# Note on Symmetry and Changes of Variables for Differential Equations

Eric A. Carlen
Rutgers University

February 17, 2011

## 1 Changes of variables and diffeomorphisms

The key step in the solution to many mathematical problems is to show that two problems are equivalent to one another in the sense that there is an explicit method for turning every solution of one problem into a solution of the other problem, and *vice-versa*.

In linear algebra, this is particularly simple: If two systems of linear equations are related to one another by a sequence of elementary row operations, every solution of one system simply *is* a solution of the other system, and *vice-versa*. Therefore, to solve a system of linear equations, you can apply repeated elementary row operations to simplify the system – by eliminating variables – and then solve the simplified system.

Our goal in these notes is to explain some important ideas concerning this strategy in the context of solving differential equations.

Changes of variables that are useful for studying differential equations are given by *diffeomorphisms*; that is, functions $h$ that are invertible and such that both $h$ and its inverse function $h^{-1}$ are continuously differentiable. We will begin by studying the case in which the domain of $h$ is an open set $U \subset \mathbb{R}^2$, and the range is an open set $V \subset \mathbb{R}^2$.

Recall that the derivative of a function

$$h(x,y) = (u(x,y), v(x,y))$$

from $U \subset \mathbb{R}^2$ to $V \subset \mathbb{R}^2$ is expressed by is *Jacobian matrix* $[J_h(\mathbf{x})]$ at $\mathbf{x} \in U$, which is the $2 \times 2$ matrix

$$[J_h(\mathbf{x})] = \begin{bmatrix} \nabla u(\mathbf{x}) \\ \nabla v(\mathbf{x}) \end{bmatrix} . \tag{1.1}$$

That is, the first row of $[J_h(\mathbf{x})]$ is $\nabla u(x)$ while the second row is $\nabla v(x)$. Written out explicitly as a $2 \times 2$ matrix,

$$[J_h(\mathbf{x})] = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{bmatrix} . \tag{1.2}$$

It is usually more useful to think of matrices as lists of vectors instead of arrays of numbers, and hence (1.1) will be useful to us, and is probably more memorable that (1.2).

When all of the partial derivatives in $[J_h(\mathbf{x})]$ exist *and are continuous* at a point $\mathbf{x}_0 \in U$, the function $h$ is *differentiable* there, meaning that

$$h(\mathbf{x}) = h(\mathbf{x}_0) + [J_h(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x}; \mathbf{x}_0) \tag{1.3}$$

where the *remainder term* $r(\mathbf{x}; \mathbf{x}_0)$ satisfies

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{r(\mathbf{x}; \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0 \ .$$

That is,

$$h(\mathbf{x}) \approx h(\mathbf{x}_0) + [J_h(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) \ ,$$

up to an error which is a vanishingly small fraction of the size $\|[J_h(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|$ of the main term on the right in (1.3). Differentiability of a function at a point means exactly that this linear approximation is a good approximation sufficiently close to the point.

Now, if $g$ is a differentiable map from $V$ to $W \subset \mathbb{R}^2$, we can form the composition $g \circ h$ which is a function from $U$ to $W$. If both functions are differentiable, so is their composition. Indeed, by the differentiability of $g$, and then the differentiability of $h$,

$$\begin{aligned} g(h(\mathbf{x})) &\approx g(h(\mathbf{x}_0)) + [J_g(h(\mathbf{x}_0))](h(\mathbf{x}) - h(\mathbf{x}_0)) \\ &\approx g(h(\mathbf{x}_0)) + [J_g(h(\mathbf{x}_0))]([J_h(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)) \\ &= g(h(\mathbf{x}_0)) + ([J_g(h(\mathbf{x}_0))][J_h(\mathbf{x}_0)])\,(\mathbf{x} - \mathbf{x}_0) \end{aligned}$$

Going back trough this and keeping track of the remainder terms, one sees that $g \circ h$ is differentiable at $\mathbf{x}_0$, and we obtain the *chain rule:*

$$[J_{g \circ h}(\mathbf{x}_0)] = [J_g(h(\mathbf{x}_0))][J_h(\mathbf{x}_0)] \ . \tag{1.4}$$

That is, the Jacobian of the the composition is the product of the Jacobians.

Note also that if $h$ is a linear transformation, so that for some $2 \times 2$ matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ ,$$

$$h(\mathbf{x}) = A\mathbf{x} = (ax + by, cx + dy) \ ,$$

we have

$$[J_h(\mathbf{x})] = A \ .$$

Thus, the Jacobian of a linear transformation is simply the constant matrix $A$ that represents that linear transformation. In particular, if $h$ is invertible, $h^{-1} \circ h$ is the identity transformation, which is the linear transformation represented by the identity matrix

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ .$$

Therefore, by the chain rule (1.4), supposing that both $h$ and $h^{-1}$ are differentiable,

$$I = [J_{h^{-1} \circ h}(\mathbf{x})] = [J_{h^{-1}}(h(\mathbf{x}_0))][J_h(\mathbf{x}_0)] \ .$$

It then follows that

$$[J_{h^{-1}}(h(\mathbf{x}_0))] = [J_h(\mathbf{x}_0)]^{-1} \ . \tag{1.5}$$

In fact, the *Inverse Function Theorem* says that if $h$ is differentiable at a point $\mathbf{x}_0$, and if $[J_h(\mathbf{x}_0)]$ is an invertible matrix, then $h$ is invertible on a neighborhood of $\mathbf{x}_0$, and this inverse is differentiable, and satisfies (1.5).

Finally, there is another case of the multivarible chain rule that will be useful to us here: If $\mathbf{x}(t)$ is a differentiable curve in $U$, and we define $\mathbf{u}(t) = h(\mathbf{x}(t))$ where $h$ is a differentiable function from $U$ to $V$, then $\mathbf{u}(t)$ is a differentiable curve in $V$, and

$$\mathbf{u}'(t) = [J_h(\mathbf{x}(t))]\mathbf{x}'(t) \ . \tag{1.6}$$

Now that we have reviewed these facts about differentiability in to variables, let us look at some examples of coordinate transformations.

**1 Example** (The polar coordinate transformation)**.** Let

$$U = \{(x,y)\ :\ x \neq 0 \quad \text{or } y > 0\ \}\ ,$$

which is $\mathbb{R}^2$ with the non-positive $y$-axis removed Let

$$V = \left\{ (r,\theta)\ :\ r > 0 \quad \text{and } -\frac{\pi}{4} < \theta < \frac{3\pi}{4}\ \right\}\ .$$

Define

$$h(x,y) = (r(x,y), \theta(x,y))$$

where

$$r(x,y) = \sqrt{x^2+y^2} \qquad \text{and} \qquad \theta(x,y) = \begin{cases} \arctan(y/x) & x > 0 \\ \pi/2 & x = 0 \\ \pi + \arctan(y/x) & x < 0\ . \end{cases}$$

The inverse transformation $h^{-1}$ is

$$h^{-1}(r,\theta) = (x(r,\theta), y(r,\theta))$$

where

$$x(r,\theta) = r\cos\theta \qquad \text{and} \qquad y(r,\theta) = r\sin\theta\ .$$

Not only are both $h$ and $h^{-1}$ differentiable, we can readily calculate the derivatives, which are expressed as the Jacobian matrices $[J_h]$ and $[J_{h^{-1}}]$ respectively.

