Math 396. Submersions and transverse intersections

Fix $1 \leq p \leq \infty$, and let $f : X' \to X$ be a C^p mapping between C^p premanifolds (no corners or boundary!). In a separate handout there is proved an important local structure theorem: if $\mathrm{d}f(x') : \mathrm{T}_{x'}(X') \to \mathrm{T}_{f(x')}(X)$ has rank r that is independent of x' then for any $x' \in X'$ there exist small opens $U' \subseteq X'$ around x' and $U \subseteq X$ around $x = f(x')$ with $f(U') \subseteq U$ such that U and U' admit C^p coordinate systems $\phi' : U' \to \mathbf{R}^{n'}$ and $\phi : U \to \mathbf{R}^n$ with respect to which f looks like $(a_1, \dots, a_{n'}) \mapsto (a_1, \dots, a_r, 0, \dots, 0)$. That is,

$$\phi \circ f \circ \phi'^{-1} : \phi'(U') \to \phi(U) \subseteq \mathbf{R}^n$$

is the map $(a_1, \dots, a_{n'}) \mapsto (a_1, \dots, a_r, 0, \dots, 0)$. In class we explained two important consequences, the local coordinatized descriptions of submersions and immersions. The aim of this handout is to explain some further applications of submersions via this local structure theorem. In particular, we analyze transverse intersections of embedded C^p submanifolds and give a clean geometric proof of the theorem on Lagrange multipliers in its natural setting on premanifolds.

1. Preliminaries on topologies and tangent spaces

Before we study transversality for submanifolds of a manifold, generalizing the notion of transversality for subspaces of a vector space, we need to clarify the topological properties of submersions and the behavior of tangent spaces with respect to product decompositions. We begin with the latter.

Let X_1 and X_2 be C^p premanifolds with $1 \leq p \leq \infty$. Pick a point $\xi = (\xi_1, \xi_2) \in X_1 \times X_2$. We would like to make precise how the tangent space to $X_1 \times X_2$ at ξ is related to the tangent spaces of X_1 and X_2 at ξ_1 and ξ_2 respectively. Let us record a theorem proved in an earlier handout:

Theorem 1.1. *Let $\iota_j : X_j \to X_1 \times X_2$ be the C^p maps $\iota_1(x_1) = (x_1, \xi_2)$ and $\iota_2(x_2) = (\xi_1, x_2)$. The linear map*

$$\mathrm{T}_{\xi_1}(X_1) \oplus \mathrm{T}_{\xi_2}(X_2) \to \mathrm{T}_\xi(X_1 \times X_2). \tag{1.1}$$

defined by $(v_1, v_2) \mapsto \mathrm{d}\iota_1(\xi_1)(v_1) + \mathrm{d}\iota_2(\xi_2)(v_2)$ is an isomorphism, with inverse given by

$$v \mapsto (\mathrm{d}\pi_1(\xi)(v), \mathrm{d}\pi_2(\xi)(v)),$$

where $\pi_j : X_1 \times X_2 \to X_j$ is the C^p projection.

Our interest in tangent spaces to products is due to the fact that the submersion theorem tells us that C^p submersions locally (on the source) look like projections to the factor of a product of C^p premanifolds. In general, when studying the projection $\pi : P = X \times X' \to X$ to a factor of a product, at any point $\xi = (x, x')$ the tangent map $\mathrm{d}\pi(\xi) : \mathrm{T}_\xi(P) \to \mathrm{T}_x(X)$ is a surjection and its kernel is the subspace $\mathrm{T}_{x'}(X')$ via $\mathrm{d}\iota'(x')$, with $\iota' : X' \to P$ the embedding that is an C^p isomorphism onto the fiber $\pi^{-1}(x) = \{x\} \times X'$ as an embedded submanifold of P. That is, for such a projection π the kernel of its tangent map at ξ is the tangent space $\mathrm{T}_\xi(\pi^{-1}(x))$ along the fiber $\pi^{-1}(\pi(\xi))$. This latter description is intrinsic to the projection map $\pi : P \to X$ in the sense that it *makes no reference* to the product decomposition of P. Hence, we arrive at the second part of the following result for general submersions:

Theorem 1.2. *Let $\pi : P \to X$ be a C^p submersion between C^p premanifolds with $1 \leq p \leq \infty$.*

1. *The map π is an open map, and a subset S in the open set $\pi(P) \subseteq X$ is open in $\pi(P)$ (or equivalently, in X) if and only if $\pi^{-1}(S) \subseteq P$ is an open set. In particular, the topology on $\pi(P)$ is determined by the topology on P and the set-theoretic map π.*

(2) *For $x \in \pi(P)$ the fiber $\pi^{-1}(x)$ is a closed C^p subpremanifold in P, and for all $\xi \in \pi^{-1}(x)$ the kernel of the surjection $\mathrm{d}\pi(\xi) : \mathrm{T}_\xi(P) \twoheadrightarrow \mathrm{T}_x(X)$ is the tangent space $\mathrm{T}_\xi(\pi^{-1}(x))$ along the fiber through ξ.*

Proof. We first show that π is an open map. By the submersion theorem, for each $\xi \in P$ over $x \in X$ there exist opens $U_0 \subseteq X$ around x and $U \subseteq \pi^{-1}(U_0)$ around ξ such that there is a C^p isomorphism $\varphi : U \simeq U_0 \times U_0'$ with U_0' an open box in a Euclidean space and $\pi|_U \circ \varphi^{-1} : U_0 \times U_0' \to U_0$ equal to the standard projection. Such projections are visibly open maps, and hence $\pi|_U : U \to U_0$ is open. Since the inclusion $U_0 \to X$ is open, it follows that $\pi|_U : U \to X$ is an open map. Such opens U cover P, so it follows that $\pi : P \to X$ is open.

We next check that the subspace topology on $\pi(P) \subseteq X$ is given by the condition for openness of preimages in P. Since $\pi(P)$ is open in X, when it is given its standard C^p-structure induced by X we see that π uniquely factors through a C^p mapping $\overline{\pi} : P \to \pi(P)$ that is also clearly a submersion. We may therefore replace X with $\pi(P)$ to reduce to the case when π is surjective. In this case, we need to prove that a subset $S \subseteq X$ is open if and only if its preimage $\pi^{-1}(S) \subseteq P$ is open. One implication follows from continuity of π, and conversely if $\pi^{-1}(S)$ is open then $\pi(\pi^{-1}(S))$ is open (since π is an open map). But surjectivity of π implies $S = \pi(\pi^{-1}(S))$, so S is open as desired.

