/dx)/x \tag{109}$$

The general solution of such an equation is given by

$$x = \exp\left(\int^{y/x} \frac{du}{g(u)}\right), \tag{110}$$

where $v_1 = g(v)$ is a solution of the reduced ODE

$$\frac{dv_1}{dv} = \omega(v, v_1 + v)/v_1 - 1. \tag{111}$$

Proof. Let us rectify

$$x\partial_x + y\partial_y = \partial_u$$

by taking

$$u = \log(x), \quad v = y/x$$

as a change of coordinates. Applying the prolongation formulas (104) and (105) gives

$$\begin{aligned} v_1 &= \frac{-y/x^2 + y_1/x}{1/x} = -y/x + y_1, \\ v_2 &= \frac{1/x^2}{1/x^3} y_2 + (\partial_x + y_1\partial_y)[y_1 - y/x] \\ &= xy_2 + (y/x^2 - y_1/x)/(1/x) \\ &= xy_2 + y/x - y_1 \end{aligned}$$

Now in the u, v coordinates, an autonomous ODE (one that admits ∂_u as a symmetry) has the form

$$\frac{d^2v}{du^2} = \tilde{\omega}(v, v_1).$$

Setting $v_2 = \tilde{\omega}(v, v_1)$ and applying the above formula gives

$$y_2 = v_2/x$$

Writing

$$v_2 = \tilde{\omega}(v, v_1), \quad y_2 = (v_2 + v_1)/x = \omega(v, y_1)/x,$$

we obtain the transformation formula

$$\tilde{\omega}(v, v_1) = \omega(v, v_1 + v) - v_1.$$

The reduction method now follows from doing the reduction of an autonomous ODE in (u, v) coordinates, and then re-expressing the solution in (x, y) coordinates. $\square$

We expect the flow generated by an infinitesimal symmetry of a 2nd order equation to be a symmetry of that equation. We will not take the time to prove this in full generality, but will make note of the following. Recall that the vector field $x\partial_x + y\partial_y$ generates the flow

$$\Phi_t(x, y) = (e^t x, e^t y).$$

Proposition 15.12. *Let $k > 0$ be a constant. The transformation $(u, v) = (kx, ky)$ preserves a homogeneous 2nd order ODE* (109).

Proof. Apply the transformation formulas (104) (105). ☐

Example 15.13. Consider the following homogeneous 2nd order ODE

$$\frac{d^2y}{dx^2} = \frac{1}{y} + \frac{dy/dx}{x}.$$

The above is a homogeneous equation with

$$\omega(y/x, y_1) = (y/x)^{-1} + y_1.$$

The reduced equation is therefore,

$$\frac{dv_1}{dv} = (1/v + v_1 + v)/v_1 - 1 = v + 1/v.$$

The general solution of the reduced equation is

$$v_1 = \pm\sqrt{v^2 + 2\log(v) + C_1}.$$

Since $y_1 = v_1 + v$, the general solution of the given 2nd order equation is

$$x = \exp\left(\pm \int_{C_2}^{y/x} \frac{ds}{\sqrt{s^2 + 2\log(s) + C_1}}\right) \tag{112}$$

One could also say that this homogeneous system is equivalent to a mechanical system in u, v coordinates with potential $-v^2/2 - \log(v)$.

Let's derive the reduced equation directly. Set

$$\begin{aligned}
v &= y/x \\
v_1 &= x\frac{dv}{dx} = \frac{dy}{dx} - v \\
v_2 &= x\frac{dv_1}{dx} = x\frac{d^2y}{dx^2} - v_1 \\
&= x\frac{d^2y}{dx^2} - \frac{dy}{dx} + v
\end{aligned}$$

The given ODE can now be expressed as

$$\begin{aligned}
v_2 &= \frac{1}{v} + \frac{dy}{dx} - \frac{dy}{dx} + v \\
&= v + \frac{1}{v}
\end{aligned}$$

We seek a function $v_1 = g(v)$ that is compatible with the solution of the above equation. This implies that

$$v_2 = g'(v)g(v) = v + \frac{1}{v}$$

and hence that

$$g(v)^2 = v^2 + 2\log(v) + C$$

Having obtained $g(v)$ we can now solve for v as a function of x by noting that

$$\frac{dv}{dx} = \frac{g(v)}{x}$$

Hence,

$$\log(x) = \int^v \frac{ds}{g(s)};$$

another way to write the solution above.