Appendix D

Coordinate Systems and Canonical Forms

D.1. Local Coordinates

Let O be an open set in $\mathbf{R}^n$. We say that an n-tuple of smooth real-valued functions defined in O, $(\phi_1, \ldots, \phi_n)$, forms a *local coordinate system* for O if the map $\phi : p \mapsto (\phi_1(p), \ldots, \phi_n(p))$ is a *diffeomorphism*, that is, if it is a one-to-one map of O onto some other open set U of $\mathbf{R}^n$ and if the inverse map $\psi := \phi^{-1} : U \to \mathbf{R}^n$ is also smooth. The relation of ψ to ϕ is clearly completely symmetrical, and in particular ψ defines a coordinate system $(\psi_1, \ldots, \psi_n)$ in U.

D.1.1. Remark. By the Inverse Function Theorem, the necessary and sufficient condition for ϕ to have a smooth inverse in some neighborhood of a point p is that the $D\phi_p$ is an invertible linear map, or equivalently that the differentials $(d\phi_i)_p$ are linearly independent and hence a basis for $(\mathbf{R}^n)^*$. In other words, given n smooth real-valued functions $(\phi_1, \ldots, \phi_n)$ defined near p and having linearly independent differentials at p, they always form a coordinate system in some neighborhood O of p.

The most obvious coordinates are the "standard coordinates", $\phi_i(p) = p_i$, with O all of $\mathbf{R}^n$ (so ϕ is just the identity map). If $e_1, \ldots, e_n$ is the standard basis for $\mathbf{R}^n$, then $\phi_1, \ldots, \phi_n$ is just the dual basis for $(\mathbf{R}^n)^*$. We will usually denote these standard coordinates by $(x^1, \ldots, x^n)$. More generally, given any basis $f_1, \ldots, f_n$ for $\mathbf{R}^n$, we can let $(\phi_1, \ldots, \phi_n)$ be the corresponding dual basis. Such coordinates are called *Cartesian*. In this case, ϕ is the linear isomorphism of $\mathbf{R}^n$ that maps e_i to f_i. If the f_i are orthonormal, then these are called orthogonal Cartesian coordinates and ϕ is an orthogonal transformation.

Why not just always stick with the standard coordinates? One reason is that once we understand a concept in $\mathbf{R}^n$ in terms of arbitrary coordinates, it is easy to make sense of that concept on a general "differentiable manifold". But there is another important reason. Namely, it is often possible to simplify the analysis of a problem considerably by choosing a well-adapted coordinate system. In more detail, various kinds of geometric and analytic objects have numerical descriptions in terms of a coordinate system. This observation by Descartes is of course the basis of the powerful "analytic geometry" approach to studying geometric questions. Now, the precise numerical description of an object is usually highly dependent on the choice of coordinate system, and it can be more or less complicated depending on that choice. Frequently, there will be certain special "adapted" coordinates with respect to which the numerical description of the object has a particularly simple so-called "canonical form", and facts that are difficult to deduce from the description with respect to general coordinates can be obvious from the canonical form.

Here is a well-known simple example. An ellipse in the plane, $\mathbf{R}^2$, is given by an implicit equation of the form $ax^2+by^2+cxy+dx+ey+f=0$, but if we choose the diffeomorphism that translates the origin to the center of the ellipse and rotates the coordinate axes to be the axes of the ellipse, then in the resulting coordinates ξ, η the implicit equation for the ellipse will have the simpler form $\alpha^2\xi^2 + \beta^2\eta^2 = 1$. Notice that this diffeomorphism is actually a Euclidean motion, so ξ and η are orthogonal Cartesian coordinates and even the metric properties of the ellipse are preserved by this change of coordinates. If that is not important in some context, we could instead use $u := \alpha\xi$ and $v := \beta\eta$ as our coordinates and work with the even simpler equation $u^2 + v^2 = 1$.

This example illustrates the general idea behind choosing coordinates adapted to a particular object Ω in some open set O. Namely, you should think intuitively of finding a diffeomorphism ϕ that moves, bends, and twists Ω into an object Ω^ϕ in $U = \phi(O)$, one that is in a "canonical configuration" having a particularly simple description with respect to standard coordinates. While one can apply this technique to all kinds of geometric and analytic objects, here we will concentrate on three of the objects of greatest interest to us, namely real-valued functions, $f : O \to \mathbf{R}$; smooth curves, $\sigma : I \to O$; and vector fields V defined in O.