Now we turn to the second part of the theorem. To see that $\pi^{-1}(x)$ is closed in P, it suffices that x is closed. Rather generally, any point z in a locally Hausdorff space Z is closed. Indeed, if $\{Z_i\}$ is a covering of Z by Hausdorff open sets then $Z_i \cap \{z\}$ is either a point or the empty set and hence is closed in Z_i for every i. Thus, $\{z\}$ is closed in Z. With $\pi^{-1}(x)$ now known to be a closed subset of P, the problem of whether or not it is an embedded C^p subpremanifold in P (and the description of its tangent spaces inside tangent spaces to P as kernels of maps $\mathrm{d}\pi(\xi)$ for $\xi \in \pi^{-1}(x)$) is a problem local on $\pi^{-1}(x)$ within P. (See the beginning of the proof of Theorem 2.1 for a general discussion of the local nature of the problem of constructing C^p subpremanifold structures on a locally closed set in a C^p premanifold.) Working locally on P and X, the submersion theorem identifies the situation $\pi : P \to X$ with a product projection $\pi : U \times U' \to U$. However, this is exactly the situation that we have already considered prior to the statement of the theorem, in which case the identification of the kernel of the tangent map with the tangent space along the fiber was deduced from Theorem 1.1. ■

One very useful consequence of this theorem is that *surjective* C^p submersions satisfy a simple mapping property with respect to C^p maps:

Corollary 1.3. *Let $\pi : P \to X$ be a surjective C^p submersion, with $1 \le p \le \infty$. If $f : P \to Y$ is a C^p map to a C^p premanifold Y and f has constant value on each fiber of π, then the unique set-theoretic map $\overline{f} : X \to Y$ satisfying $\overline{f} \circ \pi = f$ is in fact a C^p mapping. In particular, it is continuous.*

Before we prove the corollary, we note that the key aspect of the corollary is to ensure that $\overline{f}$ is a C^p mapping. This is false without the submersion property. For example, consider the map $t \mapsto t^3$ that is a C^∞ surjection $\pi : P = \mathbf{R} \to \mathbf{R} = X$ but fails to be a submersion at the origin. Any C^p map $f : P \to Y$ (such as the identity on $P = \mathbf{R}$) trivially has constant value on each fiber of π (as the fibers $\pi^{-1}(t) = \{t^{1/3}\}$ are 1-point sets), and $\overline{f}(t) = f(t^{1/3})$, so if the corollary were to apply to π then the conclusion would be that $t \mapsto f(t^{1/3})$ is C^p whenever $t \mapsto f(t)$ is C^p. This already fails even for the identity map $f : \mathbf{R} \to \mathbf{R}$.

As a simple example to illustrate the usefulness of the corollary, let V be a finite-dimensional $\mathbf{R}$-vector space with dimension ≥ 2 and let $\pi : V^\vee - \{0\} \to \mathbf{P}(V)$ be the natural submersion (sending each nonzero linear functional to its kernel hyperplane). By the corollary, if $h : V^\vee - \{0\} \to Y$ is a C^p mapping to a C^p premanifold Y such that h is constant on lines through the origin, then the unique set-theoretic factorization $h = \overline{h} \circ \pi$ has $\overline{h} : \mathbf{P}(V) \to Z$ a C^p mapping. The key is that we do *not* need to muck around with explicit local coordinate systems on $\mathbf{P}(V)$ or descriptions of h in local coordinates to verify this; once we have used coordinate systems on $\mathbf{P}(V)$ to verify that the surjective map h is a C^p submersion, the rest comes for free. To summarize, the problem of factoring a C^p map through a surjective C^p submersion is entirely *set-theoretic*: the submersion theorem does all of the hard work of keeping track of the C^p aspects of the problem.

Now we prove the corollary.

Proof. By Theorem 1.2, a subset $S \subseteq X$ is open if and only if its preimage in P is open. Thus, to see that $\overline{f} : X \to Y$ is continuous we have to show that for any open set $U \subseteq Y$ the preimage $\pi^{-1}(\overline{f}^{-1}(U))$ is open in P. But this is $f^{-1}(U)$ since $f = \overline{f} \circ \pi$, so continuity of f gives what we need. This takes care of the topological aspect of the problem, and to verify that f is a C^p mapping we may work locally near a point $x \in X$. Pick $\xi \in P$ over x, so by the submersion theorem we can find opens $U_0 \subseteq X$ around x and $U \subseteq \pi^{-1}(U_0)$ around ξ such that there is a C^p isomorphism $\varphi : U \simeq U_0 \times U_0'$ with $\pi|_U \circ \varphi^{-1} : U_0 \times U_0' \to U_0$ the standard projection.

Clearly $\overline{f}|_{U_0}$ arises as the factorization of $f|_U \circ \varphi^{-1}$ through $\pi|_U \circ \varphi^{-1}$. To verify that $\overline{f}|_{U_0}$ is a C^p map (which certainly suffices to get the C^p property for $\overline{f}$ near our arbitrary choice of point $x \in X$ around which U_0 is a small open), we may therefore replace $\pi : P \to X$ with the projection $\pi_0 : U_0 \times U_0' \to U_0$. Let $(x, x') \in U_0 \times U_0'$ correspond to ξ, so the "section" $s : U_0 \to U_0 \times U_0'$ defined by $s(u) = (u, x')$ is a C^p map (this is one of the "slice inclusions" considered in our general study of products above). Since $\pi_0 \circ s$ is the identity on U_0, we have

$$\overline{f} = \overline{f} \circ (\pi_0 \circ s) = (\overline{f} \circ \pi_0) \circ s = f \circ s$$

with $f : U_0 \times U_0' \to Y$ the given C^p mapping. Hence, we have presented $\overline{f}$ as the composite of the C^p maps s and f, so $\overline{f}$ is a C^p map. ■

2. The Jacobi criterion

We now turn to a very important application of the local structure theorem for C^p maps, essentially a criterion for fibers of C^p mappings to be C^p submanifolds. Here is the general problem. Let $f : X' \to X$ be a C^p mapping between C^p premanifolds ($1 \le p \le \infty$), and pick a point $x \in f(X')$, so the fiber $f^{-1}(x)$ is a non-empty closed subset of X'. Is it a C^p subpremanifold? In special situations, fibers of a map can become "degenerate". For example, consider the map $f : \mathbf{R}^2 \to \mathbf{R}$ defined by $f(u, v) = u^2 - v^2$. For $t \in \mathbf{R} - \{0\}$, the fiber $f^{-1}(t)$ is the hyperbola $H_t : u^2 - v^2 = t$ in $\mathbf{R}^2$, and this is certainly a smooth submanifold (as the differential of the defining equation of H_t is non-vanishing along H_t, and hence the implicit function theorem kicks in as usual). However, the fiber $f^{-1}(0)$ is the union of the coordinate axes and this has a problem at the point where the lines cross. In this case, for points $x' \in H_0$ we have $\mathrm{d}h(x') \neq 0$ except at $x' = (0, 0)$. Geometrically, the fibers H_t degenerate to something singular as $t \to 0$.

The one tool we have had so far to make embedded submanifolds is the case of vanishing for a single function: if the differential of C^p function h is nonvanishing at the points in the set $\{h = 0\}$ then $\{h = 0\}$ a (codimension-1) C^p submanifold. We seek a version that works for simultaneous zero loci of several C^p functions, but we first emphasize that we can only give sufficient conditions and not necessary ones. Indeed, even in the case of a single equation $\{h = 0\}$, the condition of non-vanishing on the differential along the zero locus is merely sufficient and *not* necessary for the zero locus to be a submanifold, since we could be using the "wrong" equations: the equation $h^2 = 0$

defines the set $\{h = 0\}$ but it never satisfies the differential criterion to be a submanifold – eve when $\{h = 0\}$ is a C^p submanifold – because $\partial_{x_j}(h^2) = h \cdot \partial_{x_j} h$ vanishes where h^2 vanishes.)

Just as the sufficient condition for $\{h = 0\}$ to be a submanifold even gave that it is "codimension 1", we expect a codimension aspect in the conclusion for any reasonable criterion for smoothness of simultaneous zero loci. Thus, let us first give a precise definition of codimension in the setting of embedded premanifolds. For a C^p subpremanifold Z in a C^p premanifold Z', the *codimension* of Z in Z' at a point $z \in Z$ is the codimension of $\mathrm{T}_z(Z)$ in $\mathrm{T}_z(Z')$; in other words, it is the difference of the pointwise dimensions of Z and Z' at z. If this number is a constant c independent of z, we say Z has *codimension* c in Z'. For example, if Z and Z' are connected then there is necessarily such a constant codimension (as the dimensions of tangent spaces on Z and Z' are the same at all points of Z and Z' respectively, and so the codimension at each $z \in Z$ is just the difference of these constant tangent space dimensions).