We find

$$[J_h(x,y)] = \begin{bmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial\theta/\partial x & \partial\theta/\partial y \end{bmatrix} = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix}\ .$$

Likewise,

$$[J_{h^{-1}}(r,\theta)] = \begin{bmatrix} \partial x/\partial r & \partial x/\partial\theta \\ \partial y/\partial r & \partial y/\partial\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}\ .$$

Note that

$$[J_h(x,y)] \circ h^{-1}(r,\theta) = \begin{bmatrix} \frac{x(r,\theta)}{\sqrt{x^2(r,\theta)+y^2(r,\theta)}} & \frac{y(r,\theta)}{\sqrt{x^2(r,\theta)+y^2(r,\theta)}} \\ \frac{-y(r,\theta)}{x^2(r,\theta)+y^2(r,\theta)} & \frac{x(r,\theta)}{x^2(r,\theta)+y^2(r,\theta)} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -(\sin\theta)/r & (\cos\theta)/r \end{bmatrix} .$$

We cah check our work by verifying that

$$[J_{h^{-1}}(r,\theta)] = [J_h \circ h^{-1}(r,\theta)]^{-1} ,$$

which must be true according to (1.5).

**2 Example** (The homogeneous coordinate transformation)**.** Let

$$U = \{(x,y) \ : \ x > 0 \ \} ,$$

which is called *the right half-plane.* Let $V = \mathbb{R}^2$.

Define

$$h(x,y) = (u(x,y), v(x,y))$$

where

$$u(x,y) = \ln(x) \qquad \text{and} \qquad v(x,y) = \frac{y}{x} .$$

The inverse transformation $h^{-1}$ is

$$h^{-1}(r,\theta) = (x(u,v), y(u,v))$$

where

$$x(u,v) = e^u \qquad \text{and} \qquad y(u,v) = e^u v .$$

Again, it is easy to compute the derivatives; i.e., Jacobian matrices. We find

$$[J_h(x,y)] = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{bmatrix} = \begin{bmatrix} 1/x & 0 \\ -y/x^2 & 1/x \end{bmatrix} ,$$

and

$$[J_{h^{-1}}(u,v)] = \begin{bmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial u/\partial u & \partial y/\partial v \end{bmatrix} = \begin{bmatrix} e^u & 0 \\ ve^u & e^u \end{bmatrix} .$$

We see that

$$[J_h \circ h^{-1}(u,v)] = \begin{bmatrix} e^{-u} & 0 \\ -ve^{-u} & e^{-u} \end{bmatrix}$$

and hence

$$[J_{h^{-1}}(u,v)] = [J_h \circ h^{-1}(u,v)]^{-1} ,$$

as must be the case.

## 2 How vector fields and direction fields transform under changes of variables

Let

$$h(x,y) = (u(x,y), v(x,y))$$

be a diffeomorphism of an open set $U$ in $\mathbb{R}^2$ onto $V$, another open set in $\mathbb{R}^2$.

Let $\mathbf{x}(t)$, $a < t < b$, be a continuously differentiable curve in $U$. Then define

$$\mathbf{u}(t) = h(\mathbf{x}(t)) \ .$$

As we have noted above, this is a differentiable curve in $V$, and by the chain rule,

$$\mathbf{u}'(t) = [J_h(\mathbf{x}(t))]\mathbf{x}'(t) \ .$$

Let

$$\mathbf{v}(x,y) = (f(x,y), g(x,y))$$

be a smooth vector field on an open set $U$ in $\mathbb{R}^2$. Suppose that $\mathbf{x}(t)$ is an integral curve of $\mathbf{v}$, meaning that

$$\mathbf{x}'(t) = \mathbf{v}(\mathbf{x}(t)) \ .$$

Then

$$\begin{aligned} \mathbf{u}'(y) &= [J_h(\mathbf{x}(t))]\mathbf{v}(\mathbf{x}(t)) \\ &= [J_h(h^{-1}(\mathbf{u}(t)))]\mathbf{v}(h^{-1}(\mathbf{x}(t))) \\ &= \mathbf{w}(\mathbf{u}(t)) \ , \end{aligned} \tag{2.1}$$

where $\mathbf{w}(u,v)$ is the vector field on $V$ defined by

$$\mathbf{w}(u,v) = [J_h(h^{-1}(u,v)]\mathbf{v}(h^{-1}(u,v)) \ . \tag{2.2}$$

Thus, whenever $\mathbf{v}$ and $\mathbf{w}$ are related by (2.2), the diffeomophism $h$ carries integral curves of $\mathbf{v}$ onto integral curves of $\mathbf{w}$. Moreover, since $[J_{h^{-1}}(u,v)] = [J_h \circ h(u,v)]^{-1}$, whenever $\mathbf{v}$ and $\mathbf{w}$ are related by (2.2), we also have

$$\mathbf{v}(x.y) = [J_{h^{-1}}(h(x,y)]\mathbf{w}(h(x,y)) \ , \tag{2.3}$$

and *vice-versa*.

Thus (2.2) and (2.3) are equivalent to one another, and whenever either holds, $h$ carries integral curves of $\mathbf{v}$ onto integral curves of $\mathbf{w}$, and $h^{-1}$ carries integral curves of $\mathbf{w}$ onto integral curves of $\mathbf{v}$. This brings us to the following definition:

**2.1 DEFINITION** (Conjugate vector fields)**.** Let $U$ and $V$ be open sets in $\mathbb{R}^2$ with a diffeomorphism $h$ from $U$ onto $V$. Let $\mathbf{v}$ and $\mathbf{w}$ be continuously differentiable vector fields on $U$ and $V$ respectively. Then $\mathbf{w}$ is conjugate to $\mathbf{v}$ under $h$ in case (2.2) is true.

Since, as we have observed above (2.2) is equivalent to (2.3), $\mathbf{w}$ is conjugate to $\mathbf{v}$ under $h$ if and only if $\mathbf{v}$ is conjugate to $\mathbf{w}$ under $h^{-1}$ What we have noted above also proves the following theorem:

**2.2 THEOREM** (Conjugate vector fields)**.** *Let $U$ and $V$ be open sets in $\mathbb{R}^2$ with a diffeomorphism $h$ from $U$ onto $V$. Let $\mathbf{v}$ and $\mathbf{w}$ be continuously differentiable vector fields on $U$ and $V$ respectively. Suppose that $\mathbf{v}$ and $\mathbf{w}$ are conjugate to one another under $h$ and $h^{-1}$ respectively. Then a curve $\mathbf{x}(t)$ in $U$ solves*

$$\mathbf{x}'(t) = \mathbf{v}(\mathbf{x}(t)) \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.4}$$

*if and only if the curve $\mathbf{u}(t)$ in $V$ given by $\mathbf{u}(t) = h(\mathbf{x}(t))$ satisfies*

$$\mathbf{u}(t) = \mathbf{w}(\mathbf{u}(t)) \qquad \mathbf{u}(t_0) = \mathbf{u}_0 := h(\mathbf{x}_0)\ . \tag{2.5}$$

The point of the definition is that whenever $\mathbf{v}$ and $\mathbf{w}$ are conjugate, and we know how to solve (2.5), then we know how to solve (2.4): The solution is $h^{-1}(\mathbf{u}(t))$.

**3 Example** (Conjugacy under the polar diffeomorphism)**.** Let $U$ and $V$ and $h$ be given as in Example 1. Let

$$\mathbf{v}(x,y) = (y + x(1-x^2-y^2), -x + y(1-x^2-y^2))\ . \tag{2.6}$$

Define

$$\mathbf{w}(r,\theta) := h_*\mathbf{v}(r,\theta) = [J_h(h^{-1}(r,\theta))]\mathbf{v}(h^{-1}(r,\theta)\ ,$$

which makes $\mathbf{w}$ conjugate to $\mathbf{v}$ under $h$.