First let us consider how to define f^ϕ, σ^ϕ, and V^ϕ. For functions and curves the definition is almost obvious; namely $f^\phi : U \to \mathbf{R}$ is defined by $f^\phi := f \circ \phi^{-1}$ (so that $f^\phi(\phi_1(x), \ldots \phi_n(x)) = f(x_1, \ldots, x_n)$ for all $x \in O$) and $\sigma^\phi : I \to U$ is defined by $\sigma^\phi := \phi \circ \sigma$.

Defining the vector field V^ϕ in U from the vector field V in O is slightly more tricky. At p in O, V defines a tangent vector, $(p, V(p))$, that the differential of ϕ at p maps to a tangent vector at $q = \phi(p)$, $V^\phi(q) := (q, D\phi_p(V(p)))$. Written explicitly in terms of q (and dropping the first component) gives the somewhat ugly formula $V^\phi(q) := D\phi_{\phi^{-1}(q)} V(\phi^{-1}(q))$, but note that if the diffeomorphism ϕ is linear (i.e., if the coordinates $(\phi_1, \ldots, \phi_n)$ are Cartesian), then $D\phi_p = \phi$, so the formula simplifies to $V^\phi = \phi V \phi^{-1}$.

▷ **Exercise D–1.** Recall that the radial (or Euler) vector field R on $\mathbf{R}^n$ is defined by $R(x) = x$, or equivalently, written as a differential operator using standard coordinates, $R = \sum_{i=1}^n x^i \frac{\partial}{\partial x^i}$. Show that if ϕ is any linear diffeomorphism of $\mathbf{R}^n$, then $R^\phi = R$, or equivalently, if $(y_1, \ldots, y_n)$ is any Cartesian coordinate system, then $R = \sum_{i=1}^n y_i \frac{\partial}{\partial y_i}$. That is, the radial vector field has the remakable property that it "looks the same" in all Cartesian coordinate systems. Show that any linear vector field L on $\mathbf{R}^n$ with this property must be a constant multiple of the radial field. (Hint: The only linear transformations that commute with all linear isomorphisms of $\mathbf{R}^n$ are constant multiples of the identity.)

▷ **Exercise D–2.** Let f, σ, and V be as above.

a) Show that if σ is a solution curve of V, that is, if $\sigma'(t) = V(\sigma(t))$, then σ^ϕ is a solution of V^ϕ.

b) Show that if f is a constant of the motion for V (i.e., $Vf \equiv 0$), then f^ϕ is a constant of the motion for V^ϕ.

A common reason for using a particular coordinate system is that these coordinates reflect the symmetry properties of a geometrical problem under consideration. While Cartesian coordinates are good for problems with translational symmetry, they are not well adapted to problems with rotational symmetry, and the analysis of such problems can often be simplified by using some sort of polar coordinates. In $\mathbf{R}^2$ we have the standard polar coordinates $\phi(x, y) = (r(x, y), \theta(x, y))$ defined in $O =$ the complement of the negative x-axis by $r := \sqrt{x^2 + y^2}$ and $\theta :=$ the branch of $\arctan(\frac{y}{x})$ taking values in $-\pi$ to π. In this case U is the infinite rectangle $(0, \infty) \times (-\pi, \pi)$ and the inverse diffeomorphism is given by $\psi(r, \theta) = (r\cos\theta, r\sin\theta)$. Similarly, in $\mathbf{R}^3$ we can use polar cylindrical coordinates r, θ, z to deal with problems that are symmetric under rotations about the z-axis or polar spherical coordinates r, θ, φ for problems with symmetry under all rotations about the origin.