Theorem 2.1. *With notation as given above, suppose that for all $x' \in X'$ the tangent map* $\mathrm{df}(x') : \mathrm{T}_{x'}(X') \to \mathrm{T}_x(X)$ *has constant rank $r \geq 0$, where $x = f(x')$. For each $x \in X$, the fiber $f^{-1}(x) \subseteq X'$ is a closed C^p subpremanifold of X' with codimension r in X' at all points.*

This theorem is the *Jacobi criterion* for smoothness of fibers of a map. (The same name is also used for a variant in Theorem 3.3.) The failure of the submersion condition usually means that some fiber is "bad" (such as in our hyperbola example, where the submersion property fails at exactly the point $(0, 0)$ that is "bad" in its fiber H_0), but the situation can be more complicated. For example, if $C \subseteq \mathbf{R}^2$ is the parabola $y = x^2$ then the projection $f : C \to \mathbf{R}$ to the x-axis fails to be a submersion at $(0, 0)$ but the fiber is perfectly nice: a single point. However, the geometry shows that something has still "gone wrong" here: the fiber $f^{-1}(t)$ for $t > 0$ consists of two points $\{(t, \pm\sqrt{t})\}$ and as $t \to 0^+$ these "come together" to the single point $(0, 0)$ that morally ought to count twice; even worse, for $t < 0$ the fibers are empty. Thus, the fiber situation at $t = 0$ is geometrically a bit more complicated than one would guess from considering $f^{-1}(0)$ in isolation without noticing the nature of $f^{-1}(t)$ for $t < 0$ and $t > 0$ near 0.

Proof. We fix $x \in f(X')$ and aim to prove that the non-empty closed subset $f^{-1}(x)$ in X' is a C^p subpremanifold with codimension r at all points. Our first task is to show that the problem is local on X' around points of $f^{-1}(x)$, as this will open the door to using the local structure theorem for C^p maps with constant Jacobian rank. We will exploit a powerful general principle, which is that to solve global construction problems admitting unique solutions it suffices to make solutions locally, as the uniqueness will ensure that local solutions agree on overlaps and hence "glue" to give a global solution! In our present circumstances, the *a priori* uniqueness result is the fact that locally closed sets in X' admit at most one structure of C^p subpremanifold of X'. Thus, if $f^{-1}(x)$ is to admit a structure of C^p subpremanifold, such a structure is unique (and we can try to work locally to check its codimension is r at all points). Let us first explain how this uniqueness lets us reduce the global problem to the local case.

Let $\{U'_i\}$ be a collection of opens in X' that cover $f^{-1}(x)$, and let $f_i : U'_i \to X$ be the restriction of f. Clearly $f_i^{-1}(x') = U'_i \cap f^{-1}(x)$, so this is a closed set in U'_i and hence is a locally closed set in X'. Suppose we can solve the problem for each f_i, so we get a C^p-structure on each $f_i^{-1}(x') = U'_i \cap f^{-1}(x)$ making it an embedded C^p subpremanifold of U'_i and hence of X' (as U'_i is open in X'). These C^p-structures on the *open* subsets $U'_i \cap f^{-1}(x)$ in the topological space $f^{-1}(x)$ *must* agree on overlaps. (In the more classical language of atlases, this just says that C^p atlases arising on the overlap from the respective maximal C^p-atlases on $f_i^{-1}(x)$ and $f_j^{-1}(x)$ must be C^p-compatible with each other.) Indeed, on $(U'_i \cap f^{-1}(x)) \cap (U'_j \cap f^{-1}(x))$ we get two induced C^p-structures, one from its nature as an open set in $U'_i \cap f^{-1}(x)$ and one from its nature as an open set in $U'_j \cap f^{-1}(x)$, and so both structures make it a locally closed C^p subpremanifold in X'. But we have seen in class that locally closed sets in a C^p premanifold (such as X') have *at most one* structure of C^p subpremanifold, and thus we get the asserted agreement on overlaps. The C^p-structures on the $f_i^{-1}(x)$'s therefore arise from a unique C^p-structure on $f^{-1}(x)$ (proof: the C^p-atlases on the open subsets $f_i^{-1}(x) \subseteq f^{-1}(x)$ are C^p-compatible on all overlaps $f_i^{-1}(x) \cap f_j^{-1}(x)$, and hence their union is a C^p-atlas on $f^{-1}(x)$; the associated C^p-structure is the only possibility, and clearly it does work).

Continuing to assume that we can solve the problem for the maps $f_i : U'_i \to X$, we claim that the resulting C^p-structure that we have just constructed on $f^{-1}(x)$ makes the inclusion map $f^{-1}(x) \to X'$ an immersion (and hence a C^p embedding). Indeed, $f^{-1}(x)$ is covered by the open subsets $f_i^{-1}(x)$ on which the induced C^p-structure is the one we assumed to exist as a local solution to our original problem, and so since the inclusion maps $f_i^{-1}(x) \to U'_i$ are C^p immersions (by assumption!) we deduce the same for the inclusion maps $f_i^{-1}(x) \to X'$ and hence for the inclusion map $f^{-1}(x) \to X'$. Likewise, the property that $f^{-1}(x)$ with its C^p structure has codimension r in X' may be inferred from the same codimension property for the inclusion maps $f_i^{-1}(x) \to X'$. To summarize, we have shown that to solve the original global problem for $f : X' \to X$ it suffices to solve it for the maps $f_i : U'_i \to X$ for opens $\{U'_i\}$ in X' whose union contains $f^{-1}(x)$. In other words, for each $x' \in f^{-1}(x)$ it suffices to find an open $U' \subseteq X'$ around x' such that we can solve the problem for the map $f|_{U'} : U' \to X$. Also, if $U \subseteq X$ is an open set around x then $f^{-1}(U) \subseteq X'$ is an open set around $f^{-1}(x)$ and so we can replace X' and X with $f^{-1}(U)$ and U respectively. Hence, we can work locally in X' around a fixed point $x' \in f^{-1}(x)$ and locally in X around x. That is, for opens $U \subseteq X$ around x and $U' \subseteq f^{-1}(U)$ around $x' \in f^{-1}(x)$ it suffices to solve the problem for the map $f : U' \to U$ *as we vary* $x' \in f^{-1}(x)$.

Now we bring in the local structure theorem: since it is assumed that f has constant Jacobian rank r, we can find small opens $U \subseteq X$ around x and $U' \subseteq X'$ around x' admitting respective C^p coordinate systems $\phi : U \to \mathbf{R}^n$ and $\phi' : U' \to \mathbf{R}^{n'}$ satisfying $\phi(x) = 0 \in \mathbf{R}^n$ and $\phi'(x') = 0 \in \mathbf{R}^{n'}$ such that in terms of the associated C^p coordinate systems we have the following "coordinatized" description of the map $f_{U',U} : U' \to U$ induced by f:

$$\widetilde{f} = \phi \circ f_{U',U} \circ \phi'^{-1} : (a'_1, \dots, a'_{n'}) = (a'_1, \dots, a'_r, 0, \dots, 0) \in \phi(U) \subseteq \mathbf{R}^n$$

on $\phi'(U')$. Thus, as a subset of the open domain $\phi'(U') \subseteq \mathbf{R}^{n'}$, we have the following coordinatized description of the part of the fiber $f^{-1}(x)$ that meets U' (i.e., $f_{U',U}^{-1}(x)$):