Computing we find, using results from Example 1,

$$\begin{aligned}\mathbf{w}(u,v) &= \begin{bmatrix} \cos\theta & \sin\theta \\ -(\sin\theta)/r & (\cos\theta)/r \end{bmatrix} (r\sin\theta + r(1-r^2)\cos\theta, -r\cos\theta + r(1-r^2)\sin\theta) \\ &= (r(1-r^2), -1)\ .\end{aligned}$$

Thus, to solve the equation $\mathbf{x}'(t) = \mathbf{v}(\mathbf{x}(t))$, it suffices to solve

$$(r', \theta') = (r(1-r^2), -1)\ . \tag{2.7}$$

This is much simpler, and we have solved (2.7) already in exercises. By making the inverse transformation, we recover the solution of the original differential equation.

We will frequently be working with transformations of vector fields under diffeomorphisms. The following definition will be useful.

**2.3 DEFINITION** (The push-forward of a vector field)**.** Let Let $U$ and $V$ be open sets in $\mathbb{R}^2$ with a diffeomorphism $h$ from $U$ onto $V$. Let $\mathbf{v}$ be a smooth vector filed on $U$, Then the *push-forward* of $\mathbf{v}$ under $h$ is the vector field on $V$ denoted by $h_*\mathbf{v}$ and defined by

$$h_*\mathbf{v} = \mathbf{w}(u,v) = [J_h(h^{-1}(u,v))]\mathbf{v}(h^{-1}(u,v))\ . \tag{2.8}$$

Now consider a diffeomorphism $h$ from $U$ to $V$, and then a diffeomorphism $g$ from $V$ to $W$ where $U$, $V$ and $W$ are all open sets in $\mathbb{R}^2$. Then, by the chain rule (1.4) $g \circ h$ is a diffeomorphism from $U$ onto $W$, and its derivative is given by $[J_{g\circ h}(\mathbf{x})] = [J_g(h(\mathbf{x}))][J_h(\mathbf{x})]$. Using this and the fact that $(g \circ h)^{-1} = h^{-1} \circ g^{-1}$, is is easy to compute $g \circ h_*$ in terms of $g_*$ and $h_*$: For any vector field $\mathbf{v}$ on $U$, and any $\mathbf{z} \in W$

$$\begin{aligned}
(g \circ h)_* \mathbf{v}(\mathbf{z}) &= [J_{g \circ h}((g \circ h)^{-1}(\mathbf{z}))]\mathbf{v}((g \circ h)^{-1}(\mathbf{z})) \\
&= [J_g(h(h^{-1}(g^{-1}(\mathbf{z})))][J_h(h^{-1}(g^{-1}(\mathbf{z})))]\mathbf{v}((h^{-1}(g^{-1}(\mathbf{z}))) \\
&= [J_g(g^{-1}(\mathbf{z}))][J_h(h^{-1}(g^{-1}(\mathbf{z})))]\mathbf{v}((h^{-1}(g^{-1}(\mathbf{z}))) \\
&= [J_g(g^{-1}(\mathbf{z}))]h_* \mathbf{v}((g^{-1}(\mathbf{z})) \\
&= g_*(h_* \mathbf{v})(\mathbf{z}) \ .
\end{aligned} \tag{2.9}$$

What this computation shows is that, essentially as a consequence of the chain rule (**??**), *the push forward of a composition of diffeomorphisms is the composition of their push-forwards.* This will turn out to be a useful fact, so we record it in a theorem:

**2.4 THEOREM** (The push-forward and composition)**.** *For any diffeomorphism $h$ from $U$ to $V$, and any diffeomorphism $g$ from $V$ to $W$ where $U$, $V$ and $W$ are all open sets in $\mathbb{R}^2$,*

$$(g \circ h)_* = g_* \circ h_* \ . \tag{2.10}$$

We close this section by discussing a strategy for solving the equation $\mathbf{x}'(t) = \mathbf{v}(\mathbf{x}(t))$ on and open set $U \subset \mathbb{R}^2$: Find a diffeomorphism $h$ of $U$ onto $V \subset \mathbb{R}^2$ such that

$$h_* \mathbf{v}(u, v) = (a(v), b(v)) \ .$$

That is, such that $h_* \mathbf{v}(u, v)$ depends on $v$ alone. Then we can solve

$$\mathbf{u}' = (u', v') = (a(v), b(v))$$

by separately solving

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} u &= a(v) \\
\frac{\mathrm{d}}{\mathrm{d}t} v &= b(v) \ .
\end{aligned}$$

Note that the second of these equations may be solved by separation of variables. Then with $v(t)$ known, so that the right hand side of the first equation is known, it can be solved by integration.

# 3 How direction fields transform under changes of variables

Consider the differential equations

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y) \tag{3.1}$$

and

$$\mathbf{x}'(t) = a(\mathbf{x}(t))(1, f(\mathbf{x}(t)) \tag{3.2}$$

where $a(x, t)$ is some smooth function that is never zero anywhere on its domain of definition, and of course $\mathbf{x}'(t) = (x'(t), y'(t))$.

Given a curve $\mathbf{x}(t)$ that satisfies (3.2), define

$$x(t) = x_0 + \int_{t_0}^{t} a(\mathbf{x}(s))\mathrm{d}s \ .$$

Since $a(\mathbf{x})$ is never zero, $x(t)$ is a strictly monotone function of $t$, and hence the inverse function $t(x)$ is defined and differentiable. Then

$$y(x) := y(t(x))$$

solves (3.1) since

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}x} = a(x,y)f(x,y)\frac{1}{a(x,y)} = f(x,y) \ .$$

Conversely, suppose that a function $y(x)$ solves (3.1). Define the curve $\mathbf{x}(t)$ by

$$\mathbf{x}(t) = (t, y(t)) \ . \tag{3.3}$$

Then differentiating,

$$\mathbf{x}'(t) = (1, y'(t)) = (1, f(t, y(t))) = (1, f(x(t), y(t))) \ ,$$

so that $\mathbf{x}(t)$ solves (3.2) with $a(x,y) = 1$. This gives us a recipe for passing back and forth between the equations (3.1) and (3.2).

The difference between (3.1) and (3.2) is that (3.1) involves only two variables, $x$ and $y$, and is given by a direction field, while (3.2) involves three variables, $x$, $y$ and $t$, and is given by a vector field. The difference in the number of variables involved accounts for the fact that for any given (smooth and non-vanishing) function $a(\mathbf{x})$, there is a one to one correspondence between solutions of (3.1) and (3.2).

Recall that a direction field on an open set $U$ in $\mathbb{R}^2$ is *non-singular* in case none of the lines it associates to any point of $U$ is either vertical or horizontal. In this case, the line that the direction field associates to the point $(x,y)$ has a non-zero and finite slope $f(x,y)$. The vector field $\mathbf{v}(x,y) = (1, f(x,y))$ is called the *slope field* of the corresponding direction field. This brings us to the following definition;

**3.1 DEFINITION** (Conjugate direction fields)**.** Let $U$ and $V$ be open sets in $\mathbb{R}^2$ with a diffeomorphism $h$ from $U$ onto $V$. Let non-singular direction fields be given on $U$ and $V$ whose slope fields are $\mathbf{v}(x,y) = (1, f(x,y))$ and $\mathbf{w}(u,v) = (1, g(u,v))$ respectively. Then these directions are *conjugate* in case there is a smooth function $a(u,v)$ such that $a(u,v) \neq 0$ anywhere on $V$ and

$$\mathbf{w}(u,v) = a(u,v)h_*\mathbf{v}(u,v) \ , \tag{3.4}$$

where $h_*\mathbf{v}$ is defined in (2.8).

The following theorem is a direct consequence of Theorem 2.2 and what we have said above about the relation between equations (3.1) and (3.2).

**3.2 THEOREM** (Conjugate direction fields)**.** *Let $U$ and $V$ be open sets in $\mathbb{R}^2$ with a diffeomorphism $h$ from $U$ onto $V$. Let continuously differentiable direction fields on $U$ and $V$ be given that are conjugate under $h$. Then $y(x)$ is an integral curve of the given direction field on $U$ if and only if $h(x, y(x))$ is an integral curve of the conjugate direction field on $V$.*

**4 Example** (Homogeneous equations)**.** Consider the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F\left(\frac{y}{x}\right) \tag{3.5}$$

on $x > 0$.