▷ **Exercise D–3.** If $f(x, y)$ is a real-valued function, then $f^\phi(r, \theta) = f \circ \phi^{-1}(r, \theta) = f(r\cos\theta, r\sin\theta)$, so if f is C^1, then, by the chain rule, $\frac{\partial f^\phi}{\partial r}(r, \theta) = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$, and similarly, $\frac{\partial f^\phi}{\partial \theta}(r, \theta) = -\frac{\partial f}{\partial x} r\sin\theta + \frac{\partial f}{\partial y} r\cos\theta = -y\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial y}$. Generalize this to find formulas for $\frac{\partial f^\phi}{\partial \phi_i}$ in terms of the $\frac{\partial f}{\partial x^i}$ for a general coordinate system ϕ.

▷ **Exercise D–4.** Show that a C^1 real-valued function in the plane, $f : \mathbf{R}^2 \to \mathbf{R}$, is invariant under rotation if and only if $y\frac{\partial f}{\partial x} = x\frac{\partial f}{\partial y}$. (Hint: The condition for f to be invariant under rotation is that $f^\phi(r, \theta)$ should be a function of r only, i.e., $\frac{\partial f^\phi}{\partial \theta} \equiv 0$.)

D.2. Some Canonical Forms

Now let us look at the standard canonical form theorems for functions, curves, and vector fields.

Recall that if f is a C^1 real-valued function defined in some open set G of $\mathbf{R}^n$, then a point $p \in G$ is called a critical point of f if $df_p = 0$ and otherwise it is called a regular point of f. If p is a regular point, then we can choose a basis $\ell_1, \ldots, \ell_n$ of $(\mathbf{R}^n)^*$ with $\ell_1 = df_p$, and by the remark at the beginning of this appendix it follows that $(f, \ell_2, \ldots, \ell_n)$ is a coordinate system in some neighborhood O of p. This proves the following canonical form theorem for smooth real-valued functions:

D.2.1. Proposition. *Let f be a smooth real-valued function defined in an open set G of $\mathbf{R}^n$ and let $p \in G$ be a regular point of f. Then there exists a coordinate system $(\phi_1, \ldots, \phi_n)$ defined in some neighborhood of p such that $f^\phi = \phi_1$.*

Informally speaking, we can say that near any regular point a smooth function looks linear in a suitable coordinate system.

Next we will see that a straight line is the canonical form for a smooth curve $\sigma : I \to \mathbf{R}^n$ at a regular point, i.e., a point $t_0 \in I$ such that $\sigma'(t_0) \neq 0$.

D.2.2. Proposition. *If t_0 is a regular point of the smooth curve $\sigma : I \to \mathbf{R}^n$, then there is a diffeomorphism ϕ of a neighborhood of $\sigma(t_0)$ into $\mathbf{R}^n$ such that $\sigma^\phi(t) = \gamma(t)$, where $\gamma : \mathbf{R} \to \mathbf{R}^n$ is the straight line $t \mapsto (t, 0, \ldots, 0)$.*

Proof. Without loss of generality we can assume that $t_0 = 0$. Also, since we can anyway translate $\sigma(t_0)$ to the origin and apply a linear isomorphism mapping $\sigma'(t_0)$ to $e_1 = (1, 0, \ldots, 0)$, we will assume that $\sigma(0) = 0$ and $\sigma'(0) = e_1$. Then if we define a map ψ near the origin of $\mathbf{R}^n$ by $\psi(x_1, \ldots, x_n) = \sigma(x_1) + (0, x_2, \ldots, x_n)$, it is clear that $D\psi_0$ is the identity, so by the inverse mapping theorem, ψ maps some neighborhood U of the origin diffeomorphically onto another neighborhood O. By definition, $\psi \circ \gamma(t) = \sigma(t)$, so if $\phi = \psi^{-1}$, then $\sigma^\phi(t) = \phi \circ \sigma(t) = \gamma(t)$. ∎

Notice a pattern in the canonical form theorems for functions and curves. If we keep away from "singularities", then locally a function or curve looks like the simplest example. This pattern is repeated for vector fields. Recall that a singularity of a vector field V is a point p where $V(p) = 0$, and the simplest vector fields are the constant vector fields, such as $\frac{\partial}{\partial x^1}$. The canonical form theorem for vector fields, often called the "Straightening Theorem", just says that near a nonsingular point a smooth vector field looks like a constant vector field. Let's try to make this more precise.