$$\phi'(U' \cap f^{-1}(x)) = \widetilde{f}^{-1}(0) = \phi'(U') \cap \{x'_1 = \cdots = x'_r = 0\}$$

inside of $\phi'(U')$. This is an open subset of the codimension-r linear subspace

$$W = \mathbf{R}^{n'-r} = \{x'_1 = \cdots = x'_r = 0\}$$

in $\mathbf{R}^{n'}$, and the standard linear coordinates $x'_{r+1}|_W, \dots, x'_{n'}|_W$ on this subspace provide the solution to our problem. More precisely, by taking $x'_j \circ \phi'|_{U' \cap f^{-1}(x)} = (x'_j|_W) \circ \phi'$ for $r+1 \leq j \leq n'$ as coordinates on $U' \cap f^{-1}(x)$, we get a C^p atlas on $f^{-1}(x) \cap U'$ with respect to which its inclusion map into U' is a codimension-r C^p subpremanifold (because $\{x'_1 \circ \phi', \dots, x'_{n'} \circ \phi'\}$ is a C^p coordinate system on U' with respect to which $U' \cap f^{-1}(x)$ is the zero locus of the first r coordinates). This gives a solution to our problem on U', including the property of having codimension r. ■

3. Transversal submanifolds

To apply the theorems in the previous section, we turn to an important topic in the geometry of submanifolds that generalizes the notion of transversality for linear subspaces of a finite-dimensional vector space. As a special case, we will deduce the theorem on Lagrange multipliers in its natural geometric setting (recovering the "traditional" version on open sets in $\mathbf{R}^n$ as a special case). The starting point is:

Definition 3.1. Let X be a C^p premanifold, $1 \le p \le \infty$. A collection $Z_1, \dots, Z_n$ of embedded C^p subpremanifolds in X is *mutually transverse* if for each $z \in Z = \cap_j Z_j$ the subspaces $\mathrm{T}_z(Z_j)$ in $\mathrm{T}_z(X)$ are mutually transverse.

Note that this definition is vacuously satisfied (and not interesting) if $\cap Z_j$ is empty.

Example 3.2. Let $f_1, \dots, f_n : X \to \mathbf{R}$ be C^p mappings such that $\{\mathrm{d}f_j(x)\}$ is an independent set in $\mathrm{T}_x(X)^\vee$ for all $x \in X$. (In particular, $\mathrm{d}f_j(x) \neq 0$ for all $x \in X$.) By the differential criterion for zero loci to be subpremanifolds, the non-vanishing of $\mathrm{d}f_j(x)$ for all $x \in X$ implies that for any $c_1, \dots, c_n \in \mathbf{R}$ the level sets $Z_j = f_j^{-1}(c_j) = \{f_j - c_j = 0\}$ are closed C^p subpremanifolds of X with codimension 1 at all points when $Z_j \neq \emptyset$. We claim that the Z_j's are mutually transverse in X.

By Theorem 1.2, for any $z \in Z_j$ the subspace $\mathrm{T}_z(Z_j)$ in $\mathrm{T}_z(X)$ is the hyperplane kernel $\ker \mathrm{d}f_j(z)$ of the nonzero linear functional $\mathrm{d}f_j(z) \in \mathrm{T}_z(Z)^\vee$. Thus, the condition on mutual transversality of the Z_j's at $z \in \cap Z_j$ reduces to the claim that if V is a vector space (such as $\mathrm{T}_z(X)$) and $\ell_1, \dots, \ell_n \in V^\vee$ are linearly independent functionals on V (such as the $\mathrm{d}f_j(z)$'s) then the hyperplanes $\ker \ell_j$ are mutually transverse in V. This fact from linear algebra was proved on an earlier homework.

The most classical instance of this example is to take X to be an open set U in some $\mathbf{R}^N$ (with its usual C^p-structure). In this case, the result is that if $f_1, \dots, f_n$ are C^p functions on U with $n \le N$ and $c_1, \dots, c_n \in \mathbf{R}$ are numbers such that the Jacobian matrix $(\partial f_i / \partial x_j)$ has rank n (i.e., independent columns) at all points on $\cap f_j^{-1}(c_j)$, then if this set is non-empty it is a closed C^p submanifold of U with codimension n in U (i.e., dimension $N - n$) at all of its points. Beware that just checking this "Jacobi condition" at points on the mutual overlap of the level sets $Z_j = f_j^{-1}(c_j)$ does not tell us anything about the global geometry of overlaps of fewer than n of the Z_j's. Of course, since linear independence of a collection of n vectors in $\mathbf{R}^N$ is an open condition on $(\mathbf{R}^N)^n$, it does follow automatically from the continuity of the $\partial f_i / \partial x_j$'s that for some open U' around $\cap Z_j$ in U each $Z'_j = Z_j \cap U'$ is a closed C^p submanifold in U' and that these are mutually transverse. However, "far away" there could of course be nasty singularities on the Z_j's if we do not examine the $\mathrm{d}f_j$'s away from $Z_1 \cap \cdots \cap Z_n$.

The preceding example can be pushed a little further:

Theorem 3.3. *Let $f_1, \dots, f_n \in \mathscr{O}(X)$ be global C^p functions on X such that the functionals $\mathrm{d}f_j(x)$ in $\mathrm{T}_x(X)^\vee$ are linearly independent for all $x \in X$. Let $Z_j = f_j^{-1}(c_j)$ for $c_1, \dots, c_n \in \mathbf{R}$, and assume $Z = \cap Z_j$ is non-empty, so the Z_j's are mutually transverse C^p hypersurfaces in X. The closed set Z a closed C^p subpremanifold with codimension n at all points and $\mathrm{T}_z(Z) = \cap \ker \mathrm{d}f_j(z)$ for all $z \in Z$.*

Proof. Let $f = (f_1, \dots, f_n) : X \to \mathbf{R}^n$, so $Z = f^{-1}(c_1, \dots, c_n)$. For any $x \in X$, the standard linear isomorphism $\mathrm{T}_{f(x)}(\mathbf{R}^n) \simeq \mathbf{R}^n$ identifies $\mathrm{d}f(x)$ with the linear map $\mathrm{T}_x(X) \to \mathbf{R}^n$ whose component functions are the functionals $\mathrm{d}f_j(x)$, and so $\ker \mathrm{d}f(x)$ is the intersection of the kernels of the $\mathrm{d}f_j(x)$'s. Thus, the theorem says that for $c = (c_1, \dots, c_n) \in \mathbf{R}^n$ with $f^{-1}(c)$ non-empty, the fiber $f^{-1}(c)$ is a closed C^p subpremanifold of X with codimension n at all points and for all

$z \in f^{-1}(c)$ the tangent space $\mathrm{T}_z(f^{-1}(c))$ to the fiber is $\ker \mathrm{d}f(z)$. By Theorem 1.2, it suffices to show that f is a submersion. That is, for every $x \in X$ we want $\mathrm{d}f(x) : \mathrm{T}_x(X) \to \mathrm{T}_{f(x)}(\mathbf{R}^n)$ to be surjective.