Associated to this equation is the slope field

$$\mathbf{v}(x,y) = (1, F(y/x)) \ .$$

Let $h$ be the homeomorphism of the right half plane onto $\mathbb{R}^2$ considered in Example 2. That is;

$$h(x,y) = (u(x,y), v(x,y) = (\ln(x), y/x) \ .$$

Let us compute $h_*\mathbf{v}$. We find:

$$h_*\mathbf{v}(u,v) = \begin{bmatrix} e^{-u} & 0 \\ -ve^{-u} & e^{-u} \end{bmatrix} (1, F(v)) = e^{-u}(1, F(v) - v) \ .$$

Since $e^{-u}$ is never zero, the slope field $(1, F(v) - v)$ is conjugate to the slope field $(1, F(x/y))$. Consequently, if $(u, v(u))$ is any integral curve of the slope field $(1, F(v) - v)$,

$$h^{-1}(u, v(u))$$

is an integral curve of the slope field $(1, F(x/y))$.

To make this more concrete, let us carry through a specific example. Take $F(y/x) = (y/x)^2$ so that (3.6) becomes

$$\frac{\mathrm{d}}{\mathrm{d}x}y = \left(\frac{y}{x}\right)^2 \ . \tag{3.6}$$

Then the conjugate slope field is $(1, v^2 - v)$, and we find integral equations of this slope field by solving

$$\frac{\mathrm{d}}{\mathrm{d}u}v = v^2 - v \ . \tag{3.7}$$

Note that $v = 0$ and $v = 1$ are equilibrium solutions. All other solutions remain forever in one of the intervals $(-\infty, 0)$, $(0,1)$ or $(1,\infty)$. Since $u$ is not present on the right hand side, we can separate variables to find

$$\mathrm{d}u = \frac{\mathrm{d}v}{v^2 - v} = \left(\frac{1}{v-1} - \frac{1}{v}\right)\mathrm{d}v \ .$$

Integrating, we find, for $v > 1$,

$$u + C = \ln\left(\frac{v-1}{v}\right) \ .$$

Letting $v_0$ denote $v(0)$, we solve to find

$$v(u) = \frac{v_0}{e^u(1 - v_0) + v_0} \ .$$

This function defines an integral curve for $-\infty < u < \ln(v_0/(v_0 - 1)$, at which value of $u$ the curve has a vertical asymptote.

Now simply undo the change of variables to express this curve in terms of $x$ and $y$. Since for $u = 0$, $x = 1$, $y(1) = v(0) = v_0$. Hence we find

$$y(x) = \frac{y_1 x}{x(1 - y_1) + y_1} \ .$$

You may now easily check that $y(x)$ solves

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \left(\frac{y}{x}\right)^2 \qquad \text{and} \qquad y(1) = y_1$$

for $y_1 > 1$. In fact, it solves it for *all* $y_1$, as you can verify by direct computation. (This is another reason to check our work as we go along: It saves us from explicitly considering the cases $0 < v_0 < 1$ and $-\infty < v_0 < 0$.

There is another interesting thing to note about this example. The equation (3.7) has two equilibrium solutions, namely $v = 0$ and $v = 1$, while the equation (3.6) has only one equilibrium solution, namely $y = 0$. The $y = 0$ solution of (3.7) corresponds to the $v = 0$ solution of (3.6) since

$$h(x, 0) = (\ln x, 0) = (u, 0) \ .$$

Now compute

$$h^{-1}(u, 1) = (e^u, e^u) = (x, x) \ ,$$

and hence the $v = 1$ solution of (3.7) corresponds to the solution $y(x) = x$ of (3.6). Although the original slope field is singular only at $y = 0$, the conjugate slope field is singular at $v = 0$ and $v = 1$. Hence we see that even though this diffeomorphism produces singularities, it still enables us to find *all* solutions of (3.6).

# 4 Symmetries of vector fields

Very often, the source of a change of variables; i.e., coordinate transformation, that enables one to solve the differential equation

$$\mathbf{v}'(t) = \mathbf{v}(\mathbf{x}(t))$$

is a *symmetry property* of the vector field $\mathbf{v}$.

**4.1 DEFINITION** (Symmetry of a vector field)**.** Let $U$ be an open set in $\mathbb{R}^2$ and let $\mathbf{v}$ be a continuously differentiable vector field on $U$. A homeomorphism $h$ of $U$ onto $U$ is a *symmetry of* $\mathbf{v}$ in case

$$h_* \mathbf{v} = \mathbf{v} \ .$$

That is, $h$ is a symmetry of $\mathbf{v}$ in case the push forward of $\mathbf{v}$ under $h$ is $\mathbf{v}$ itself.

A one parameter group $\{g^t\}_{t \in \mathbb{R}}$ of diffeomorphisms of $U$ onto itself is a *one parameter symmetry group* of $\mathbf{v}$ in case each $g^t$ is a symmetry of $\mathbf{v}$.

**5 Example** (Rotation invariance)**.** Let us consider one of the most familiar one groups of diffeomorphisms of $\mathbb{R}^2$ into itself: The planar rotation group. This is given by

$$g^\theta(\mathbf{x}) = [R(\theta)]\mathbf{x} \tag{4.1}$$

where $[R(\theta)]$ is the $2 \times 2$ rotation matrix

$$[R(\theta)] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} . \tag{4.2}$$

For any vector $\mathbf{x} \in \mathbb{R}^2$, $[R(\theta)]\mathbf{x}$ is the result of rotating $\mathbf{x}$ counter-clockwise through that angle $\theta$.

Note that since $g^\theta$ is a linear transformation, its derivative; i.e., its Jacobian matrix, is simply

$$[J_{g^\theta}] = [R(\theta)] ,$$

and the right hand side is independent of $\mathbf{x}$. Also, it is clear that the inverse of $g^\theta$ is $g^{-\theta}$. Hence, for any vector field $\mathbf{v}$ on $\mathbb{R}^2$,

$$g^\theta_*\mathbf{v}(\mathbf{x}) = [R(\theta)]\mathbf{v}([R(-\theta)]\mathbf{x}) . \tag{4.3}$$

Now, given a vector field $\mathbf{v}(\mathbf{x}) = (a(\mathbf{x}), b(\mathbf{x}))$, when is $\mathbf{v}$ symmetric under rotations? Let us write $a$ and $b$ in terms of polar coordinates $r$ and $\varphi$:

$$a(\mathbf{x}) = a(r,\varphi) \qquad \text{and} \qquad b(\mathbf{x}) = b(r,\varphi)$$

Then

$$\mathbf{v}([R(-\theta)]\mathbf{x}) = (a(r,\varphi-\theta), b(r,\varphi-\theta)) ,$$

it follows from (4.3), this is the case if and only if

$$(a(r,\varphi), b(r,\varphi)) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} (a(r,\varphi-\theta), b(r,\varphi-\theta)) ,$$

for each $r$, $\varphi$ and $\theta$, which means

$$\begin{aligned} a(r,\varphi) &= \cos\theta a(r,\varphi-\theta) - \sin\theta b(r,\varphi-\theta) \\ b(r,\varphi) &= \sin\theta a(r,\varphi-\theta) + \cos\theta b(r,\varphi-\theta) . \end{aligned}$$