D.2.3. Definition. Let V be a vector field defined in an open set O of $\mathbf{R}^n$, and let $\phi = (\phi_1, \dots, \phi_n)$ be local coordinates in O. We call ϕ the *flow-box coordinates* for V in O if $V^\phi = \frac{\partial}{\partial \phi_1}$.

D.2.4. Remark. Let $\phi(x, y) = (r(x, y), \theta(x, y))$ denote polar coordinates in $\mathbf{R}^2$. We saw above that if $V = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$, then $V^\phi = \frac{\partial}{\partial \theta}$, so that polar coordinates are flow-box coordinates for V.

D.2.5. The Straightening Theorem. *If V is a smooth vector field defined in an open set O of $\mathbf{R}^n$ and $p_0 \in O$ is not a singularity of V, then there exist flow-box coordinates for V in some neighborhood of p_0.*

Proof. This is a very strong result; it easily implies both local existence and uniqueness of solutions and smooth dependence on initial conditions. And as we shall now see, these conversely quickly give the Straightening Theorem. Without loss of generality, we can assume that p_0 is the origin and $V(0) = e_1 = (1, 0, \dots, 0)$. Choose $\epsilon > 0$ so that for $\|p\| < \epsilon$ there is a unique solution curve of V, $t \mapsto \sigma(t, p)$, defined for $|t| < \epsilon$ and satisfying $\sigma(0, p) = p$. The existence of ϵ follows from the local existence and uniqueness theorem for solutions of ODE (Appendix B). Let U denote the disk of radius ϵ in $\mathbf{R}^n$ and define $\psi : U \to \mathbf{R}^n$ by $\psi(x) = \sigma(x_1, (0, x_2, \dots, x_n))$. It follows from smooth dependence on initial conditions (Appendix G) that ψ is a smooth map.

$\triangleright$ **Exercise D–5.** Complete the proof of the Straightening Theorem by first showing that $D\psi_0$ is the identity map (so ψ does define a local coordinate system near the origin) and secondly showing that $V^\phi = \frac{\partial}{\partial \phi_1}$, where $\phi := \psi^{-1}$. (Hint: Note that $\sigma(0, (0, x_2, \dots, x_n)) = (0, x_2, \dots, x_n)$, while $\frac{\partial}{\partial x_1}\sigma(x_1, (0, \dots, 0)) = e_1$. The fact that $D\psi_0$ is the identity is an easy consequence. Since $t \mapsto \sigma(t, p)$ is a solution curve of V, it follows that $V(\psi(x)) = \frac{\partial}{\partial x_1}\psi(x)$, and it follows that $D\psi_p$ maps $(\frac{\partial}{\partial x_1})_p$ to $V(\psi(p))$. Use this to deduce $V^\phi = \frac{\partial}{\partial \phi_1}$. ∎

D.2.6. Definition. Let $V : \mathbf{R}^n \to \mathbf{R}^n$ be a smooth vector field. If O is an open set in $\mathbf{R}^n$, then a smooth real-valued function $f : O \to \mathbf{R}$ is called a *local constant of the motion* for V if $Vf \equiv 0$ in O.

Notice that $x^2, x^3, \dots, x^n$ are constants of the motion for the vector field $\frac{\partial}{\partial x_1}$. Hence,

D.2.7. Corollary of the Straightening Theorem. *If V is a smooth vector field on $\mathbf{R}^n$ and p is any nonsingular point of V, then there exist $n-1$ functionally independent local constants of the motion for V defined in some neighborhood O of p.*

(Functionally independent means that they are part of a coordinate system.)

D.2.8. CAUTION. There are numerous places in the mathematics and physics literature where one can find a statement to the effect that every vector field on $\mathbf{R}^n$ has $n-1$ constants of the motion. What is presumably meant is something like the above corollary, but it is important to realize that such statements should **not** be taken literally—there are examples of vector fields with no global constants of the motion except constants. A local constant of the motion is very different from a global one. If V is a vector field in $\mathbf{R}^n$ and f is a local constant of the motion for V, defined in some open set O, then if σ is a solution of V, $f(\sigma(t))$ will be constant on any interval I such that $\sigma(I) \subseteq O$; however it will in general have different constant values on different such intervals.