We have noted above that under the standard linear isomorphism $\mathrm{T}_{f(x)}(\mathbf{R}^n) \simeq \mathbf{R}^n$, $\mathrm{d}f(x)$ has component functionals $\mathrm{d}f_j(x)$. Thus, our problem is a special case of a more general claim in linear algebra: if V is a finite-dimensional vector space over a field F (such as $\mathrm{T}_x(X)$ over $\mathbf{R}$) and $\ell_1, \dots, \ell_n \in V^\vee$ are linearly independent functionals (such as the $\mathrm{d}f_j(x)$'s) then the map $V \to F^n$ with ℓ_j's as component functions is surjective. Dualizing, it is equivalent to say that the map $F^n \to V^\vee$ sending the standard (dual!) basis vectors to the ℓ_j's is injective, and this is exactly the hypothesis of linear independence on the ℓ_j's. ■

A very interesting corollary is the theorem on Lagrange multipliers:

Corollary 3.4. *Let $f_1, \dots, f_n : X \to \mathbf{R}$ be C^p functions on a C^p premanifold. Fix $c_1, \dots, c_n \in \mathbf{R}$ and let $Z = \cap f_j^{-1}(c_j)$. Assume Z is non-empty, so it is a closed C^p subpremanifold of X with codimension n at all points, by Theorem 3.3. Assume that $\{\mathrm{d}f_i(x)\}$ is linearly independent in $\mathrm{T}_x(X)^\vee$ for all $x \in Z$.*

For a C^p function $g : X \to \mathbf{R}$, if $g|_Z$ has a local extremum at $z \in Z$ then $\mathrm{d}g(z) = \sum \lambda_j \mathrm{d}f_j(z)$ in $\mathrm{T}_z(X)^\vee$ for some $\lambda_1, \dots, \lambda_n$.

Taking X to be an open set in $\mathbf{R}^N$, this corollary is exactly the theorem on Lagrange multipliers in multivariable calculus. There is no doubt that the theory of transverse submanifolds of a manifold (as used in the following proof) provides a powerful geometric intuition for the real meaning of this theorem. It is precisely the possibility of viewing the "constraint locus" $Z = \cap f_j^{-1}(c_j)$ as a manifold in its own right (compatibly with the manifold structure on X as well) that is the key to the whole proof.

Proof. Since the $\mathrm{d}f_j(z)$'s are linearly independent functionals, the intersection $\mathrm{T}_z(Z)$ of their kernels is annihilated by exactly those functionals on $\mathrm{T}_z(X)$ that are linear combinations of the $\mathrm{d}f_j(z)$'s in $\mathrm{T}_z(X)^\vee$. (This is a general fact in linear algebra that is easily proved by extending the independent functionals to a basis of the dual space, and then computing in a dual basis of the original space.) Thus, the meaning of the conclusion in the theorem is simply that $\mathrm{d}g(z) \in \mathrm{T}_z(X)^\vee$ vanishes on $\mathrm{T}_z(Z)$.

Let $i : Z \to X$ be the C^p inclusion, so $\mathrm{d}i(z) : \mathrm{T}_z(Z) \to \mathrm{T}_z(X)$ is the linear injection by which we are identifying $\mathrm{T}_z(Z)$ with a linear subspace of $\mathrm{T}_z(X)$. The restriction of $\mathrm{d}g(z)$ to $\mathrm{T}_z(Z)$ is thereby identified with the composite $\mathrm{d}(z) \circ \mathrm{d}i(z)$ (we have $i(z) = z$), and by the Chain Rule this is $\mathrm{d}(g \circ i)(z)$. Hence, our goal is to prove $\mathrm{d}(g \circ i)(z) = 0$ as a functional on $\mathrm{T}_z(Z)$. But $g \circ i$ is the C^p function on Z obtained by restricting g to Z, and so our problem is now intrinsic to the C^p premanifold Z (and the ambient X and constraint functions f_j may now be discarded from consideration): we have a C^p function $h : Z \to \mathbf{R}$ with a local extremum at a point $z \in Z$, and we want to prove that the linear functional $\mathrm{d}h(z) : \mathrm{T}_z(Z) \to \mathbf{R}$ vanishes.

For local C^p coordinates $\{x_1, \dots, x_m\}$ on Z near z, the tangent vectors $\partial_{x_j}|_z$ are a basis of $\mathrm{T}_z(Z)$, and so it is equivalent to say that $\mathrm{d}h(z)(\partial_{x_j}|_z) = 0$ for all j. But we know that this value is $(\partial h/\partial x_j)(z)$, and so the problem becomes one in multivariable calculus: if a C^p function on an open set in $\mathbf{R}^m$ has a local extremum at some point, its partials in every coordinate direction vanish at the point. But this is obvious: taking a partial derivative (of a C^p function) in a coordinate direction at a point in $\mathbf{R}^m$ is just a 1-variable derivative for the restriction of the function to the line through the point parallel to the coordinate axis, and so the statement is that a C^p function on an *open* interval in $\mathbf{R}$ has vanishing derivative at any local extremum. This is the first interesting theorem about derivatives that one learns in calculus. ■

So far we have examined transversality for hypersurfaces in some detail, but we now want to show that the general situation for transversality also accords with our geometric intuition. The interesting feature of transversality in general (not just in the case of hypersurfaces) is that it ensures intersections are "nice" and have the "expected" codimension when non-empty:

Theorem 3.5. *Let $Z_1, \dots, Z_n$ be mutually transverse embedded C^p subpremanifolds in a C^p premanifold X. If the intersection $\cap Z_j$ is non-empty then it is an embedded C^p subpremanifold of X with tangent space $\cap \mathrm{T}_z(Z_j)$ inside of $\mathrm{T}_z(X)$ for all $z \in \cap Z_j$. In particular, if Z_j has constant codimension c_j in X for each j then $\cap Z_j$ has constant codimension $\sum c_j$ at all points when it is non-empty.*

As the proof of the theorem will show, the local situation near points in $\cap Z_j$ is rather similar to that in Theorem 3.3. Also, the transversality hypothesis is crucial for relations such as $\mathrm{T}_z(\cap Z_j) = \cap T_z(Z_j)$ and $\mathrm{codim}_X(\cap Z_j) = \sum \mathrm{codim}_X(Z_j)$ to hold, even if the subset $\cap Z_j$ happens to be a C^p subpremanifold of X for "accidental" reasons. For example, consider $X = \mathbf{R}^3$, $Z_1 = \{x^2 + y^2 = z\}$, and $Z_2 = \{x^2 + y^2 = -z\}$. These two paraboloid C^∞ surfaces in X are non-transverse at the unique point $(0, 0, 0)$ in the intersection $Z_1 \cap Z_2$ whose dimension is "lower than expected" (a generic intersection of surfaces in 3-space "ought" to be a curve): the tangent planes $\mathrm{T}_0(Z_1)$ and $\mathrm{T}_0(Z_2)$ inside of $\mathrm{T}_0(X) \simeq \mathbf{R}^3$ coincide with the xy-plane $\mathbf{R}\partial_x|_0 + \mathbf{R}\partial_y|_0$. This failure of transversality of tangent planes at the origin is "why" this intersection fails to exhibit the pleasant geometric features as in the transversal case addressed in the theorem.

Proof. The calculation of the codimension follows from the rest because of an earlier homework result in linear algebra: intersections of mutually transverse linear subspaces of a finite-dimensional vector spaces necessarily have the "expected" codimension. Recall also from the homework that the property of mutual transversality for a collection of linear subspaces $W_1, \dots, W_n$ in a finite-dimensional vector space V is equivalent to pairwise transversality for $\cap_{j \le i} W_j$ and W_{i+1} for $1 \le i < n$. Thus, in view of the asserted description of the tangent spaces on $\cap Z_j$ (if non-empty), we may proceed by induction on n, and more specifically we immediately reduce to the special case $n = 2$: if Z and Z' are transverse embedded C^p subpremanifolds in X, and if the locally closed set $Z \cap Z'$ in X is non-empty, then it is an embedded C^p subpremanifold with tangent space $\mathrm{T}_x(Z) \cap \mathrm{T}_x(Z')$ inside of $\mathrm{T}_x(X)$ for each $x \in Z \cap Z'$.