Differentiating both sides of both equations in $\theta$, and evaluating at $\theta = 0$, we find

$$\begin{aligned} 0 &= -a(r,\varphi) + b'(r,\varphi) \\ 0 &= -a'(r,\varphi) + b(r,\varphi) \end{aligned}$$

where primes denote derivatives with respect to the angular variable. Hold $r$ fixed as a parameter, and consider $a$ and $b$ as functions of $\varphi$. Then we see that $\mathbf{v}$ is invariant if and only if these functions satisfy the system of differential equations

$$\begin{aligned} b'(r,\varphi) &= -a(r,\varphi) \\ a'(r,\varphi) &= b(r,\varphi) \end{aligned}$$

But we know the general solution to this system: It is

$$\begin{aligned} a(r,\varphi) &= a(r,0)\cos\varphi + b(r,0)\sin\varphi \\ b(r,\varphi) &= -a(r,0)\sin\varphi + b(r,0)\cos\varphi \end{aligned} \tag{4.4}$$

Now define, for $r > 0$,

$$f_1(r) = \frac{b(r,0)}{r} \qquad \text{and} \qquad f_2(r) = \frac{a(r,0)}{r} \ .$$

Then we can rewrite (4.4) as

$$(a(r,\varphi), b(r\varphi)) = (f_2(\|\mathbf{x}\|)y + f_1(\|\mathbf{x}\|)x \ , \ -f_2(\|\mathbf{x}\|)x + f_1(\|\mathbf{x}\|)y) \ .$$

We have proved the following theorem:

**4.2 THEOREM** (Characterization of rotation symmetric vector fields)**.** *A vector field* $\mathbf{v}$ *on* $\mathbb{R}^2$ *is symmetric under rotations if and only if it has the form*

$$\mathbf{v}(x,y) = (f_2(\|\mathbf{x}\|)y + f_1(\|\mathbf{x}\|)x \ , \ -f_2(\|\mathbf{x}\|)x + f_1(\|\mathbf{x}\|)y) \ . \tag{4.5}$$

*for some functions* $f_1$ *and* $f_2$*.*

In particular, the vector field $\mathbf{v}$ given in (2.6) of Example 3 has this form: In this case $f_1(r) = 1 - r^2$ and $f_2(r) = 1$. Hence it is symmetric under the planar rotation group. As we shall see, this is why the change to polar coordinates lead to an explicit solution of the corresponding equation.

More generally, let $h$ be the polar coordinate diffeomorphism, and let us compute $h * \mathbf{v}$ for $\mathbf{v}$ given by (**??**). Computing as in Example 3, we find

$$\begin{aligned} h_*\mathbf{v}(r,\theta) &= \begin{bmatrix} \cos\theta & \sin\theta \\ -(\sin\theta)/r & (\cos\theta)/r \end{bmatrix} (f_2(r)r\sin\theta + f_1(r)r\cos\theta \ , \ -f_2(r)r\cos\theta + f_1(r)r\sin\theta) \\ &= (f_1(r)r, -f_2(r) \ . \end{aligned}$$

Therefore, solving the equation $\mathbf{x}'(t)$ is reduced to solving

$$\begin{aligned} r' &= rf_1(r) \\ \theta' &= -f_2(r) \ . \end{aligned}$$

## 5 Symmetries of direction fields

We are also concerned with symmetries of direction fields. Again, symmetry means invariance under a transformation, but since we are only concerned with invariance of directions and not magnitudes, the notion of symmetry for direction fields is more comprehensive.

**5.1 DEFINITION** (Symmetry of a direction field)**.** Let $U$ be an open set in $\mathbb{R}^2$ and consider a non-singular direction field on $U$ whose slope field is $\mathbf{v} = (1, f(x,y))$. A homeomorphism $h$ of $U$ onto $U$ is a *symmetry of this direction field* in case

$$h_*\mathbf{v} = a\mathbf{v}$$

where $a(x,y)$ is some function that is not zero for any $(x,y)$ in $U$.

A one parameter group $\{g^t\}_{t\in\mathbb{R}}$ of diffeomorphisms of $U$ onto itself is a *one parameter symmetry group of this direction field* in case each $g^t$ is a symmetry of of this direction field. Note that a vector field is symmetric under a diffeomorphism if and only if it is *invariant* under the action of the diffeomorphism. The terms symmetry and invariance are often used interchangeably in this context.

**6 Example** (Scale invariance)**.** Here is another simple diffeomorphism group that is relevant to the solution of many differential equations: Fot $t \in \mathbb{R}$, define

$$g^t(x,y) = (e^t x, e^t y) \ . \tag{5.1}$$

It is easy to check that $g^t \circ g^s = g^{t+s}$, that $g^0$ is the identity transformation, and that $g^{-t} = (g^t)^{-1}$. Finally, $g^t$ is linear, given by

$$g^t(\mathbf{x}) = [S(t)]\mathbf{x} \tag{5.2}$$

where $S(t)$ is the $2 \times 2$ *scaling matrix*

$$[S(t)] = \begin{bmatrix} e^t & 0 \\ 0 & e^t \end{bmatrix} \ . \tag{5.3}$$

Again since for each $t$, $g^t$ is a linear transformation of $\mathbb{R}^2$ to $\mathbb{R}^2$, it is differentiable and its derivative; i.e., its Jacobian matrix, is given by

$$[J_{g^t}] = [S(t)] = \begin{bmatrix} e^t & 0 \\ 0 & e^t \end{bmatrix} \ .$$

Again, the right hand side is independent of $\mathbf{x}$.

Hence, for any vector field $\mathbf{v}$ on $\mathbb{R}^2$,

$$g^t_*\mathbf{v}(\mathbf{x}) = [S(t)]\mathbf{v}([S(-t)]\mathbf{x}) \ .$$

If the vector field $\mathbf{v}$ is linear; i.e., if there is a $2 \times 2$ (constant) matrix $A$ so that

$$\mathbf{v}(\mathbf{x}) = A\mathbf{x} \ ,$$

then

$$g^t_*\mathbf{v}(\mathbf{x}) = [S(t)]A([S(-t)]\mathbf{x}) = [S(t)][S(-t)](A\mathbf{x}) = A\mathbf{x} = \mathbf{x}(\mathbf{x}) \ ,$$

since, while matrix multiplication is not generally commutative, $[S(t)]$ does commute with every other $2 \times 2$ matrix – it is a multiple of the identity.

Thus, every linear vector field is conjugate to itself under scale transformations. However, a *much broader* class of direction fields are conjugate to themselves under scale transformations.

Recall that a nonsingular direction field is conjugate to itself under a transformation $h$ in case the corresponding slope field $\mathbf{v} = (1, f(x,y))$ satisfies

$$h_*\mathbf{v}(\mathbf{x}) = a(\mathbf{x})\mathbf{v}(\mathbf{x}) \tag{5.4}$$

for some function $a(\mathbf{x})$ with $a(\mathbf{x})$ non-zero everywhere.

Hence, the direction field whose slope field is $\mathbf{v} = (1, f(x,y))$ is self-conjugate under scaling transformations in case for each $t$,

$$[S(t)]\mathbf{v}([S(-t)]\mathbf{x})) = a(\mathbf{x})\mathbf{v}(\mathbf{x})$$

for some function $a(\mathbf{x})$ with $a(\mathbf{x})$ non-zero everywhere.

For $\mathbf{v}(x,y) = (1, f(x,y))$,

$$\mathbf{v}([S(-t)]\mathbf{x})) = (1, f(e^t x, e^t y)) \ .$$

Therefore, if

$$f(x,y) = F\left(\frac{y}{x}\right) \ ,$$

$\mathbf{v}([S(-t)]\mathbf{x})) = \mathbf{v}(\mathbf{x})$ and then

$$g_*^t \mathbf{v}(\mathbf{x}) = e^t \mathbf{v}(\mathbf{x}) \ .$$

Thus, with $g^t$ in place of $h$, (5.4) is satisfied with $a(\mathbf{x}) = e^t$ – independent of $\mathbf{x}$, and certainly non-zero.