In view of the uniqueness of C^p subpremanifold structures on locally closed subsets of X, the problem of constructing the asserted C^p-structure on $Z \cap Z'$ is local on $Z \cap Z'$ in X. (This goes by the same localization method as in the proof of Theorem 2.1.) Hence, for each $x \in Z \cap Z'$ it suffices to solve the problem after replacing X with an open set U around x. In particular, since the pointwise dimension on a premanifold is locally constant and both Z and Z' have the subspace topology from X, we may shrink around x so that X, Z, and Z' have constant dimensions. Thus, Z and Z' have constant codimensions in X, say c and c'. In view of the immersion theorem (that gives a local description of embedded C^p subpremanifolds in a C^p premanifold), we may therefore shrink X so that there exist C^p functions $f_1, \dots, f_c$ and $f'_1, \dots, f'_{c'}$ on X such that $\{f_1, \dots, f_c\}$ and $\{f'_1, \dots, f'_{c'}\}$ are each subsets of C^p coordinate systems (not necessarily the same ones, *a priori*) with $Z = \cap f_j^{-1}(0)$ and $Z' = \cap f_j'^{-1}(0)$.

Since these f_i's and f'_j's are subsets of C^p coordinate systems on X, it follows that the subsets $\{\mathrm{d}f_i(x)\}_i$ and $\{\mathrm{d}f'_j(x)\}$ in $\mathrm{T}_x(X)^\vee$ are linearly independent sets. Pick $x \in Z \cap Z'$ and let $H_i = \ker \mathrm{d}f_i(x)$ and $H'_j = \ker \mathrm{d}f'_j(x)$ be the kernel hyperplanes in $\mathrm{T}_x(X)$. The collection of

H_i's is mutually transverse in $\mathrm{T}_x(X)$ (as the f_i's are part of a C^p coordinate system), and likewise the collection of H'_j's is mutually transverse in $\mathrm{T}_x(X)$. Moreover, by Theorem 3.3, we have $\cap \ker \mathrm{d}f_j(z) = \mathrm{T}_z(Z)$ for each $z \in Z$, and likewise $\cap \ker \mathrm{d}f'_i(z') = \mathrm{T}_{z'}(Z')$ for each $z' \in Z'$. Thus, for our $x \in Z \cap Z'$ we have $\cap H_i = \mathrm{T}_x(Z)$ and $\cap H'_j = \mathrm{T}_x(Z')$. These respective subspaces of $\mathrm{T}_x(X)$ have codimensions c and c', and the transversality hypothesis on Z and Z' at x says that $\mathrm{T}_x(Z) \cap \mathrm{T}_x(Z')$ has codimension $c + c'$ in $\mathrm{T}_x(X)$. Thus, the $c + c'$ hyperplanes $H_1, \dots, H_c, H'_1, \dots, H'_{c'}$ in $\mathrm{T}_x(X)$ have intersection with codimension $c + c'$, whence the *combined* collection of hyperplanes is mutually transverse in $\mathrm{T}_x(X)$. By the homework problem on transversality of linear subspaces of a finite-dimensional vector space, this implies that the combined collection of functionals given by the $\mathrm{d}f_i(x)$'s and $\mathrm{d}f'_j(x)$'s span lines whose direct sum injects into the dual space $\mathrm{T}_x(X)^\vee$ and hence these functionals are linearly independent.

Such a linear independence property at the point x says exactly that the map

$$f : (f_1, \dots, f_c, f'_1, \dots, f'_{c'}) : X \to \mathbf{R}^{c+c'}$$

is a submersion at x (as the dual map on cotangent spaces $\mathbf{R}^{c+c'} = \mathrm{T}_{f(x)}(\mathbf{R}^{c+c'})^\vee \to \mathrm{T}_x(X)^\vee$ is injective, since it sends the collection of dual basis vectors $\mathrm{d}t_k(f(x))$ to the linearly independent collection of cotangent vectors given by the $\mathrm{d}f_i(x)$'s and $\mathrm{d}f'_j(x)$'s at x). By the submersion theorem, this submersion property at our point $x \in Z \cap Z'$ extends to an open set in X around x, so by shrinking X around x we may arrange that for *all* points $\xi \in X$ the collection of functionals given by the $\mathrm{d}f_i(\xi)$'s and $\mathrm{d}f'_j(\xi)$'s in $\mathrm{T}_\xi(X)^\vee$ are independent, or equivalently that f is a submersion on X. Hence, Theorem 1.2 ensures that the fiber $f^{-1}(0)$ is a C^p subpremanifold of X and its tangent space at any point ξ is the kernel of $\mathrm{d}f(\xi) : \mathrm{T}_\xi(X) \to \mathrm{T}_0(\mathbf{R}^{c+c'})$. Identifying $\mathrm{T}_0(\mathbf{R}^{c+c'})$ with $\mathbf{R}^{c+c'}$ in the usual manner gives the $\mathrm{d}f_i(\xi)$'s and $\mathrm{d}f'_j(\xi)$'s as the component functions of $\mathrm{d}f(\xi)$, and so the kernel $\mathrm{T}_\xi(f^{-1}(0))$ of $\mathrm{d}f(\xi)$ is the intersection of the kernels of the $\mathrm{d}f_i(\xi)$'s and $\mathrm{d}f'_j(\xi)$'s. But we rigged the f_i's and f'_j's so that $f^{-1}(0) = Z \cap Z'$, and hence for our point $x \in Z \cap Z'$ we have

$$\mathrm{T}_x(Z \cap Z') = (\cap H_i) \cap (\cap H'_j) = \mathrm{T}_x(Z) \cap \mathrm{T}_x(Z')$$

inside of $\mathrm{T}_x(X)$. ■

Remark 3.6. Even if X and the Z_j's in the theorem are connected, the overlap $\cap Z_j$ may have infinitely many connected components. For example, consider $X = \mathbf{R}^2$ and Z and Z' the graphs $y = \cos(x)$ and $y = \sin(x)$ respectively. (Check using the rigorous definition and not just drawing of pictures these really are transverse at their points of intersection.) In more complicated geometric settings, finding the connected components of $\cap Z_j$ can be extremely hard!

We conclude by recording an important observation that is implicit in the preceding proof: if $\{Z_1, \dots, Z_n\}$ is a collection of mutually transverse embedded C^p subpremanifolds of a C^p premanifold X ($1 \le p \le \infty$), then around each $z \in \cap Z_j$ there exists a local C^p coordinate system in terms of which the collection of Z_j's is carried over to a collection of mutually transverse linear subspaces in some $\mathbf{R}^N$, at least in a small open neighborhood of the origin. Thus, in a sense, the local theory of mutually transverse subpremanifolds is a non-linear version of the theory of mutually transverse linear subspaces of a finite-dimensional vector space. In particular, from Theorem 3.5 and the linear theory in the homework we get:

Corollary 3.7. *Let $Z_1, \dots, Z_n$ be a collection of embedded C^p subpremanifolds of a C^p premanifold X. If $\cap_{j \le i} Z_j$ and Z_{i+1} are transverse in X for $1 \le i < n$, then the Z_j's are mutually transverse in X.*

Remark 3.8. Theorem 3.5 ensures (by induction on i) that the intersections $\cap_{j\le i} Z_i$ are embedded C^p subpremanifolds of X for $2 \le i \le n$ due to the assumption at stage $i-1$. Hence, the hypothesis in the corollary "makes sense" when considered as an ordered set of hypotheses for $i = 1, \dots, n-1$.

4. Applications to Lie groups

A *Lie group* is a C^∞ manifold G equipped with a group structure such that the multiplication map $m : G \times G \to G$ and the inversion map $i : G \to G$ in the group law are C^∞ maps. (It is a remarkable theorem in Lie theory that weakening C^∞ to C^0 does not give a new concept: any topological group manifold admits a *unique* C^∞-structure with respect to which it is a C^∞ Lie group.) The theory of Lie groups pervades most of pure mathematics and large parts of theoretical physics, even it even touches parts of pure algebra where it would seem to a beginner that ideas from differential geometry ought to have no relevance. It is awesome, and you should study it.

The most basic example of a Lie group is $G = \mathrm{GL}(V)$ for a finite-dimensional vector space V, say of dimension $n > 0$. This is considered with the C^∞-structure that it acquires as an open subset of the vector space $\mathrm{Hom}(V, V)$. Concretely, if we choose linear coordinates on V then $\mathrm{Hom}(V, V)$ is just the space of $n\times n$ matrices (with "matrix entries" as its own linear coordinates) and $\mathrm{GL}(V)$ becomes the open subset $\mathrm{GL}_n(\mathbf{R})$ of invertible $n \times n$ matrices. The explicit universal algebraic formulas for matrix multiplication and inversion in terms of matrix entries show that matrix multiplication (and inversion) makes this a Lie group with respect to the given C^∞ structure. (This can also be proved withour recourse to coordinates, a task we leave to the interested reader.)

There are two rather interesting applications of Theorem 2.1 in this context. As our first application, we shall prove that in the definition of a Lie group, it is not necessary to assume that the inversion map is C^∞: this follows automatically from the hypothesis that the group law $m : G\times G \to G$ is C^∞! To see why this is, we require some lemmas. In what follows, note that we only use the C^∞ hypothesis on m and not for the inversion map.

Lemma 4.1. *The manifold G has the same dimension at all points.*

Though the conclusion of this lemma is valid for any connected topological premanifold whatsoever, we emphasize that there are interesting examples of disconnected Lie groups (such as $\mathrm{GL}_n(\mathbf{R})$).

Proof. If $e \in G$ is the identity element and $g \in G$ is an arbitrary point, the left multiplication map $\ell_g : G \to G$ defined by $x \mapsto gx$ is C^∞ (as it is the composite of the C^∞ slice embedding $G \to G\times G$ onto $\{g\} \times G$ and the group law m that we are assuming to be C^∞). Since $\ell_{g^{-1}}$ is a C^∞ inverse, it follows that ℓ_g is a C^∞ automorphism of G and it carries e to g, so its differential at the identity gives a linear isomorphism of $\mathrm{T}_e(G)$ onto $\mathrm{T}_g(G)$. Thus, the pointwise dimension at the arbitrary g agrees with the one at the identity e. ■

The preceding proof made crucial use of the C^∞ translation automorphisms of the manifold underlying G, and the proofs that follow will demonstrate the power of these maps; in effect, they make G a rather homogeneous space.

Lemma 4.2. *With respect to the natural isomorphism $\mathrm{T}_e(G)\oplus\mathrm{T}_e(G) \simeq \mathrm{T}_{(e,e)}(G\times G)$ as in Theorem 1.1, the differential*

$$\mathrm{d}m(e, e) : \mathrm{T}_e(G) \oplus \mathrm{T}_e(G) \to \mathrm{T}_e(G)$$

of the group law at the identity is the addition mapping $(v, v') \mapsto v + v'$.

Proof. The map $\mathrm{d}m(e, e)$ is a linear map between the two indicated spaces, and in general if $T : V \oplus V \to V$ is a linear map then since $T(v, v') = T((v, 0) + (0, v')) = T(v, 0) + T(0, v')$ we see that T is $(v, v') \mapsto v + v'$ if and only if the restrictions of T to the embedded subspaces

$V \rightrightarrows V \oplus V$ via $v \mapsto (v, 0)$ and $v \mapsto (0, v)$ are each the identity map. By the very construction of the isomorphism in Theorem 1.1 (applied to m and $(e, e) \in G \times G$), these embeddings of $\mathrm{T}_e(G)$ into the "factors" of $\mathrm{T}_{(e,e)}(G \times G)$ are given by the tangent maps of the slice inclusions $G \rightrightarrows G \times G$ defined by $R : g \mapsto (g, e)$ and $L : g \mapsto (e, g)$. Thus, by the Chain Rule, restricting $\mathrm{d}m(e, e)$ to these slices amounts to computing $\mathrm{d}m(e, e) \circ \mathrm{d}R(e) = \mathrm{d}(m \circ R)(e)$ and $\mathrm{d}m(e, e) \circ \mathrm{d}L(e) = \mathrm{d}(m \circ L)(e)$ as self-maps of $\mathrm{T}_e(G)$. We want these self-maps to be the identity. But this is clear because by the identity axioms for the group law we have $m \circ L$ and $m \circ R$ as self-maps of G are the identity maps (and each tangent map induced by the identity map on the manifold is the identity map on the corresponding tangent space). ■

We now claim that $m : G \times G \to G$ is a submersion at all points. At (e, e) this follows from the preceding lemma, as the addition map $V \oplus V \to V$ is clearly surjective for any vector space V. We now deduce the result at an arbitrary point $(g, g') \in G \times G$ by translations, as follows. We have $(g, g') = (\ell_g \times \rho_{g'})(e, e)$ where $\ell_g, \rho_{g'} : G \rightrightarrows G$ are the C^∞ automorphisms given by left multiplication by g ($x \mapsto gx$) and right multiplication by g' ($x \mapsto xg'$). By the group axioms,

$$m \circ (\ell_g \times \rho_{g'}) = \ell_g \circ \rho_{g'} \circ m$$

as C^∞ maps from $G \times G$ to G (this just says $(gx)(yg') = g(xy)g'$ for all $x, y \in G$). Applying the Chain Rule at (e, e), we get

$$\mathrm{d}m(g, g') \circ \mathrm{d}(\ell_g \times \rho_{g'})(e, e) = \mathrm{d}\ell_g(g') \circ \mathrm{d}\rho_{g'}(e) \circ \mathrm{d}m(e, e).$$

Since ℓ_g and $\rho_{g'}$ are C^∞ automorphisms, the right side is surjective because $\mathrm{d}m(e, e)$ is surjective, and so from the left side it follows that $\mathrm{d}m(g, g')$ is surjective because $\ell_g \times \rho_{g'}$ is a C^∞ automorphism (so its tangent maps are linear automorphisms). This completes the verification that m is a submersion, so its tangent maps have constant rank equal to the constant pointwise dimension of G, say d.