The fact that direction fields whose slope fields have the form $(1, F(y/x))$ are self-conjugate under scale transformations is the source of the change of coordinates we used in Example 4.

# 6 Constructing canonical coordinates for a symmetry

As we shall now show, whenever one can find a one parameter group of symmetries of a vector field or a non-singular direction field on $\mathbb{R}^2$, one can construct a *canonical* set of coordinates so that changing variables to the new coordinates reduces the problem of solving the differential equation to integration. In this section, we explain how to construct these coordinates. In the next section we explain why these coordinates are useful for solving differential equations.

The *orbits* of the one parameter transformation group $g^t$ are the points on the curves $\mathbf{x}(u) = g^u(\mathbf{x}_0)$ for each fixed $\mathbf{x}_0$. Any two orbits either coincide or are disjoint. Suppose that $\mathbf{x}(v)$, $a < v < b$, is a curve such that $\mathbf{x}(v)$ passes through each orbit exactly once. (We allow, but do not requires, $a = -\infty$ or $b = \infty$.) It is usually easy to find such a curve, and once this is done, the rest of the construction is canonical; i.e., by the rules. Here is the rule:

Define a transformation $h^{-1}(u,v)$ from

$$V := \mathbb{R} \times (a,b)$$

to $U$ by

$$h^{-1}(u,v) = g^u(\mathbf{x}(v)) \ .$$

Since each $(x,u) \in U$ lies on exactly one orbit of the group $\{g^t\}$, and since $\mathbf{x}(v)$ intersects each orbit exactly once, for each $(x,y) \in U$, there is a unique $u \in \mathbb{R}$ and $v \in (a,b)$ such that

$$(x,y) = g^u(\mathbf{x}(v)) = h^{-1}(u,v) \ .$$

Thus $h^{-1}$ is an invertible transformation from $V$ to $U$, and its inverse $h$ is an invertible transformation from $U$ to $V$.

To investigate the differentiability, introduce the vector field $\mathbf{z}(\mathbf{x})$ associated to $\{g^t\}$:

$$\mathbf{z}(\mathbf{x}) = \lim_{t\to 0} \frac{1}{t}(g^t(\mathbf{x}) - \mathbf{x}) \ . \tag{6.1}$$

We then have:

**6.1 THEOREM** (Coordinates from symmetry groups)**.** *With $h$ defined as above, through $h^{-1}$, $h$ is a diffeomorphism provided that for each $v$,*

$$\det([\mathbf{z}(\mathbf{x}(v)), \mathbf{x}'(v)]) \neq 0 \ . \tag{6.2}$$

**Proof:** From what we have said above, it suffice to show that $h^{-1}$ and $H$ are both differentiable. We start with $h^{-1}$, which is directly defined.

$$\begin{aligned}
\frac{\partial}{\partial u} h^{-1}(u,v) &= \lim_{t\to 0} \frac{1}{t}(g^t(h^{-1}(u,v)) - h^{-1}(u,v)) \\
&= \lim_{t\to 0} \frac{1}{t}(g^t(g^u(\mathbf{x}(v))) - g^u(\mathbf{x}(v))) \\
&= \lim_{t\to 0} \frac{1}{t}(g^{t+u}(\mathbf{x}(v))) - g^u(\mathbf{x}(v))) \\
&= \lim_{t\to 0} \frac{1}{t}(g^u(g^t(\mathbf{x}(v))) - g^u(\mathbf{x}(v))) \\
&= [J_{g^u}(\mathbf{x}(v)]\mathbf{z}(\mathbf{x}(v)) \ ,
\end{aligned}$$

where in the last line we have used the chain rule.

Much more directly, by the chain rule,

$$\frac{\partial}{\partial v} h^{-1}(u,v) = \lim_{t\to 0} \frac{1}{t}(g^u(\mathbf{x}(v+t)) - g^u(\mathbf{x}(v))) = [J_{g^u}(\mathbf{x}(v)]\mathbf{x}'(v) \ .$$

Thus $h^{-1}$ is differentiable, and its Jacobian is

$$[J_{h^{-1}}(u,v)] = \begin{bmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{bmatrix} = \big[[J_{g^u}(\mathbf{x}(v)]\mathbf{z}(\mathbf{x}(v)), [J_{g^u}(\mathbf{x}(v)]\mathbf{x}'(v)\big] \ ,$$

where for any $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^2$, $[\mathbf{a}, \mathbf{b}]$ denotes the $2\times 2$ matrix whose fist column is $\mathbf{a}$, and whose second column is $\mathbf{b}$. Then since for any $2\times 2$ matrix $A$, $[A\mathbf{a}, A\mathbf{b}] = A[\mathbf{a}, \mathbf{b}]$, we have

$$[J_{h^{-1}}(u,v)] = [J_{g^u}(\mathbf{x}(v)][\mathbf{z}(\mathbf{x}(v)), \mathbf{x}'(v)] \ .$$

Since $g^u$ is a diffeomorphism, $[J_{g^u}(\mathbf{x}(v)]$ is invertible. Hence as long as $[\mathbf{z}(\mathbf{x}(v)), \mathbf{x}'(v)]$ is invertible for each $v$, $[J_{h^{-1}}(u,v)]$ is the product of invertible matrices, and hence is invertible. It then follows from the Inverse Function Theorem that $h$ is also differentiable. Now note that the condition that the condition $\det([\mathbf{z}(\mathbf{x}(v)), \mathbf{x}'(v)]) \neq 0$ implies that $[\mathbf{z}(\mathbf{x}(v)), \mathbf{x}'(v)]$ is invertible. □

**7 Example** (From rotation symmetry to polar coordinates)**.** Consider the planar rotation group $\{g^\theta\}$ defined by (4.1) and (4.2). The orbits of this group are the circles centered on the origin. The positive $x$-axis intersects each of the orbits exactly once. Parameterize the positive $x$-axis by $\mathbf{x}(r) = (r, 0)$, $r > 0$. Here, we use $r$ in place of $v$. Then

$$h^{-1}(r,\theta) = \{g^\theta\}(\mathbf{x}(r)) = [R(\theta)](r,0) = (r\cos\theta, r\sin\theta) \ .$$

Writing $(x, y) = h^{-1}(r, \theta)$, we have

$$x = r\cos\theta \qquad \text{and} \qquad y = r\sin\theta \ .$$

Thus, our method of constructing coordinate systems out of symmetries leads from rotation symmetry to polar coordinates.

**8 Example** (From scale symmetry to homogeneous coordinates)**.** Consider the scale transformation group $\{g^t\}$ defined by (5.1), (5.2) and (5.3). The line $\mathbf{x}(v) = (1, v)$ intersects each orbit of $\{g^u\}$ in the right half-plane exactly once. Hence we define

$$(x, y) = h^{-1}(u, v) = g^u(1, v) = \begin{bmatrix} e^u & 0 \\ 0 & e^u \end{bmatrix} (1, v) = (e^u, e^u v) \ .$$

Solving for $u$ and $v$ in terms of $x$ and $y$, we find

$$u = \ln(x) \qquad \text{and} \qquad v = \frac{y}{x} \ .$$

Thus, our method of constructing coordinate systems out of symmetries leads from scale symmetry to homogeneous coordinates.

In the last two examples, we have seen that our method of constructing coordinates out of symmetries leads to some familiar coordinate systems. But what is so spacial about this method? The answer is that the action of a symmetry group is very simple when expressed in the corresponding coordinates.

Suppose that $U$ and $V$ are open sets in $\mathbb{R}^2$, and $h$ is a diffeomorphism from $U$ onto $V$. Suppose also that $\{g^t\}$ is a one parameter family of diffeomorphisms of $U$ onto itself. Define

$$\widetilde{g}^t = h \circ g^t \circ h^{-1} \ . \tag{6.3}$$

Since compositions of diffeomorphisms are diffeomorphisms, $\widetilde{g}^t$ is a diffeomorphism from $V$ onto $V$. Moreover,

$$\widetilde{g}^s \circ \widetilde{g}^t = (h \circ g^s \circ h^{-1}) \circ (h \circ g^t \circ h^{-1}) = h \circ g^{s+t} \circ h^{-1} = \widetilde{g}^{s+t} \ ,$$

and clearly $\widetilde{g}^0$ is the identity transformation. Hence $\{\widetilde{g}^t\}$ is a one parameter group of diffeomorphisms of $V$.

**6.2 DEFINITION** (Conjugate groups of diffeomorphisms)**.** Suppose that $U$ and $V$ are open sets in $\mathbb{R}^2$, and $h$ is a diffeomorphism from $U$ onto $V$. Suppose also that $\{g^t\}$ is a one parameter family of diffeomorphisms of $U$ onto itself. Then with $\widetilde{g}^t$ defined by (6.3), $\{\widetilde{g}^t\}$ is the one parameter group of diffeomorphisms of $V$ conjugate to $\{g^t\}$ under $h$.

**6.3 THEOREM** (Transformation to translation)**.** *Suppose that $\{g^t\}$ is a one parameter family of diffeomorphisms of an open set $U \subset \mathbb{R}^2$ onto itself. Suppose that $\mathbf{x}(v)$, $a < v < b$ is a continuously differentiable curve that intersects each orbit of $\{g^u\}$ exactly once and such that for each $v$, (6.2) is true where $\mathbf{z}$ is defined by (6.1). Let $h(\mathbf{x})$ be the diffeomorphism from $U$ onto $V$ defined implicitly by $h^{-1}(u, v) = g^u(\mathbf{x}(v)$, noting that Theorem 6.1 ensures that this is a diffeomorphism. Then*

$$\widetilde{g}^s(u, v) = h(g^s(h^{-1}(u, v))) = (u + s, v) \ .$$

*That is, the one parameter group of diffeomorphisms of $V$ conjugate to $\{g^t\}$ under $h$ acts on $V$ by translation.*

**Proof:** We compute

$$\begin{aligned} h(g^s(h^{-1}(u, v))) &= h(g^s(g^u(\mathbf{x}(v)) \\ &= h((g^{s+u}(\mathbf{x}(v)) \\ &= (u + s, v) \ . \end{aligned}$$

□

# 7 How symmetries lead to useful coordinate transformations

Let $\{g^t\}$ be a one parameter group of diffeomorphisms of an open set $U \subset \mathbb{R}^2$ onto itself. Suppose that $h$ is a diffeomorphism of $U$ onto another open set $V \subset \mathbb{R}^2$ so that the conjugate one parameter group of diffeomorphisms $\{\widetilde{g}^t\}$, where $\widetilde{g}^t = h \circ g^t \circ h^{-1}$, acts on $V$ by translation in one coordinate:

$$\widetilde{g}^t(u,v) = (u+t, v) \ .$$

We have already seen how to construct such a diffeomorphism $h$, starting from $\{g^t\}$. Note that the Jacobian of $\widetilde{g}^t$ is simply the identity:

$$J_{\widetilde{g}^t}(u,v) = \left[\frac{\partial g^t(u,v)}{\partial u}, \frac{\partial g^t(u,v)}{\partial v}\right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ .$$

Therefore, if $\mathbf{w}(u,v)$ is *any* vector field on $V$, its push-forward under $\widetilde{g}^t$ is very easy to work out:

$$\widetilde{g}^t_*\mathbf{w}(u,v) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{w}(u-t,v) = \mathbf{w}(u-t,v) \ .$$

We observe that a vector field $\mathbf{w}$ on $V$ is invariant under the one parameter group of diffeomorphisms $\{\widetilde{g}^t\}$ if and only if $\mathbf{w}$ does not depend on $u$; i.e., if and only if

$$\mathbf{w}(u,v) = (a(v), b(v))$$

for some functions $a$ and $b$.

Next, we observe that if $\mathbf{v}$ is any vector field on $U$ that is invariant under the original one parameter group of diffeomorphisms $\{g^t\}$, then the push-forward of $\mathbf{v}$ under $h$ is invariant under the conjugate group $\{\widetilde{g}^t\}$.

To see this, note that since $\widetilde{g}^t = h \circ g^t \circ h^{-1}$, Theorem 2.4 gives us

$$\widetilde{g}^t_* = h_* \circ g^t_* \circ h^{-1}_* \ .$$

But then

$$\widetilde{g}^t_*(h_*\mathbf{v}) = (h_* \circ g^t_* \circ h^{-1}_*)(h_*\mathbf{v}) = h_*(g^t_*\mathbf{v}) = h_*\mathbf{v} \ .$$

We have proved:

**7.1 THEOREM** (From symmetry to good coordinates)**.** *Let $\mathbf{v}$ be vector field on $U$, and open subset of $\mathbb{R}^2$. Let $\{g^t\}$ be a one parameter group of diffeomorphisms of $U$ such that each $g^t$ is a symmetry of $\mathbf{v}$. Let $h$ be a diffeomorphism of $U$ onto $V \subset \mathbb{R}^2$ so that the conjugate one parameter group of diffeomorphisms $\{\widetilde{g}^t\}$ acts on $V$ by translation in one coordinate:*

$$\widetilde{g}^t(u,v) = (u+t, v) \ .$$

*Then, the conjugate vector field $\mathbf{w} := h_*\mathbf{v}$ has the special form*

$$\mathbf{w}(u,v) = (a(v), b(v))$$

*for some functions $a$ and $b$.*

The same sort of thing works for direction fields. Consider a direction field on $U$ whose slope fields is $\mathbf{v} = (1, f(x,y))$. Suppose that $\{g^t\}$ is a one parameter group of diffeomorphisms of $U$ such that for each $t$,

$$g^t_* \mathbf{v}(\mathbf{x}) = c_t(\mathbf{x})\mathbf{v}(\mathbf{x}) \tag{7.1}$$

where $c_t(\mathbf{x})$ does not equal zero for any $t$ or $\mathbf{x} \in U$. That is. suppose that each $g^t$ us a symmetry of the given direction field. Let $h$ be a diffeomorphism of $U$ onto $V \subset \mathbb{R}^2$ so that the conjugate one parameter group of diffeomorphisms $\{\widetilde{g}^t\}$ acts on $V$ by translation in one coordinate: $\widetilde{g}^t(u,v) = (u+t, v)$, as above. Define $\mathbf{w} = h_*\mathbf{w}$. Then, computing as above,

$$\widetilde{g}^t_* \mathbf{w}(u,v) = h_* g^t_* \mathbf{v}(u,v) = h_* c_t \mathbf{v}(u,v) = c_t(h^{-1}(u,v))\mathbf{w}(u,v) \ .$$

But we also have

$$\widetilde{g}^t_* \mathbf{w}(u,v) = \mathbf{w}(u+t,v) \ .$$

Since the slope associated to the vector field $c_t(h^{-1}(u,v))\mathbf{w}(u,v)$ is the same as the slope associated to the vector field $\mathbf{w}$, since $c_t$ is never zero, it follows that $\mathbf{w}(u+t,v)$ has the same slope as $\mathbf{x}(u,v)$ for all $t$ and all $u$ and $v$.