By Theorem 2.1 (!), we conclude that the fiber $m^{-1}(e)$ as a closed subset of $G \times G$ is a C^∞ submanifold of $G \times G$ with constant codimension d and hence constant dimension $2d - d = d$. The C^∞ projection $G \times G \to G$ onto the first factor restricts to a C^∞ map $m^{-1}(e) \to G$ (using the C^∞ submanifold structure on $m^{-1}(e)$). By group theory, $m^{-1}(e)$ is the set of points (g, g^{-1}) in $G \times G$, so the projection map $m^{-1}(e)$ is a bijection. We can say more:

Lemma 4.3. *The C^∞ projection map $p_1 : m^{-1}(e) \to G$ to the first factor is a C^∞ isomorphism with respect to the C^∞ submanifold structure on $m^{-1}(e)$.*

Proof. Since p_1 is bijective, if it is a local C^∞ isomorphism then its set-theoretic inverse is C^∞ and so p_1 is a C^∞ isomorphism. Our problem is therefore to show that p_1 is a local C^∞ isomorphism, and for this the inverse function theorem provides necessary and sufficient conditions: the tangent maps should be linear isomorphisms at all points of $m^{-1}(e)$. Let us first use translations to show that it suffices to check this linear isomorphism property at the point $(e, e) \in m^{-1}(e)$. For an arbitrary point $(g, g^{-1}) \in m^{-1}(e)$, the map $\ell_g \times \rho_{g^{-1}} : G \times G \to G \times G$ is a C^∞ map that carries the C^∞ submanifold $m^{-1}(e)$ back into itself (since $(gx)(x^{-1}g^{-1}) = e$ for all $x \in G$), so by the mapping property of embedded submanifolds it restricts to a C^∞ self-map of $m^{-1}(e)$. As such, this restricted map is a C^∞ automorphism φ_g of $m^{-1}(e)$ because $\ell_{g^{-1}} \times \rho_g$ induces a C^∞ inverse. The C^∞ automorphism φ_g of $m^{-1}(e)$ carries (e, e) to (g, g^{-1}), and by the definitions we see $p_1 \circ \varphi_g = \ell_g \circ p_1$ as C^∞ maps from $m^{-1}(e)$ to G. Applying the Chain Rule at $(e, e) \in m^{-1}(e)$ to both sides, we get

$$\mathrm{d}p_1(g, g^{-1}) \circ \mathrm{d}\varphi_g(e, e) = \mathrm{d}\ell_g(e) \circ \mathrm{d}p_1(e, e).$$

Since ℓ_g and φ_g are C^∞ automorphisms (of G and $m^{-1}(e)$ respectively), it follows that $\mathrm{d}p_1(g, g^{-1})$ is an isomorphism if and only if $\mathrm{d}p_1(e, e)$ is an isomorphism, as desired.

We now wish to compute $\mathrm{d}p_1(e,e) : \mathrm{T}_{(e,e)}(m^{-1}(e)) \to \mathrm{T}_e(G)$ to see that this is a linear isomorphism. The map p_1 is the composite of the C^∞ embedding $\iota : m^{-1}(e,e) \to G \times G$ and the projection $\pi_1 : G \times G \to G$ onto the first factor. Hence, by the Chain Rule,

$$\mathrm{d}p_1(e,e) = \mathrm{d}\pi_1(e,e) \circ \mathrm{d}\iota(e,e).$$

By Theorem 1.1, under the natural isomorphism

$$\mathrm{T}_{(e,e)}(G \times G) \simeq \mathrm{T}_e(G) \oplus \mathrm{T}_e(G)$$

the map $\mathrm{d}\pi_1(e,e)$ is projection onto the first factor and $\mathrm{d}m(e,e)$ is the addition map. By Theorem 1.2, the kernel of $\mathrm{d}m(e,e)$ is identified with the tangent space $\mathrm{T}_{(e,e)}(m^{-1}(e))$ of the fiber of m over $m(e,e) = e$, via the injection $\mathrm{d}\iota(e,e)$. Hence, we conclude that the composite of the inclusion $\mathrm{d}\iota(e,e)$ and the isomorphism $\mathrm{T}_{(e,e)}(G \times G) \simeq \mathrm{T}_e(G) \oplus \mathrm{T}_e(G)$ identifies $\mathrm{T}_{(e,e)}(m^{-1}(e))$ with the subspace of points $(v, -v)$ in $\mathrm{T}_e(G) \oplus \mathrm{T}_e(G)$. Our problem is therefore reduced to the obvious fact that if V is a vector space then the projection map $V \oplus V \to V$ onto the first factor restricts to a linear isomorphism from the subspace of points $(v, -v)$ onto the target V. ■

By the preceding lemma, the set-theoretic inverse map $G \to m^{-1}(e)$ given by $g \mapsto (g, g^{-1})$ is a C^∞ map with respect to the C^∞-structures on source and target. Composing this with the C^∞ inclusion map $m^{-1}(e) \to G \times G$ and the C^∞ second projection $G \times G \to G$ yields a C^∞ composite map $G \to G$ that is $g \mapsto g^{-1}$! Hence, we have proved that inversion on G is indeed automatically a C^∞ map (and even necessarily a C^∞ automorphism, as it is its own inverse map).

We now turn to a more striking application of Theorem 2.1 in the context of Lie groups:

Theorem 4.4. *Let $f : G \to G'$ be a group homomorphism between Lie groups, and assume that f is a C^∞ map. The kernel $K = \ker f$ is a closed C^∞ submanifold of G, and in particular has a natural structure of Lie group.*

It is a remarkable theorem in Lie theory that any continuous homomorphism between Lie groups is automatically C^∞, but we shall not consider this issue here. As a simple application of the theorem, note that the determinant map $\det : \mathrm{GL}_n(\mathbf{R}) \to \mathbf{R}^\times$ is a C^∞ group homomorphism (via the explicit polynomial formula for matrix determinants), and so the theorem implies that the closed subgroup $\mathrm{SL}_n(\mathbf{R})$ of matrices with determinant 1 is a C^∞ submanifold. This can also be proved "by hand" via the implicit function theorem and some considerations with translations and the formula for determinants (try!), but the theorem shows that such explicit considerations are not necessary. Note also that even if G and G' in the theorem are connected, the kernel K can have infinitely many connected components (and so this gives motivation for avoiding connectivity conditions in the foundations of the theory). For example, if $G = \mathbf{C}$ (considered as a Lie group via addition) and $G' = \mathbf{C}^\times$ (considered as a Lie group via multiplication), then the mapping

$$z = x + \sqrt{-1} \cdot y \mapsto e^z = e^x \cos(y) + \sqrt{-1} \cdot e^x \sin(y)$$

is a C^∞ group homomorphism (that is independent of the choice of $\sqrt{-1}$!) and its kernel is the infinite discrete subgroup $2\pi\sqrt{-1} \cdot \mathbf{Z}$ inside of $\mathbf{C}$.

Proof. Let $e \in G$ and $e' \in G'$ be the respective identity elements. The kernel is $f^{-1}(e')$, so by Theorem 2.1 it suffices to show that f has constant Jacobian rank at all points. Choose $g \in G$. The rank of the tangent map $\mathrm{d}f(g) : \mathrm{T}_g(G) \to \mathrm{T}_{f(g)}(G')$ is the difference of the dimensions of $\mathrm{T}_g(G)$ and $\ker \mathrm{d}f(g)$, so since $\dim \mathrm{T}_g(G)$ is independent of g (Lemma 4.1) it is equivalent to show that $\dim \ker \mathrm{d}f(g)$ is independent of g. We shall again use translations to show that this dimension is equal to the value in the case $g = e$. For an arbitrary $g \in G$, since f is a group homomorphism we have $f \circ \ell_g = \ell_{f(g)} \circ f$ (this says $f(gx) = f(g)f(x)$ for any $x \in G$), so using the Chain Rule at e and the equality $f(e) = e'$ we get

$$\mathrm{d}f(g) \circ \mathrm{d}\ell_g(e) = \mathrm{d}\ell_{f(g)}(e') \circ \mathrm{d}f(e).$$

The maps $\mathrm{d}\ell_g(e)$ and $\mathrm{d}\ell_{f(g)}(e')$ are linear isomorphisms (as left translations are C^∞ automorphisms), so it follows that $\mathrm{d}f(g)$ and $\mathrm{d}f(e)$ must have kernels with the same dimension. (Concretely, the linear isomorphism $\mathrm{d}\ell_g(e) : \mathrm{T}_e(G) \simeq \mathrm{T}_g(G)$ must carry $\ker \mathrm{d}f(e)$ isomorphically over to $\ker \mathrm{d}f(g)$.) ■