Thus, the change of variables associated to $h$ eliminates one variable, namely $u$, from the slope field. The transformed equation can then be solved by separation of variables

**9 Example** (Quasi-homogenous coordinates)**.** Let non-zero real numbers $\alpha$ and $\beta$ be given, and consider the one parameter group $\{g^t\}$ of transformation of the right half plane; i.e. $U = \{(x,y) : x > 0\}$ given by

$$g^t(x,y) = (e^{\alpha t}x, e^{\beta t}y) \ . \tag{7.2}$$

Again, this is linear, and so

$$[J_{g^t}(\mathbf{x})] = \begin{bmatrix} e^{\alpha t} & 0 \\ 0 & e^{\beta t} \end{bmatrix} \ .$$

Hence, for any vector field $\mathbf{v}(\mathbf{x})$ on $U$,

$$g^t_* \mathbf{v}(\mathbf{x}) = \begin{bmatrix} e^{\alpha t} & 0 \\ 0 & e^{\beta t} \end{bmatrix} \mathbf{v}(e^{-\alpha t}x, e^{-\beta t}y) \ . \tag{7.3}$$

We now prove:

**7.2 THEOREM** (Quasi-homogeneous direction fields)**.** *The direction field with slope field* $\mathbf{v} = (1, f(x,y))$ *is symmetric under the quasi-homogenous scale transformation group* $\{g^t\}$ *given by (7.2) if and only if*

$$f(e^{-\alpha t}x, e^{-\beta t}y) = e^{(\alpha-\beta)t} f(x,y) \ , \tag{7.4}$$

*for all* $t$.

**Proof:** Let $\mathbf{v}(\mathbf{x}) = (1, f(\mathbf{x}))$. Then (7.3) simplifies to

$$\begin{aligned} g^t_* \mathbf{v}(\mathbf{x}) &= (e^{t\alpha}, e^{t\beta} f(e^{-\alpha t}x, e^{-\beta t}y) \\ &= e^{t\alpha}(1 \ , \ e^{(\beta-\alpha)t} f(e^{-\alpha t}x, e^{-\beta t}y)) \end{aligned}$$

Thus, (7.1) is satisfied for this $\mathbf{v}$ if and only if $e^{(\beta-\alpha)t}f(e^{-\alpha t}x, e^{-\beta t}y) = f(x,y)$, in which case (7.1) is true with $c_t(\mathbf{x}) = e^{\alpha t}$. $\square$

Hence, whenever (7.4) is true, the canonical coordinates associated to $\{g^t\}$ will render the equation

$$\frac{\mathrm{d}}{\mathrm{d}x}y = f(x,y) \tag{7.5}$$

separable.

Here is a useful characterization of a class of functions that satisfy (7.4).

**7.3 THEOREM** (Quasi-homogeneous slopes)**.** *Let $f(x,y)$ have the form*

$$f(x,y) = \sum_{j=1}^{m} a_j x^{p_j} y^{q_j} \ ,$$

*For some positive integer $m$, non-negative integers $q_j$, and some integers $p_j$. Define $p_0 = 1$ and $q_0 = 1$. Then $f(x,y)$ satisfies (7.4) if and only if for some number $s$,*

$$\alpha p_j + \beta q_j = s \qquad \text{for all} \qquad j = 0, 1, \ldots, m \ .$$

*That is, $f(x,y)$ satisfies (7.4) if and only if all of the points $(p_0,q_0), (p_1,q_1), \ldots, (p_m,q_m)$ lie on a common line in the $p,q$ plane.*

**Proof:** By direct computation we find

$$e^{(\beta-\alpha)t}f(e^{-\alpha t}x, e^{-\beta t}y) = \sum_{j=1}^{m} e^{(\beta-\alpha)t-\alpha p_j t-\beta q_j t} a_j x^{p_j} y^{q_j} \ .$$

The right hand side can equal $f(x,y)$ (which is independent of $t$) if and only if

$$e^{[(\beta-\alpha)-\alpha p_j-\beta q_j]t} = 1$$

for each $j = 1, \ldots, m$, and this is the case if and only if

$$\alpha p_j + \beta q_j = \beta - \alpha$$

for each $j = 1, \ldots, m$. Since $(p_0, q_0)$ satisfies this equation, we see that this condition is equivalent to the condition that all of the points $(p_0,q_0), (p_1,q_1), \ldots, (p_m,q_m)$ lie on a common line in the $p,q$ plane. $\square$

**10 Example** (A quasi-homogeneous slope field)**.** Let For example, consider

$$f(x,y) = -x^{-3} - \frac{2}{x}y + xy^2 \ . \tag{7.6}$$

We have

$$(p_0, q_o) = (-1,1) \qquad (p_1,q_1) = (-3,0) \qquad (p_2,q_2) = (-1,1) \qquad \text{and} \qquad (p_3,q_3) = (1,2) \ .$$

These points all lie on a common line of slope 1/2. Since the slope is 1/2, the vector $(-1,2)$ in the $p,q$ plane is orthogonal to this line, and so (7.4) is satisfied with

$$\alpha = -1 \qquad \text{and} \qquad \beta = 2 \ .$$

Notice that plotting the points not only tells us whether $f(x,y)$ satisfies (**??**) for some $\alpha$ and $\beta$; it also gives us the values of $\alpha$ and $\beta$.

**11 Example** (Constructing quasi-homogeneous coordinates)**.** For constructing the coordinates, notice that each orbit of $\{g^t\}$ intersects the line $x = 1$ exactly once. This line is parameterized by $\mathbf{x}(v) = (1, v)$, We define $h^{-1}(u, v) = g^u\mathbf{x}(v)$. We find

$$(x, y) = h^{-1}(u, v) = \begin{bmatrix} e^{u\alpha} & 0 \\ 0 & e^{u\beta} \end{bmatrix} (1, v) = (e^{u\alpha}, e^{u\beta}v) \ .$$

Thus,

$$x(u, v) = e^{u\alpha} \qquad \text{and} \qquad y(u, v) = e^{u\beta}v \ .$$

Inverting we find

$$u(x, y) = \frac{1}{\alpha}\ln x \qquad \text{and} \qquad v(x, y) = x^{-\beta/\alpha}y \ .$$

The diffeomorphism $h$ then is

$$h(x, y) = \left(\frac{1}{\alpha}\ln x \ , \ x^{-\beta/\alpha}y\right) \ . \tag{7.7}$$

**12 Example** (Solving an equation via quasi-homogeneous coordinates)**.** Consider the function $f(x, y)$ given by (7.6). Then as we have seen, (7.4) is satisfied with $\alpha = -1$ and $\beta = 2$. Then from (7.7), we find that quasi-homogeneous coordinate transformation to use for this example is

$$h(x, y) = (-\ln x, x^2y) \ .$$

Thus,

$$[J_h] = \begin{bmatrix} -1/x & 0 \\ 2xy & x^2 \end{bmatrix} = \begin{bmatrix} -e^u & 0 \\ 2e^uv & e^{-2u} \end{bmatrix} \ .$$

Hence

$$h_*(1, f)(u, v) = \begin{bmatrix} -e^u & 0 \\ 2e^uv & e^{-2u} \end{bmatrix} (1, -e^{3u} - 2e^{3u}v + e^{3u}v^2) = e^u(-1, -1 + v^2) \ .$$

Hence the transformed slope field is simply

$$(1, v^2 - 1) \ .$$

Thus, this change of variables reduces the solution of the Riccatit equation

$$\frac{\mathrm{d}}{\mathrm{d}x}y = -x^{-3} - \frac{2}{x}y + xy^2$$

to the solution of

$$\frac{\mathrm{d}}{\mathrm{d}u}v = v^2 - 1 \ .$$

This can be solved separating variables. computing as usual, one finds

$$v(u) = \frac{(v_0 + 1) + (v_0 - 1)e^{2u}}{(v_0 + 1) - (v_0 - 1)e^{2u}}$$

where $v_0 = v(0)$.

Translating this back into $x$ and $y$ terms, one finds

$$y(x) = \frac{(y_1 + 1)x^2 + (y_1 - 1)}{(y_1 + 1)x^4 - (y_1 - 1)x^2}$$

where $y_1 = y(1)$.